RESEARCH CENTER

FIELD

Activity Report 2013

Section New Results

Edition: 2014-03-19

ABS Project-Team

6. New Results

6.1. Modeling Interfaces and Contacts

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

The work undertaken in this vein in 2013 will be finalized in 2014.

6.2. Modeling Macro-molecular Assemblies

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

6.2.1. Connectivity Inference in Mass Spectrometry based Structure Determination

Participants: Frédéric Cazals, Deepesh Agarwal.

In collaboration with J. Araujo, and C. Caillouet, and D. Coudert, and S. Pérennes, from the COATI project-team (Inria-CNRS).

In [14], we consider the following MINIMUM CONNECTIVITY INFERENCE problem (MCI), which arises in structural biology: given vertex sets $V_i \subseteq V, i \in I$, find the graph $G = (V, E)$ minimizing the size of the edge set E, such that the sub-graph of G induced by each V_i is connected. This problem arises in structural biology, when one aims at finding the pairwise contacts between the proteins of a protein assembly, given the lists of proteins involved in sub-complexes. We present four contributions.

First, using a reduction of the set cover problem, we establish that MCI is APX-hard. Second, we show how to solve the problem to optimality using a mixed integer linear programming formulation (`MILP`). Third, we develop a greedy algorithm based on union-find data structures (`Greedy`), yielding a $2(\log_2 |V| + \log_2 \kappa)$-approximation, with κ the maximum number of subsets V_i a vertex belongs to. Fourth, application-wise, we use the MILP and the greedy heuristic to solve the aforementioned connectivity inference problem in structural biology. We show that the solutions of `MILP` and `Greedy` are more parsimonious than those reported by the algorithm initially developed in biophysics, which are not qualified in terms of optimality. Since MILP outputs a set of optimal solutions, we introduce the notion of *consensus solution*. Using assemblies whose pairwise contacts are known exhaustively, we show an almost perfect agreement between the contacts predicted by our algorithms and the experimentally determined ones, especially for consensus solutions.

6.3. Algorithmic Foundations

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

6.3.1. Greedy Geometric Algorithms for Collection of Balls, with Applications to Geometric Approximation and Molecular Coarse-Graining

Participants: Frédéric Cazals, Tom Dreyfus.

In collaboration with S. Sachdeva (Princeton University, USA), and N. Shah (Carnegie Mellon University, USA).

Choosing balls to best approximate a 3D object is a non trivial problem. To answer it, in [18], we first address the *inner approximation* problem, which consists of approximating an object $\mathcal{F}_{\mathcal{O}}$ defined by a union of n balls with $k < n$ balls defining a region $\mathcal{F}_{\mathcal{S}} \subset \mathcal{F}_{\mathcal{O}}$. This solution is further used to construct an *outer approximation* enclosing the initial shape, and an *interpolated approximation* sandwiched between the inner and outer approximations.

The inner approximation problem is reduced to a geometric generalization of weighted max k-cover, solved with the greedy strategy which achieves the classical $1 - 1/e$ lower bound. The outer approximation is reduced to exploiting the partition of the boundary of $\mathcal{F}_{\mathcal{O}}$ by the Apollonius Voronoi diagram of the balls defining the inner approximation.

Implementation-wise, we present robust software incorporating the calculation of the exact Delaunay triangulation of points with degree two algebraic coordinates, of the exact medial axis of a union of balls, and of a certified estimate of the volume of a union of balls. Application-wise, we exhibit accurate coarse-grain molecular models using a number of balls 20 times smaller than the number of atoms, a key requirement to simulate crowded cellular environments.

6.3.2. Towards Morse Theory for Point Cloud Data

Participants: Frédéric Cazals, Christine Roth.

In collaboration with C. Robert (IBPC / CNRS, Paris, France), and C. Mueller (ETH, Zurich).

Morse theory provides a powerful framework to study the topology of a manifold from a function defined on it, but discrete constructions have remained elusive due to the difficulty of translating smooth concepts to the discrete setting.

Consider the problem of approximating the Morse-Smale (MS) complex of a Morse function from a point cloud and an associated nearest neighbor graph (NNG). While following the constructive proof of the Morse homology theorem, we present novel concepts for critical points of any index, and the associated Morse-Smale diagram [19].

Our framework has three key advantages. First, it requires elementary data structures and operations, and is thus suitable for high-dimensional data processing. Second, it is gradient free, which makes it suitable to investigate functions whose gradient is unknown or expensive to compute. Third, in case of under-sampling and even if the exact (unknown) MS diagram is not found, the output conveys information in terms of ambiguous flow, and the Morse theoretical version of topological persistence, which consists in canceling critical points by flow reversal, applies.
On the experimental side, we present a comprehensive analysis of a large panel of bi-variate and tri-variate Morse functions whose Morse-Smale diagrams are known perfectly, and show that these diagrams are recovered perfectly.

In a broader perspective, we see our framework as a first step to study complex dynamical systems from mere samplings consisting of point clouds.

6.4. Misc

Computational Biology, Biomedicine.

6.4.1. Book

Participant: Frédéric Cazals.

Edited in collaboration with P. Kornprobst, from the Neuromathcomp project-team.

Biology and biomedicine currently undergo spectacular progresses due to a synergy between technological advances and inputs from physics, chemistry, mathematics, statistics and computer science. The goal of the book [15] is to evidence this synergy, by describing selected developments in the following fields: bioinformatics, biomedicine, neuroscience.

This book is unique in two respects. First, by the variety and scales of systems studied. Second, by its presentation, as each chapter presents the biological or medical context, follows up with mathematical or algorithmic developments triggered by a specific problem, and concludes with one or two success stories, namely new insights gained thanks to these methodological developments. It also highlights some unsolved and outstanding theoretical questions, with potentially high impact on these disciplines.

Two communities will be particularly interested. The first one is the vast community of applied mathematicians and computer scientists, whose interests should be captured by the added value generated by the application of advanced concepts and algorithms to challenging biological or medical problems. The second is the equally vast community of biologists. Whether scientists or engineers, they will find in this book a clear and self-contained account of concepts and techniques from mathematics and computer science, together with success stories on their favorite systems. The variety of systems described will act as an eye opener on a panoply of complementary conceptual tools. Practically, the resources listed at the end of each chapter (databases, software) will prove invaluable to get started on a specific topic.

ABSTRACTION Project-Team

6. New Results

6.1. Analysis of Biological Pathways

We have improved our framework to design and analyze biological networks in KAPPA. This framework focuses 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].

Keywords: Graph rewriting, Single Push-Out semantics.

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 [16], 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. Causality Analysis

We use causal analysis so as to extract minimal concurrent scenarios that lead to the activation of given events.

6.1.2.1. Implementation

Participant: Jérôme Feret.

Keywords: Causality, Counter-examples, Compression.

This year, we have re-implemented in OPENKAPPA the strong compression method that is described in [48]. The new implementation is very efficient, it has been used to extract minimal scenarios from traces of several hundred of thousands causally related events, that were generated during the simulation of a model of the WnT signaling pathway.

6.1.2.2. Framework

Participant: Jonathan Hayman.

Keywords: Abstraction, Causality, Compression.

Standard notions of independence of rule applications fail to provide adequately concise causal histories, leading to the earlier formulation of strong and weak forms of trajectory compression for Kappa. In [15], we give a simple categorical account of how forms of compression can be uniformly obtained. This generalisation also describes a way for the user to specify their own levels of compression between weak and strong, which we call filtered compression. This is based on the idea of the user specifying the part of the type graph that represents the the structure which the compression technique should track through the trace.

6.1.3. Model Reduction

Participants: Ferdinanda Camporesi, Jérôme Feret, Jonathan Hayman.

Keywords: Context-sensitivity, Differential semantics, Model reduction.

Rule-based modeling allows very compact descriptions of protein-protein interaction networks. However, combinatorial complexity increases again when one attempts to describe formally the behaviour of the networks, which motivates the use of abstractions to make these models more coarse-grained. Context-insensitive abstractions of the intrinsic flow of information among the sites of chemical complexes through the rules have been proposed to infer sound coarse-graining, providing an efficient way to find macro-variables and the corresponding reduced models.

In [12], we propose a framework to allow the tuning of the context-sensitivity of the information flow analyses and show how these finer analyses can be used to find fewer macro-variables and smaller reduced differential models.

6.2. Andromeda: Accurate and Scalable Security Analysis of Web Applications

Participants: Omer Tripp [Tel Aviv University, Israël], Marco Pistola [University of Washington, Seattle, USA], Patrick Cousot, Radhia Cousot, Salvatore Guarnieri.

Keywords: Abstract interpretation, Security, Web.

Security auditing of industry-scale software systems mandates automation. Static taint analysis enables deep and exhaustive tracking of suspicious data flows for detection of potential leakage and integrity violations, such as cross-site scripting (XSS), SQL injection (SQLi) and log forging. Research in this area has taken two directions: program slicing and type systems. Both of these approaches suffer from a high rate of false findings, which limits the usability of analysis tools based on these techniques. Attempts to reduce the number of false findings have resulted in analyses that are either (i) unsound, suffering from the dual problem of false negatives, or (ii) too expensive due to their high precision, thereby failing to scale to real-world applications.

In [21], we investigate a novel approach for enabling precise yet scalable static taint analysis. The key observation informing our approach is that taint analysis is a demand-driven problem, which enables lazy computation of vulnerable information flows, instead of eagerly computing a complete data-flow solution, which is the reason for the traditional dichotomy between scalability and precision. We have implemented our approach in Andromeda, an analysis tool that computes data-flow propagations on demand, in an efficient and accurate manner, and additionally features incremental analysis capabilities. Andromeda is currently in use in a commercial product. It supports applications written in Java, .NET and JavaScript. Our extensive evaluation of Andromeda on a suite of 16 production-level benchmarks shows Andromeda to achieve high accuracy and compare favorably to a state-of-the-art tool that trades soundness for precision.

6.3. Backward analysis

6.3.1. Automatic Inference of Necessary Preconditions

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

Keywords: Abstract interpretation, Backward analysis, Static analysis, Necessary condition inference.

In [14], 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%.

6.3.2. Under-approximations to infer sufficient program conditions

Participant: Antoine Miné.

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

In [9] we discuss the automatic inference of sufficient preconditions by abstract interpretation and sketch the construction of an under-approximating backward analysis. We focus on numeric properties of variables and revisit three classic numeric abstract domains: intervals, octagons, and polyhedra, with new under-approximating backward transfer functions, including the support for non-deterministic expressions, as well as lower widenings to handle loops. We show that effective under-approximation is possible natively in these domains without necessarily resorting to disjunctive completion nor domain complementation. 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. We built a proof-of-concept prototype implementation based on the APRON numeric domain library and experimented it on simple examples (the prototype is available for download and usable on-line at http://www.di.ens.fr/~mine/banal).

6.4. Bisimulation metrics

6.4.1. Bisimulation for MDP through Families of Functional Expressions

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

Keywords: Markov decision processes, Bisimulation, Metrics.

We have transfered a notion of quantitative bisimilarity for labelled Markov processes [54] to Markov decision processes with continuous state spaces. This notion takes the form of a pseudometric on the system states, cast in terms of the equivalence of a family of functional expressions evaluated on those states and interpreted as a real-valued modal logic. Our proof amounts to a slight modification of previous techniques [61], [60] used to prove equivalence with a fixed-point pseudometric on the state-space of a labelled Markov process and making heavy use of the Kantorovich probability metric. Indeed, we again demonstrate equivalence with a fixed-point pseudometric defined on Markov decision processes [57]; what is novel is that we recast this proof in terms of integral probability metrics [59] defined through the family of functional expressions, shifting emphasis back to properties of such families. The hope is that a judicious choice of family might lead to something more computationally tractable than bisimilarity whilst maintaining its pleasing theoretical guarantees. Moreover, we use a trick from descriptive set theory to extend our results to MDPs with bounded measurable reward functions, dropping a previous continuity constraint on rewards and Markov kernels.

This work is under submission.

6.4.2. Bisimulation Metrics are Optimal Value Functions

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

Keywords: Markov decision processes, Bisimulation, Metrics.

We have proved that a behavioural pseudometric defined on the state space of a given Markov decision process and whose kernel is stochastic bisimilarity [57] can be expressed as the optimal value function of another Markov decision process. Furthermore, this latter process can be interpreted as an optimal coupling of two copies of the original model.

This work is under submission.

6.5. 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].

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

In [18] and [19] 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.

6.6. A Galois Connection Calculus for Abstract Interpretation

Participants: Patrick Cousot, Radhia Cousot.

Keywords: Abstract interpretation, Galois connection.

In [10], we introduce a Galois connection calculus for language independent specification of abstract interpretations used in programming language semantics, formal verification, and static analysis. This Galois connection calculus and its type system are typed by abstract interpretation.

6.7. Mechanically Verifying a Shape Analysis

Participant: Arnaud Spiwack.

Keywords: Program verification, Abstract interpretation, Static analysis, Shape analysis, Coq.

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

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

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

6.8. Modular Construction of Shape-Numeric Analyzers

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

Keywords: Abstract interpretation, Memory abstraction, Shape abstract domains.

In [13], we discuss the modular construction of memory abstract domains.

The aim of static analysis is to infer invariants about programs that are precise enough to establish semantic properties, such as the absence of run-time errors. Broadly speaking, there are two major branches of static analysis for imperative programs. Pointer and *shape* analyses focus on inferring properties of pointers, dynamically-allocated memory, and recursive data structures, while *numeric* analyses seek to derive invariants on numeric values. Although simultaneous inference of shape-numeric invariants is often needed, this case is especially challenging and is not particularly well explored. Notably, simultaneous shape-numeric inference raises complex issues in the design of the static analyzer itself.

In this paper, we study the construction of such shape-numeric, static analyzers. We set up an abstract interpretation framework that allows us to reason about simultaneous shape-numeric properties by combining shape and numeric abstractions into a modular, expressive abstract domain. Such a modular structure is highly desirable to make its formalization and implementation easier to do and get correct. To achieve this, we choose a concrete semantics that can be abstracted step-by-step, while preserving a high level of expressiveness. The structure of abstract operations (i.e., transfer, join, and comparison) follows the structure of this semantics. The advantage of this construction is to divide the analyzer in modules and functors that implement abstractions of distinct features.

6.9. Reduced Product Combination of Abstract Domains for Shapes

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

Keywords: Abstract interpretation, Memory abstraction, Shape abstract domains.

In [20], we discuss the construction of shape abstract domains by reduced product.

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.

6.10. Relational Thread-Modular Static Value Analysis

Participant: Antoine Miné.

Keywords: Abstract interpretation, Concurrency, Embedded software, Rely-guarantee methods, Run-time errors, Safety.

We study in [17] thread-modular static analysis by abstract interpretation to infer the values of variables in concurrent programs. We show how to go beyond the state of the art and increase an analysis precision by adding the ability to infer some relational and history-sensitive properties of thread interferences. The fundamental basis of this work is the formalization by abstract interpretation of a rely-guarantee concrete semantics which is thread-modular, constructive, and complete for safety properties. We then show that previous analyses based on non-relational interferences can be retrieved as coarse computable abstractions of this semantics; additionally, we present novel abstraction examples exploiting our ability to reason more precisely about interferences, including domains to infer relational lock invariants and the monotonicity of counters. Our method and domains have been implemented in the ASTRÉEA static analyzer (5.3) that checks for run-time errors in embedded concurrent C programs, where they enabled a significant reduction of the number of false alarms.

6.11. Static Analyzers on the Cloud

Participants: Michael Barnett [Microsoft Research, Redmond, USA], Mehdi Bouaziz, Francesco Logozzo [Microsoft Research, Redmond, USA], Manuel Fähndrich [Microsoft Research, Redmond, USA].

A cloud-based static analyzer runs as service. Clients issue analysis requests through the local network or over the internet. The analysis takes advantage of the large computation resources offered by the cloud: the underlying infrastructure ensures scaling and unlimited storage. Cloud-based analyzers may relax performance-precision trade-offs usually associated with desktop-based analyzers. More cores enable more precise and responsive analyses. More storage enables perfect caching of the analysis results, shareable among different clients, and queryable off-line. To realize these advantages, cloud-based analyzers need to be architected differently than desktop ones. In [11], we describe our ongoing effort of moving a desktop analyzer, Clousot, into a cloud-based one, Cloudot.

6.12. Termination

We have explored the analysis of program termination and the inference of sufficient conditions to ensure the definite termination of programs using abstract interpretation techniques. Following [40], we employ a backward analysis over an abstract domain of ranking functions.

6.12.1. Abstract Domain of Segmented Ranking Functions

Participant: Caterina Urban.

We present in [24] and [23] a parameterized abstract domain that infers sufficient conditions for program termination by automatically synthesizing piecewise-defined ranking functions over natural numbers. The analysis uses over-approximations but we prove its soundness, meaning that all program executions respecting these sufficient conditions are indeed terminating. The abstract domain is parameterized by a numerical abstract domain for environments and a numerical abstract domain for functions. This parameterization allows to easily tune the trade-off between precision and cost of the analysis. We describe an instantiation of this generic domain with intervals and affine functions. We define all abstract operators, including widening to ensure convergence. To experiment with this domain, we have implemented a research prototype static analyzer FuncTion (5.6) that yielded interesting preliminary results.

6.12.2. Abstract Domain to Infer Ordinal-Valued Ranking Functions

Participants: Caterina Urban, Antoine Miné.

We observed that, in some important cases (such as programs with unbounded non-determinism), there does not exist any ranking function over natural numbers. In [22] and [29] we propose a new abstract domain to automatically infer ranking functions over ordinals. We extended the domain of piecewise-defined natural-valued ranking functions introduced in the previous section to polynomials in ω, where the polynomial coefficients are natural-valued functions of the program variables. The abstract domain is parametric in the choice of the maximum degree of the polynomial, and the types of functions used as polynomial coefficients. We have enriched the FuncTion prototype analyzer (5.6) with an instantiation of our domain using affine functions as polynomial coefficients. We successfully analyzed small but intricate examples that are out of the reach of existing methods. To our knowledge this is the first abstract domain able to reason about ordinals. Handling ordinals leads to a powerful approach for proving termination of imperative programs, which in particular subsumes existing techniques based on lexicographic ranking functions.

ACES Project-Team

5. New Results

5.1. Self-describing objects

Participants: Michel Banâtre, Nebil Ben Mabrouk, Paul Couderc [contact], Yann Glouche, Arnab Sinha.

Coupled objects enable basic integrity checking for physical objects, and use cases were demonstrated for security and logistics applications. In these applications, high reliability in the RFID reading infrastructure is assumed for the system to work. This suggest another idea for coupled objects: using control data structures distributed over the physical objects in order to improve the reliability of RFID reading protocols. This is the purpose of the Pervasive_RFID project, in collaboration with the IETR which is described in more details below 7.1.2 .

Another development in the line of the coupled objects principles are self-describing objects. While previous works enabled integrity checking over a set of physical objects, these mechanisms were limited in two aspects: expressiveness and autonomy. More precisely, coupled objects support the detection of special conditions (such as a missing element), but not the characterization of these conditions (such as describing the problem, identifying the missing element). Moreover, this compromises the autonomous feature of coupled objects, which would depend on external systems for analyzing these special conditions. Self-describing objects are an attempt to overcome these limitations, and to broaden the application perspectives of autonomous RFID systems.

The principle is to implement distributed data structure over a set of RFID tags, enabling a complex object (made of various parts) or a set of objects belonging to a given logical group to "self-describe" itself and the relation between the various physical elements. Some applications examples includes waste management, assembling and repair assistance, prevention of hazards in situations where various products / materials are combined etc. The key property of self-describing objects is, like for coupled objects, that the vital data are self-"hosted" by the physical element themselves (typically in RFID chips), not an external infrastructure like most RFID systems. This property provides the same advantages as in coupled objects, namely high scalability, easy deployment (no interoperability dependence/interference), and limited risk for privacy.

However, given the extreme storage limitation of RFID chips, designing such systems is difficult:

- data structures must be very frugal in terms of space requirements, both for the structure and for the coding.
- Data structures must be robust and able to survive missing or corrupted elements if we want to ensure the self-describing property for a damaged or incorrect object.

An application of self-describing objects has been proposed in for waste management, in the context of the bin that think project 7.1.1 . A generic graph structure applicable to RFID systems for supporting self-describing objects is proposed in Arnab Sinha's thesis document (to be defended in April 2014).

5.2. Pervasive support for Smart Homes

Participants: Andrey Boytsov, Michele Dominici, Bastien Pietropaoli, Sylvain Roche, Frederic Weis [contact].

A smart home is a residence equipped with information-and-communication-technology (ICT) devices conceived to collaborate in order to anticipate and respond to the needs of the occupants, working to promote their comfort, convenience, security and entertainment while preserving their natural interaction with the environment.

The idea of using the Ubiquitous Computing paradigm in the smart home domain is not new. However, the state-of-the-art solutions only partially adhere to its principles. Often the adopted approach consists in a heavy deployment of sensor nodes, which continuously send a lot of data to a central elaboration unit, in charge of the difficult task of extrapolating meaningful information using complex techniques. This is a *logical approach*. ACES proposed instead the adoption of a *physical approach*, in which the information is spread in the environment, carried by the entities themselves, and the elaboration is directly executed by these entities "inside" the physical space. This allows performing meaningful exchanges of data that will thereafter need a less complicate processing compared to the current solutions. The result is a smart home that can, in an easier and better way, integrate the context in its functioning and thus seamlessly deliver more useful and effective user services. Our contribution aims at implementing the physical approach in a domestic environment, showing a solution for improving both comfort and energy savings.

5.2.1. A multi-level context computing architecture

Computing context is a major subject of interest in smart spaces such as smart homes. Contextual data are necessary for services to adapt themselves to the context and to be as efficient as possible. Contextual data may be obtained via augmented appliances capable of communicating their state and a bunch of sensors. It becomes more and more real with the development of the Internet of Things. Unfortunately, the gathered data are not always directly usable to understand what is going on and to build services on them. In order to address this issue, we studied a multi-level context computing architecture divided in four layers:

- *Exploitation layer*: the highest layer, it exploits con- textual data to provide adapted services
- *Context and situation identification layer*: this is what analyzes ongoing situations and potentially predicts future situations
- *Perception layer*: it offers a first layer of abstraction for small pieces of context independent of deployed sensors
- *Sensing layer*: it mainly consists of the data gathered by sensors

In this architecture, every layer is based on the results of its underlying layers. In 2013, we studied several methods that enable the building of such levels of abstractions (see figure 2). The first level of abstraction coming to mind when describing what people are doing in a Home is high level abstractions such as "cooking". Those activities are then the highest level abstraction we want our system to be able to identify.

We proposed to use plan recognition algorithms to analyze sequences of actions and thus predict future actions of users. It is, in our case, adapted to identify ongoing activities and predict future ones. There exist different plan recognition algorithms. However, one interested us particularly, PHATT introduced by Goldman, Geib and Miller. In order to understand how PHATT is working, it is important to understand the hierarchical task network (HTN) planning problem which is "inverted" by the algorithm to perform plan recognition. It consists in automatically generating a plan starting from a set of tasks to execute and some constraints. In our case, we are able to predict future situations depending of the previously observed situations. To give an example, if we want to predict that the situation dinner will occur soon, it is sufficient to have observed situations such as cooking and/or setting the table. The performances of PHATT have been evaluated by Andrey Andrey Boytstov and Frédéric Weis. These results will be published in 2014.

5.2.2. Propagation of BFT

Context-aware applications have to sense the environment in order to adapt themselves and provide with contextual services. This is the case of Smart Homes equipped with sensors and augmented appliances. However, sensors can be numerous, heterogeneous and unreliable. Thus the data fusion is complex and requires a solid theory to handle those problems. For this purpose, we adopted the belief functions theory (BFT). The aim of the data fusion, in our case, is to compute small pieces of context we call context attributes. Those context attributes are diverse and could be for example the presence in a room, the number of people in a room or even that someone may be sleeping in a room. Since the BFT requires a substantial amount of computations, we proposed to reduce as much as possible the number of evidence required to compute a context attribute. Moreover, the number of possible worlds, *i.e.* the number of possible states for a context attribute, is also an

../../../../projets/aces/IMG/architecture.png

Figure 2. Multi-level context computing architecture

important source of computation. Thus, reducing the number of possible worlds we are working on is also important.

It is especially problematic when working on embedded systems, which may be the case when trying to observe context in smart homes. Thus, with this objective in mind, we observed that some context attributes could be used to compute others. By doing this, the number of gathered and combined evidence for each context attribute could be drastically reduced. This principle is illustrated by Figure 3 : the sets of possible worlds for "Presence" and "Posture" are seen as subsets of "Sleeping". So we proposed and implemented a method to propagate BFT through a set of possible states for a context attribute.

../../../../projets/aces/IMG/propagation.png

Figure 3. Propagation of Belief Functions Theories

5.2.3. Definition of virtual sensors

In our multi-level architecture, the sensor measures may be imperfect for multiple reasons. The most annoying reasons when deploying a system are biases and noisy measures. It requires fine tuning each type the system is deployed in a new environment. In order to prevent from doing this work again and again at levels where models are hard to build, we proposed to add a new sublayer to the sensing layer (see Figure 2): virtual sensors. Instead of modifying high level models, we created sensor abstractions such as motion sensor, sound sensor, temperature sensor, etc. It is particularly convenient when working with typed data such as temperature or sound level. It is possible to use different brands of sensors for sensors of the same type. Thus, those sensors, even if they are measuring the same physical event, can return very different data due to their range, sensibility, voltage, etc. By creating abstraction of sensors, it is possible to build models directly from typed data simplifying even more the building of models as those data have are understandable by humans. Those virtual sensors are built very simply from common heuristics and can be used for ias and noise compensation, Data aggregation and Meta-data generation.

It is also possible in these virtual sensors to implement fault and failure detection mechanisms using the BFT. It enables the detection of fault in the case of sensors of the same type. At higher level, those mechanisms will detect inconsistency between sensors of different types which is not of the same utility. Thus, those virtual virtual sensors, without disabling any features in our architecture, bring more stability for our models. Moreover, by keeping the virtual sensors very simple, they are easy to adapt and tune in a new environment and the overhead in terms of computation is reduced to the minimum and does not really impact the global system performance. Finally, the fine tuning part is always reduced to this level of our architecture and nothing else has to be changed when we move the system from one environment to another.

ADAM Project-Team

6. New Results

6.1. Self-Adaptive Software Systems

Participants: Russel Nzekwa, Romain Rouvoy [correspondant], Lionel Seinturier.

The design of self-adaptive and autonomic software systems raises many challenges. In his PhD thesis, Russel Nzekwa [12] proposes a new result with the CORONA framework that enables to build flexible autonomic systems. CORONA relies on an architectural description language which reifies the structure of the control system architecture. CORONA enables the flexible integration of non-functional-properties during the design of autonomic systems. It also provides tools for checking conflicts in the architecture of autonomic systems. Finally, the traceability between the design and the runtime implementation is carried out through the code generation of skeletons from architectural descriptions of control systems. The work on CORONA goes toward the long term objective of setting up an integrated design and programming solution for self-adaptive systems, where feedback control loops play the central role as first class elements.

6.2. Energy Management in Software Systems

Participants: Rémi Druilhe, Laurence Duchien, Lionel Seinturier [correspondant].

Energy management and saving is a concern that spans the entire domain of information and communication technologies and sciences. Recently it has been recognized that to improve its efficiency, energy has to be managed, not only at the hardware level, but also at the level of software systems, especially in distributed environments. In his PhD thesis, Rémi Druilhe [11] proposes a new result with the HOMENAP system for networked digital home environments. This work is the result of a collaboration with Orange Labs. HOMENAP takes into account three main properties: heterogeneity, dynamicity and quality of service. HOMENAP proposes an autonomic decision-making system to deal with the placement of digital services on networked devices. Based on the observation of relevant events, the system takes the decision to modify the distribution of digital services on devices in order to preserve a defined tradeoff between energy efficiency and quality of service. HOMENAP participates to the long term objective of dealing with energy as a main steering factor for self-optimizing software systems.

6.3. Automated Software Repair

Participant: Martin Monperrus [correspondant].

Automated software repair aims at assisting developers in order to improve the quality of software systems, for example by recommending some repair actions to fix bugs. In [15], we present some major results in this direction by mining fix transactions of existing software repositories. From the empirical study of 14 software repositories containing 89,993 versioning transactions, we show that we can learn a probability distribution of repair actions. We show that certain distributions over repair actions can result in an infinite time (in average) to find a repair shape while other fine-tuned distributions enable to find a repair shape in hundreds of repair attempts. We now aim at going beyond this empirical study and theoretical analysis by exploring how to use this learned knowledge for new software systems.

ALEA Project-Team

6. New Results

6.1. Sparsity-Promoting Bayesian Dynamic Linear Models

Sparsity-promoting priors have become increasingly popular over recent years due to an increased number of regression and classification applications involving a large number of predictors. In time series applications where observations are collected over time, it is often unrealistic to assume that the underlying sparsity pattern is fixed. We propose an original class of flexible Bayesian linear models for dynamic sparsity modelling. The proposed class of models expands upon the existing Bayesian literature on sparse regression using generalized multivariate hyperbolic distributions. The properties of the models are explored through both analytic results and simulation studies. We demonstrate the model on a financial application where it is shown that it accurately represents the patterns seen in the analysis of stock and derivative data, and is able to detect major events by filtering an artificial portfolio of assets.

6.2. Evolutionnary algorithms and genetic programming

The regularity of a signal can be numerically expressed using Hölder exponents, which characterize the singular structures a signal contains. In particular, within the domains of image processing and image understanding, regularity-based analysis can be used to describe local image shape and appearance. However, estimating the Hölder exponent is not a trivial task, and current methods tend to be computationally slow and complex. This work presents an approach to automatically synthesize estimators of the pointwise Hölder exponent for digital images. This task is formulated as an optimization problem and Genetic Programming (GP) is used to search for operators that can approximate a traditional estimator, the oscillations method. Experimental results show that GP can generate estimators that achieve a low error and a high correlation with the ground truth estimation. Furthermore, most of the GP estimators are faster than traditional approaches, in some cases their runtime is orders of magnitude smaller. This result allowed us to implement a real-time estimation of the Hölder exponent on a live video signal, the first such implementation in current literature. Moreover, the evolved estimators are used to generate local descriptors of salient image regions, a task for which a stable and robust matching is achieved, comparable with state-of-the-art methods. In conclusion, the evolved estimators produced by GP could help expand the application domain of Hölder regularity within the fields of image analysis and signal processing.

One of the main open problems within Genetic Programming (GP) is to meaningfully characterize the difficulty (or hardness) of a problem. The general goal is to develop predictive tools that can allow us to identify how difficult a problem is for a GP system to solve. On this topic, we identify and compare two main approaches that address this question. We denote the first group of methods as Evolvability Indicators (EI), which are measures that attempt to capture how amendable the fitness landscape is to a GP search. The best examples of current EIs are the Fitness Distance Correlation (FDC) and the Negative Slope Coefficient (NSC). The second, more recent, group of methods are what we call Predictors of Expected Performance (PEP), which are predictive models that take as input a set of descriptive attributes of a particular problem and produce as output the expected performance of a GP system. The experimental work presented here compares an EI, the NSC, and a PEP model for a GP system applied to data classification. Results suggest that the EI fails at measuring problem difficulty expressed by the performance of the GP classifiers, an unexpected result. On the other hand, the PEP models show a very high correlation with the actual performance of the GP system. It appears that while an EI can correctly estimate the difficulty of a given search, as shown by previous research on this topic, it does not necessarily capture the difficulty of the underlying problem that GP is intended to solve. Conversely, while the PEP models treat the GP system as a computational black-box, they can still provide accurate performance predictions.

Another research area is to predict the alertness of an individual by analyzing the brain activity through electroencephalographic data (EEG) captured with 58 electrodes. Alertness is characterized here as a binary variable that can be in a "normal" or "relaxed" state. We collected data from 44 subjects before and after a relaxation practice, giving a total of 88 records. After a pre-processing step and data validation, we analyzed each record and discriminate the alertness states using our proposed "slope criterion". Afterwards, several common methods for supervised classification (k nearest neighbors, decision trees (CART), random forests, PLS and discriminant sparse PLS) were applied as predictors for the state of alertness of each subject. The proposed "slope criterion" was further refined using a genetic algorithm to select the most important EEG electrodes in terms of classification accuracy. Results shown that the proposed strategy derives accurate predictive models of alertness.

6.3. Moderate Deviations for Mean Field Particle Models

Our team is interested with moderate deviation principles of a general class of mean field type interacting particle models. We discuss functional moderate deviations of the occupation measures for both the strong - topology on the space of finite and bounded measures as well as for the corresponding stochastic processes on some class of functions equipped with the uniform topology. Our approach is based on an original semigroup analysis combined with stochastic perturbation techniques and projective limit large deviation methods.

6.4. Bifurcating autoregressive processes

We investigate the asymptotic behavior of the least squares estimator of the unknown parameters of random coefficient bifurcating autoregressive processes. Under suitable assumptions on inherited and environmental effects, we establish the almost sure convergence of our estimates. In addition, we also prove a quadratic strong law and central limit theorems. Our approach mainly relies on asymptotic results for vector-valued martingales together with the well-known Rademacher-Menchov theorem.

We study also the asymptotic behavior of the weighted least square estimators of the unknown parameters of random coefficient bifurcating autoregressive processes. Under suitable assumptions on the immigration and the inheritance, we establish the almost sure convergence of our estimators, as well as a quadratic strong law and central limit theorems. Our study mostly relies on limit theorems for vector-valued martingales.

Finally, we study the asymptotic behavior of the weighted least squares estimators of the unknown parameters of bifurcating integer-valued autoregressive processes. Under suitable assumptions on the immigration, we establish the almost sure convergence of our estimators, together with the quadratic strong law and central limit theorems. All our investigation relies on asymptotic results for vector-valued martingales.

6.5. Durbin-Watson statistic and first order autoregressive processes

We investigate moderate deviations for the Durbin-Watson statistic associated with the stable first-order autoregressive process where the driven noise is also given by a first-order autoregressive process. We first establish a moderate deviation principle for both the least squares estimator of the unknown parameter of the autoregressive process as well as for the serial correlation estimator associated with the driven noise. It enables us to provide a moderate deviation principle for the Durbin-Watson statistic in the easy case where the driven noise is normally distributed and in the more general case where the driven noise satisfies a less restrictive Chen-Ledoux type condition.

We investigate the asymptotic behavior of the Durbin-Watson statistic for the general stable $p-$order autoregressive process when the driven noise is given by a first-order autoregressive process. We establish the almost sure convergence and the asymptotic normality for both the least squares estimator of the unknown vector parameter of the autoregressive process as well as for the serial correlation estimator associated with the driven noise. In addition, the almost sure rates of convergence of our estimates are also provided. Then, we prove the almost sure convergence and the asymptotic normality for the Durbin-Watson statistic. Finally, we propose a new bilateral statistical procedure for testing the presence of a significative first-order residual autocorrelation and we also explain how our procedure performs better than the commonly used Box-Pierce and Ljung-Box statistical tests for white noise applied to the stable autoregressive process, even on small-sized samples.

In a recent paper (to appear hal-00677600), we investigate the asymptotic behavior of the maximum Likelihood estimators of the unknown parameters of positive recurrent Ornstein-Uhlenbeck processes driven by Ornstein-Uhlenbeck processes.

6.6. Ornstein-Uhlenbeck process with shift

We investigate the large deviation properties of the maximum likelihood estimators for the Ornstein-Uhlenbeck process with shift. We estimate simultaneously the drift and shift parameters. On the one hand, we establish a large deviation principle for the maximum likelihood estimates of the drift and shift parameters. Surprisingly, we find that the drift estimator shares the same large deviation principle as the one previously established for the Ornstein-Uhlenbeck process without shift. Sharp large deviation principles are also provided. On the other hand, we show that the maximum likelihood estimator of the shift parameter satisfies a large deviation principle with a very unusual implicit rate function.

6.7. Markovian superquadratic BSDEs

In [Stochastc Process. Appl., 122(9):3173-3208], the author proved the existence and the uniqueness of solutions to Markovian superquadratic BSDEs with an unbounded terminal condition when the generator and the terminal condition are locally Lipschitz. In [8], we prove that the existence result remains true for these BSDEs when the regularity assumptions on the generator and/or the terminal condition are weakened.

6.8. Non-Asymptotic Analysis of Adaptive and Annealed Feynman-Kac Particle Models

Sequential and Quantum Monte Carlo methods, as well as genetic type search algorithms can be interpreted as a mean field and interacting particle approximations of Feynman-Kac models in distribution spaces. The performance of these population Monte Carlo algorithms is strongly related to the stability properties of nonlinear Feynman-Kac semigroups. We analyze these models in terms of Dobrushin ergodic coefficients of the reference Markov transitions and the oscillations of the potential functions. Sufficient conditions for uniform concentration inequalities w.r.t. time are expressed explicitly in terms of these two quantities. We provide an original perturbation analysis that applies to annealed and adaptive FK models, yielding what seems to be the first results of this kind for these type of models. Special attention is devoted to the particular case of Boltzmann-Gibbs measures' sampling. In this context, we design an explicit way of tuning the number of Markov Chain Monte Carlo iterations with temperature schedule. We also propose and analyze an alternative interacting particle method based on an adaptive strategy to define the temperature increments.

6.9. A Robbins-Monro procedure for a class of models of deformation

We are interested with the statistical analysis of several data sets associated with shape invariant models with different translation, height and scaling parameters. We propose to estimate these parameters together with the common shape function. Our approach extends the recent work of Bercu and Fraysse to multivariate shape invariant models. We propose a very efficient Robbins-Monro procedure for the estimation of the translation parameters and we use these estimates in order to evaluate scale parameters. The main pattern is estimated by a weighted Nadaraya-Watson estimator. We provide almost sure convergence and asymptotic normality for all estimators. Finally, we illustrate the convergence of our estimation procedure on simulated data as well as on real ECG data.

6.10. Individual load curves intraday forecasting

A dynamic coupled modelling is investigated to take temperature into account in the individual energy consumption forecasting. The objective is both to avoid the inherent complexity of exhaustive SARIMAX models and to take advantage of the usual linear relation between energy consumption and temperature for thermosensitive customers. We first recall some issues related to individual load curves forecasting. Then, we propose and study the properties of a dynamic coupled modelling taking temperature into account as an exogenous contribution and its application to the intraday prediction of energy consumption. Finally, these theoretical results are illustrated on a real individual load curve. The authors discuss the relevance of such an approach and anticipate that it could form a substantial alternative to the commonly used methods for energy consumption forecasting of individual customers.

ALF Project-Team

6. New Results

6.1. Processor Architecture within the ERC DAL project

Participants: Pierre Michaud, Nathanaël Prémillieu, Luis Germán Garcia Morales, Bharath Narasimha Swamy, Sylvain Collange, André Seznec, Arthur Perais, Surya Natarajan, Sajith Kalathingal, Tao Sun, Andrea Mondelli, Aswinkumar Sridharan, Alain Ketterlin, Kamil Kedzierski.

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

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

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

In the DAL research project, http://www.irisa.fr/alf/index.php?option=com_content&view=article&id=55&Itemid=3&lang=en, we explore the microarchitecture techniques that will be needed to enable high performance on such heterogeneous processor chips. Very high performance will be required on both sequential sections, -legacy sequential codes, sequential sections of parallel applications-, and critical threads on parallel applications, -e.g. the main thread controlling the application. Our research focuses essentially on enhancing single process performance.

6.1.1. Microarchitecture exploration of control flow reconvergence

Participants: Nathanaël Prémillieu, André Seznec.

After continuous progress over the past 15 years [8], [10], the accuracy of branch predictors seems to be reaching a plateau. Other techniques to limit control dependency impact are needed. Control flow reconvergence is an interesting property of programs. After a multi-option control-flow instruction (i.e. either a conditional branch or an indirect jump including returns), all the possible paths merge at a given program point: the reconvergence point.

Superscalar processors rely on aggressive branch prediction, out-of-order execution and instruction level parallelism for achieving high performance. Therefore, on a superscalar core , the overall speculative execution after the mispredicted branch is cancelled, leading to a substantial waste of potential performance. However, deep pipelines and out-of-order execution induce that, when a branch misprediction is resolved, instructions following the reconvergence point have already been fetched, decoded and sometimes executed. While some of this executed work has to be cancelled since data dependencies exist, canceling the control independent work is a waste of resources and performance. We have proposed a new hardware mechanism called SYRANT, SYmmetric Resource Allocation on Not-taken and Taken paths, addressing control flow reconvergence at a reasonable cost. Moreover, as a side contribution of this research we have shown that, for a modest hardware cost, the outcomes of the branches executed on the wrong paths can be used to guide branch prediction on the correct path [13].

6.1.2. Efficient Execution on Guarded Instruction Sets

Participants: Nathanaël Prémillieu, André Seznec.

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

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

Building on top of recent advances in branch prediction and confidence estimation, we propose a hybrid branch and guard predictor, combining a global branch history component and global branch-and-guard history component. The potential gain or loss due to the systematic use of guard prediction is dynamically evaluated at run-time. Two computing modes are enabled: systematic guard prediction use and high confidence only guard prediction use. Our experiments show that on most applications, an overwhelming majority of guarded instructions are predicted. Therefore a relatively inefficient but simple hardware solution can be used to execute the few unpredicted guarded instructions. Significant performance benefits are observed on most applications while applications with poorly predictable guards do not suffer from performance loss [35], [34], [13].

6.1.3. Revisiting Value Prediction

Participants: Arthur Perais, André Seznec.

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

First, we introduce a new value predictor VTAGE harnessing the global branch history [32]. VTAGE directly inherits the structure of the indirect jump predictor ITTAGE [8]. VTAGE is able to predict with a very high accuracy many values that were not correctly predicted by previously proposed predictors, such as the FCM predictor and the stride predictor. Three sources of information can be harnessed by these predictors: the global branch history, the differences of successive values and the local history of values. Moreover, VTAGE does not suffer from short critical prediction loops and can seamlessly handle back-to-back predictions, contrarily to previously proposed, hard to implement FCM predictors.

Second, we show that all predictors are amenable to very high accuracy at the cost of some loss on prediction coverage [32]. This greatly diminishes the number of value mispredictions and allows to delay validation until commit-time. As such, no complexity is added in the out-of-order engine because of VP (save for ports on the register file) and pipeline squashing at commit-time can be used to recover. This is crucial as adding *selective replay* in the OoO core would tremendously increase complexity.

Third, we leverage the possibility of validating predictions at commit to introduce a new microarchitecture, EOLE [31]. EOLE features *Early Execution* to execute simple instructions whose operands are ready in parallel with Rename and *Late Execution* to execute simple predicted instructions and high confidence branches just before Commit. EOLE depends on Value Prediction to provide operands for *Early Execution* and predicted instructions for *Late Execution*. However, Value Prediction requires EOLE to become truly practical. That is, EOLE allows to reduce the out-of-order issue-width by 33% without impeding performance. As such, the number of ports on the register file diminishes. Furthermore, optimizations of the register file such as *banking* further reduce the number of required ports. Overall EOLE possesses a register file whose complexity is on-par with that of a regular wider-issue superscalar while the out-of-order components (scheduler , bypass) are greatly simplified. Moreover, thanks to Value Prediction, speedup is obtained on many benchmarks of the SPEC'00/'06 suite.

6.1.4. Helper threads

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

As the number of cores on die increases with the improvements in silicon process technology, the strategy of replicating identical cores does not scale to meet the performance needs of mixed workloads. Heterogeneous Many Cores (HMC) that mix many simple cores with a few complex cores are emerging as a design alternative that can provide both high performance and power-efficient execution. The availability of many simple cores in a HMC presents an opportunity to utilize low power cores to accelerate sequential execution on the complex core. For example simple cores can execute pre-computational (or helper) code and generate prefetch requests for the main thread.

We explore the design of a lightweight architectural framework that provides instruction set support and a low-latency interface to simple-cores for efficient helper code execution. We utilize static analyses and profile data to generate helper codelets that target delinquent loads in the main thread. The main thread is instrumented to initiate helper execution ahead of time, and utilizes instruction set support to signal helper execution on the simple core, and to pass live-in values for the helper codelet. Pre-computational code executes on the simple core and generates prefetch requests that install data into a shared last-level cache. Initial experiments with a trace based simulation framework show that helper execution has the potential to cover cache-missing loads on the main thread.

The restriction of prefetching to a lower level shared cache in a loosely coupled system limits the benefits of helper execution. The main thread should have a low latency access mechanism to data prefetched by helper execution. We plan to explore direct, yet light weight, mechanisms for data communication between the helper core and the main core.

6.1.5. Adaptive Intelligent Memory Systems

Participants: André Seznec, Aswinkumar Sridharan.

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

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

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

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

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

Multi-core have become ubiquitous and industry is already moving towards the many-core era. Many open-ended questions remain unanswered for the upcoming many-core era. From the software perspective, it is unclear which applications will be able to benefit from many cores. From the hardware perspective, the tradeoff between implementing many simple cores, fewer medium aggressive cores or even only a moderate number of aggressive cores is still in debate. Estimating the potential performance of future parallel applications on the yet-to-be-designed future many cores is very speculative. The simple models proposed by Amdahl's law or Gustafson's law are not sufficient and may lead to erroneous conclusions. In this paper, we propose a still simple execution time model for parallel applications, the SNAS model. As previous models, the SNAS model evaluates the execution time of both the serial part and the parallel part of the application, but takes into account the scaling of both these execution times with the input problem size and the number of processors. For a given application, a few parameters are collected on the effective execution of the application with a few threads and small input sets. The SNAS model allows to extrapolate the behavior of a future application exhibiting similar scaling characteristics on a many core and/or a large input set. Our study shows that the execution time of the serial part of many parallel applications tends to increase along with the problem size, and in some cases with the number of processors. It also shows that the efficiency of the execution of the parallel part decreases dramatically with the number of processors for some applications. Our model also indicates that since several different applications scaling will be encountered, hybrid architectures featuring a few aggressive cores and many simple cores should be privileged.

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

Participants: Sylvain Collange, André Seznec, Sajith Kalathingal.

We aim at exploring the design of a unique core that efficiently run both sequential and massively parallel sections. We explore how the architecture of a complex superscalar core has to be modified or enhanced to be able to support the parallel execution of many threads from the same application (10's or even 100's a la GPGPU on a single core).

SIMD execution is the preferred way to increase energy efficiency on data-parallel workloads. However, explicit SIMD instructions demand challenging auto-vectorization or manual coding, and any change in SIMD width requires at least a recompile, and typically manual code changes. Rather than vectorize at compile-time, our approach is to dynamically vectorize SPMD programs at the micro-architectural level. The SMT-SIMD hybrid core we propose extracts data parallelism from thread parallelism by scheduling groups of threads in lockstep, in a way inspired by the execution model of GPUs. As in GPUs, conditional branches whose outcome differ between threads are handled with conditionally masked execution. However, while GPUs rely on explicit re-convergence instructions to restore lockstep execution, we target existing general-purpose instruction sets, in order to run legacy binary programs. Thus, the main challenge consists in detecting re-convergence points dynamically.

We proposed instruction fetch policies that apply heuristics to maximize the cycles spent in lockstep execution. We evaluated their efficiency and performance impact on an out-of-order superscalar core simulator. Results validate the viability of our approach, by showing that existing compiled SPMD programs are amenable to lockstep execution without modification nor recompilation.

6.2. Other Architecture Studies

Participants: Damien Hardy, Pierre Michaud, Ricardo Andrés Velásquez, Sylvain Collange, André Seznec, Sajith Kalathingal, Junjie Lai.

GPU, performance, simulation, vulnerability

6.2.1. Performance Upperbound Analysis of GPU applications

Participants: Junjie Lai, André Seznec.

In the framework of the ANR Cosinus PetaQCD project (ended Oct 2012), we have been modeling the demands of high performance scientific applications on hardware. GPUs have become popular and cost-effective hardware platforms. In this context, we have been addressing the gap between theoretical peak performance on GPU and the effective performance. There have been many studies on optimizing specific applications on GPU and also a lot of studies on automatic tuning tools. However, the gap between the effective performance and the maximum theoretical performance is often huge. A tighter performance upperbound of an application is needed in order to evaluate whether further optimization is worth the effort. We designed a new approach to compute the CUDA application's performance upperbound through intrinsic algorithm information coupled with low-level hardware benchmarking. Our analysis [11], [22] allows us to understand which parameters are critical to the performance and have more insights of the performance result. As an example, we analyzed the performance upperbound of SGEMM (Single-precision General Matrix Multiply) on Fermi and Kepler GPUs. Through this study, we uncover some undocumented features on Kepler GPU architecture. Based on our analysis, our implementations of SGEMM achieve the best performance on Fermi and Kepler GPUs so far (5 % improvement on average).

6.3. Microarchitecture Performance Analysis

Participants: Ricardo Andrés Velásquez, Pierre Michaud, André Seznec.

6.3.1. *Selecting benchmark combinations for the evaluation of multicore throughput*

Participants: Ricardo Andrés Velásquez, Pierre Michaud, André Seznec.

In [26], we have shown that fast approximate microarchitecture models such as BADCO [16] can be useful for selecting multiprogrammed workloads for evaluating the throughput of multicore processors. Computer architects usually study multiprogrammed workloads by considering a set of benchmarks and some combinations of these benchmarks. However, there is no standard method for selecting such sample, and different authors have used different methods. The choice of a particular sample impacts the conclusions of a study. Using BADCO, we propose and compare different sampling methods for defining multiprogrammed workloads for computer architecture. We evaluate their effectiveness on a case study, the comparison of several multicore last-level cache replacement policies. We show that random sampling, the simplest method, is robust to define a representative sample of workloads, provided the sample is big enough. We propose a method for estimating the required sample size based on fast approximate simulation. We propose a new method, workload stratification, which is very effective at reducing the sample size in situations where random sampling would require large samples.

6.3.2. *A systematic approach for defining multicore throughput metrics*

Participant: Pierre Michaud.

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

Measuring throughput is not as straightforward as measuring execution time. This has led to an ongoing debate on what forms a meaningful throughput metric for multi-program workloads. In [29], we present a method to construct throughput metrics in a systematic way: we start by expressing assumptions on job size, job distribution, scheduling, etc., that together define a theoretical throughput experiment. The throughput metric is then the average throughput of this experiment. Different assumptions lead to different metrics, so one should select the metric whose assumptions are close to the real usage he/she has in mind. We elaborate multiple metrics based on different assumptions. In particular, we identify the assumptions that lead to the commonly used weighted speedup and harmonic mean of speedups. Our study clarifies that they are actual throughput metrics, which was recently questioned. We also propose some new throughput metrics, whose calculation sometimes requires approximation. We use synthetic and real experimental data to characterize metrics and show how they relate to each other. Our study can also serve as a starting point if one needs to define a new metric based on specific assumptions, other than the ones we consider in this study. Throughput metrics should always be defined from explicit assumptions, because this leads to a better understanding of the implications and limits of the results obtained with that metric.

6.4. Compiler, vectorization, interpretation

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

6.4.1. Vectorization Technology To Improve Interpreter Performance

Participant: Erven Rohou.

Recent trends in consumer electronics have created a new category of portable, lightweight software applications. Typically, these applications have fast development cycles and short life spans. They run on a wide range of systems and are deployed in a target independent bytecode format over Internet and cellular networks. Their authors are untrusted third-party vendors, and they are executed in secure managed runtimes or virtual machines. Furthermore, due to security policies, these virtual machines are often lacking just-in-time compilers and are reliant on interpreter execution.

The main performance penalty in interpreters arises from instruction dispatch. Each bytecode requires a minimum number of machine instructions to be executed. In this work we introduce a powerful and portable representation that reduces instruction dispatch thanks to vectorization technology. It takes advantage of the vast research in vectorization and its presence in modern compilers. Thanks to a split compilation strategy, our approach exhibits almost no overhead. Complex compiler analyses are performed ahead of time. Their results are encoded on top of the bytecode language, becoming new SIMD IR (i.e., intermediate representation) instructions. The bytecode language remains unmodified, thus this representation is compatible with legacy interpreters.

This approach drastically reduces the number of instructions to interpret and improves execution time. [15]. SIMD IR instructions are mapped to hardware SIMD instructions when available, with a substantial improvement.

6.4.2. Improving sequential performance: the case of floating point computations

Participants: Erven Rohou, André Seznec, Arjun Suresh.

One way to enhance sequential performance is to consider floating point computations. Languages and instruction sets provide support for only a few representations, namely float and double, and programmers are likely to use the most accurate (unless they handle large data structures). Still, in most cases, programmers do not formally specify the precision they require from their applications, and have no guarantee on the precision they actually get. This is an opportunity for a tradeoff between performance and precision: programs could run faster at the expense of a less accurate result (note that existing compilers already embed some unsafe transformations, for example when flags such as -fast or -ffastmath are used).

The first step consisted in applying memoization to the math library libm. In this case, results are still correct. The performance improvement comes from caching results of pure functions, and retrieving them instead of recomputing a result. This shows good results on floating point intensive benchmarks. In a next step, a helper thread will monitor the patterns of parameters and precompute likely values to "prefetch" results ahead of time.

Reduced precision comes into play when no pattern can be identified, but the new value is close enough to already computed values. We plan to apply interpolation to compute the result faster than the standard code. We will also investigate how we can leverage known properties of mathematical functions, as well as programmer hints about useful properties of user-defined functions, and where reduced precision is acceptable.

6.4.3. Identifying divergence in GPU architectures

Participant: Sylvain Collange.

This research is done in collaboration with Fernando M. Q. Pereira, Diogo Sampaio and Rafael Martins de Souza, UFMG, Brazil.

GPU architectures rely on SIMD execution by vectorizing across SPMD threads. They achieve the best performance when consecutive threads take the same paths through conditional branches and access contiguous memory locations. Thus, many GPU code optimizations that target the control flow or memory access patterns necessitate accurate information about which branches and memory accesses are divergent across threads.

To enable such optimizations, we proposed divergence analysis, a compiler pass that identifies similarities in the control flow and data flow of concurrent threads [37]. This static analysis identifies program variables that are affine functions of the thread identifier and propagate this knowledge to conditional branches and memory accesses. Our analysis consistently outperforms other comparable analyses, thanks to the combination of taking into account affine relations between variables and accurately modeling control dependencies.

6.4.4. Code Obfuscation

Participant: Erven Rohou.

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

We proposed to leverage JIT compilation to make software tamper-proof. The idea is to constantly generate different versions of an application, even while it runs, to make reverse engineering hopeless. More precisely a JIT engine is used to generate new versions of a function each time it is invoked, applying different optimizations, heuristics and parameters to generate diverse binary code. A strong random number generator will guarantee that generated code is not reproducible, though the functionality is the same [38].

On-Stack-Replacement has been previously proposed to recompile functions while they run. However, it relies on compiler-generated switch points. We proposed a new technique to recompile functions at arbitraty points, thus reinforcing the Obfuscating JIT approach. A prototype is being developed [27].

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

6.4.5. Padrone

Participants: Erven Rohou, Alain Ketterlin, Emmanuel Riou.

The objective of the ADT PADRONE is to design and develop a platform for re-optimization of binary executables at run-time. Development is ongoing, and an early prototype is functional. In [24], we described the infrastructure of Padrone, and showed that its profiling overhead is minimum. We illustrated its use through two examples. The first example shows how a user can easily write a tool to identify hotspots in their application, and how well they perform (for example, by computing the number of executed instructions per cycle). In the second example, we illustrate the replacement of a given function (typically a hotspot) by an optimized version, while the program runs.

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

6.4.6. Dynamic Analysis and Re-Optimization

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

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

Dynamic binary analysis and re-optimization is specially interesting for legacy or commercial applications, but also in the context of cloud deployment, where actual hardware is unknown, and other applications competing for hardware resources can vary.

Initial results show that we are able to identify function hotspots that contain vectorized code for the Intel SSE extension, analyze them, and reoptimize the loops to target the latest and more powerful AVX ISA extension.

6.4.7. Branch Prediction and Performance of Interpreter

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

Interpreters have been used in many contexts. They provide portability and ease of development at the expense of performance. The literature of the past decade covers analysis of why interpreters are slow, and many software techniques to improve them. A large proportion of these works focuses on the dispatch loop, and in particular on the implementation of the switch statement: typically an indirect branch instruction. Conventional wisdom attributes a significant penalty to this branch, due to its high misprediction rate. We revisit this assumption [36], considering current interpreters, and modern predictors. Using both hardware counters and simulation, we show that the accuracy of indirect branch prediction is no longer critical for interpreters. We also compare the characteristics of these interpreters and analyze why the indirect branch is less important than before.

6.5. WCET estimation

Participants: Damien Hardy, Benjamin Lesage, Hanbing Li, Isabelle Puaut, Erven Rohou, André Seznec.

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

6.5.1. WCET estimation and multi-core systems

6.5.1.1. Predictable shared caches for mixed-criticality real-time systems

Participants: Benjamin Lesage, Isabelle Puaut, André Seznec.

The general adoption of multi-core architectures has raised new opportunities as well as new issues in all application domains. In the context of real-time applications, it has created one major opportunity and one major difficulty. On the one hand, the availability of multiple high performance cores has created the opportunity to mix on the same hardware platform the execution of a complex critical real-time workload and the execution of non-critical applications. On the other hand, for real-time tasks timing deadlines must be met and enforced. Hardware resource sharing inherent to multicores hinders the timing analysis of concurrent tasks. Two different objectives are then pursued: enforcing timing deadlines for real-time tasks and achieving highest possible performance for the non-critical workload.

In this work, we suggest a hybrid hardware-based cache partitioning scheme that aims at achieving these two objectives at the same time. Plainly considering inter-task conflicts on shared cache for real-time tasks yields very pessimistic timing estimates. We remove this pessimism by reserving private cache space for real-time tasks. Upon the creation of a real-time task, our scheme reserves a fixed number of cache lines per set for the task. Therefore uniprocessor worst case execution time (WCET) estimation techniques can be used, resulting in tight WCET estimates. Upon the termination of the real-time task, this private cache space is released and made available for all the executed threads including non-critical ones. That is, apart the private spaces reserved for the real-time tasks currently running, the cache space is shared by all tasks running on the processor, i.e. non-critical tasks but also the real-time tasks for their least recently used blocks. Experiments show that the proposed cache scheme allows to both guarantee the schedulability of a set of real-time tasks with tight timing constraints and enable high performance on the non-critical tasks.

This work is the main contribution of the PhD thesis of Benjamin Lesage [12].

6.5.1.2. WCET estimation for massively parallel processor arrays

Participant: Isabelle Puaut.

This is joint work with Dumitru Potop-Butucaru, Inria, EPI AOSTE.

Classical timing analysis techniques for parallel code isolates micro-architecture analysis from the analysis of synchronizations between cores by performing them in two separate analysis phases (WCET – worst-case execution time – and WCRT – worst-case response time analyses). This isolation has its advantages, such as a reduction of the complexity of each analysis phase, and a separation of concerns that facilitates the development of analysis tools. But isolation also has a major drawback: a loss in precision which can be significant. To consider only one aspect, to be safe the WCET analysis of each synchronization-free sequential code region has to consider an undetermined micro-architecture state. This may result in overestimated WCETs, and consequently on pessimistic execution time bounds for the whole parallel application. The contribution of this work [33], [23] is an *integrated* WCET analysis approach that considers at the same time micro-architectural information and the synchronizations between cores. This is achieved by extending a state-of-the-art WCET estimation technique and tool to manage synchronizations and communications between the sequential threads running on the different cores. The benefits of the proposed method are twofold. On the one hand, the micro-architectural state is not lost between synchronization-free code regions running on the same core, which results in tighter execution time estimates. On the other hand, only one tool is required for the temporal validation of the parallel application, which reduces the complexity of the timing validation toolchain.

Such a holistic approach is made possible by the use of deterministic and composable software and hardware architectures (homogeneous multi-cores without cache sharing, static assignment of the code regions on the cores). We demonstrate the interest of the approach using an adaptive differential pulse-code modulation (*adpcm*) encoder where the integrated WCET approach provides significantly tighter response time estimations than the more classical WCRT approaches, with a gain of 21% on average.

6.5.2. WCET estimation for architectures with faulty caches

Participants: Damien Hardy, Isabelle Puaut.

Semiconductor technology evolution suggests that permanent failure rates will increase dramatically with scaling, in particular for SRAM cells. While well known approaches such as error correcting codes exist to recover from failures and provide fault-free chips, they will not be affordable anymore in the future due to their non-scalable cost. Consequently, other approaches like fine grain disabling and reconfiguration of hardware elements (e.g. individual functional units or cache blocks) will become economically necessary. This fine-grain disabling will lead to degraded performance compared to a fault-free execution.

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

In [21] a method that statically calculates a probabilistic WCET bound in the presence of permanent faults in instruction caches is provided. The method, from a given program, cache configuration and probability of cell failure, derives a probabilistic WCET bound. The proposed method, because it relies on static analysis, is guaranteed to identify the longest program path, its probabilistic nature only stemming from the presence of faults. The method is computationally tractable because it does not require an exhaustive enumeration of the possible locations of faulty cache blocks. Experimental results show that it provides WCET estimates very close to, but never below, the method that derives probabilistic WCETs by enumerating all possible locations of faulty cache blocks. The proposed method not only allows to quantify the impact of permanent faults on WCET estimates, but also can be used in architectural exploration frameworks to select the most appropriate fault management mechanisms.

6.5.3. Traceability of flow information for WCET estimation

Participants: Hanbing Li, Isabelle Puaut, Erven Rohou.

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

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

The objective of this work is to address the challenging issue of the mapping and transformation of the flow information from high level down to machine code. In a first step, we have considered the issue of conveying information through the compilation flow, without any optimization. We have created our own WCET information type and used the annotation files FFX (Flow Fact in XML, provided by IRIT, partner of the W-SEPT project), and applied them to the LLVM compiler framework. We are currently studying the impact of optimizations on the traceability of annotations. We are currently designing a framework for flow fact transformation for a large panel of compiler optimizations.

6.6. HPC and mobile computing

Participant: François Bodin.

We have initiated a research action on the interaction between mobile computing and HPC. We aim at studying data representation linked to parallel programming in heterogeneous systems. In particular, we want to explore energy tradeoffs when changing hardware resources from a light mobile platform to remote execution in a datacenter.

As a test case, we are developing an application for inventorying art pieces in the public domain. This is done in collaboration with University of Rennes 2. This test case is a pluridisplinary collaboration whose goal for University of Rennes 2 is to study how mobile computing can contribute to art studies and dissemination.

6.7. Application-specific number systems

Participant: Sylvain Collange.

This research is done in collaboration with Mark G. Arnold, XLNS Research, USA.

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

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

We proposed a generalization of LNS that incorporate features analogous to subnormal floating-point [18], [28]. The Denormal LNS (DLNS) system we introduce defines a class of hybrid number systems that offer quasi-constant absolute accuracy close to zero and quasi-constant relative accuracy on larger numbers. These systems can be configured to range from pure LNS (constant relative accuracy) to fixed-point (constant absolute accuracy across the whole range).

ALGORILLE Project-Team

6. New Results

6.1. Structuring applications for scalability

In this domain we have been active on several research subjects: efficient locking interfaces, data management, asynchronism, algorithms for large scale discrete structures and the use of accelerators, namely GPU.

In addition to these direct contributions within our own domain, numerous collaborations have permitted us to test our algorithmic ideas in connection with academics of different application domains and through our association with SUPÉLEC with some industrial partners: physics and geology, biology and medicine, machine learning or finance.

6.1.1. Efficient linear algebra on accelerators.

Participants: Sylvain Contassot-Vivier, Thomas Jost.

The PhD thesis of Thomas Jost, co-supervised by S. Contassot-Vivier and Bruno Lévy (Alice INRIA team) since January 2010, dealt with specific algorithms for GPUs, in particular linear solvers [32]. He also worked on the use of GPUs within clusters of workstations via the study of a solver of non-linear problems [30], [33], [29]. The defense of this thesis was initially planned in January 2013 but Thomas decided at the end of 2012 to stop his PhD and to leave for industry.

6.1.2. Development methodologies for parallel programming of clusters.

Participants: Sylvain Contassot-Vivier, Jens Gustedt, Stéphane Vialle.

We have conducted a particular effort in merging and synthesizing our respective experiences of parallel programming of clusters (homogeneous, heterogeneous, hybrid). This has lead to two book chapters [19] and [34] (to appear).

6.1.3. Combining locking and data management interfaces.

Participants: Jens Gustedt, Stéphane Vialle, Soumeya Leila Hernane, Rodrigo Campos-Catelin.

Handling data consistency in parallel and distributed settings is a challenging task, in particular if we want to allow for an easy to handle asynchronism between tasks. Our publication [4] shows how to produce deadlock-free iterative programs that implement strong overlapping between communication, IO and computation. The thesis of Soumeya Hernane [12] has been defended in 2013. It extends distributed lock mechanisms and combines them with implicit data management.

A new implementation (ORWL) of our ideas of combining control and data management in C has been undertaken, see 5.2.1 . In 2013, work has demonstrated its efficiency for a large variety of platforms, see [22]. By using the example of dense matrix multiplication, we show that ORWL permits to reuse existing code for the target architecture, namely open source library ATLAS, Intel's compiler specific MKL library or NVidia's CUBLAS library for GPUs. ORWL assembles local calls into these libraries into efficient functional code, that combines computation on distributed nodes with efficient multi-core and accelerator parallelism.

Additionally, during the internship of Rodrigo Campos-Catelin, a detailed instrumentation of the ORWL library has been undertaken, and a new, less expensive strategy for cyclic FIFOs has been tested. This work will be continued with a master thesis at the university of Buenos Aires that will summarize and extend the results that were achieved during the internship.

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

6.1.4. Discrete and continuous dynamical systems.

Participants: Sylvain Contassot-Vivier, Marion Guthmuller.

The continuous aspect of dynamical systems has been intensively studied through the development of asynchronous algorithms for solving PDE problems. In past years, we have focused our studies on the interest of GPUs in asynchronous algorithms [29]. Also, we have investigated the possibility to insert periodic synchronous iterations inside the asynchronous scheme in order to improve the convergence detection delay. This is especially interesting on small/middle sized clusters with efficient networks. The SimGrid environment has been used to validate and evaluate load balancing strategies in parallel iterative algorithms on large scale systems [28].

In 2011, the PhD thesis of Marion Guthmuller, supervised by M. Quinson and S. Contassot-Vivier, has started on the subject of model-checking distributed applications inside the SimGrid simulator [31]. The expected results of that work may provide a very interesting tool for studying dynamical systems expressed under the form of a distributed application.

6.2. Transparent Resource Management for Clouds

Participants: Julien Gossa, Rajni Aron, Stéphane Genaud, Étienne Michon, Marc-Eduard Frîncu.

6.2.1. Provisioning strategies.

Our main achievement was the design of one comprehensive provisioning meta-strategy. This meta-strategy only use one parameter as a deadline given by the user. Contrary to other deadline-based provisioning strategies, our meta-strategy is able to combine any provisioning strategy in order to optimize the cost while meeting the deadline. This is achieved through simulation of cost and makespan of every available strategy thanks to SCHIaaS5.4.3 . It allows to apply the most inexpensive strategy as long as possible, before progressively switching to more expensive strategy when the deadline becomes closer.

The next step is to asses this meta-strategy among an important set of applications and platforms, both in real environments and simulation. The data are currently gathered and analyzed, and we should be able to draw conclusions soon.

6.2.2. User workload analysis.

We have conducted one broad study about workflows execution on the cloud, from both the theoretical and experimental point of view. In this study, we tried to discover causalities between the characteristics of workflows and the performances of provisioning strategies. We concluded that, except very peculiar cases, no causality can be identified. That is why we decided to make use of simulation to predict the strategies performances.

This predictive process is now integrated as a module of our cloud broker. It can be invoked by a user to help him decide which strategy should be used before any actual resource leasing.

We are now convinced that workload analysis is not a suitable approach because of its lack of generality.

6.2.3. Experimentations.

Given the very large consumption of CPU hours, the above work was supported mostly by simulation. We have assessed the gap between the performances of real executions on a private cloud and simulation. The latter proved to be very accurate, predicting almost perfectly the cost and makespan of every strategy on a wide range of workloads.

However, we have also shown that the simulation can be very sensitive to user defined input parameters (such as task runtimes) and may be mislead in borderline cases. Identifying the pitfalls and limitations of the simulation is very important and should end up in recommendations for a wise interpretation of simulation results.

We have also extended the range of experimentations to assess our simulator. First, we have extended the set of simulations with new applications, mostly workflows that are both generated and real applications (i.e. Montage). Second, we have conducted intensive experimentations on new platforms (i.e. Bonfire). The experimental data we have recently gathered in both cases is to be analyzed to further validate our approach.

6.3. Experimental methodologies for the evaluation of distributed systems

This year, M. Quinson defended his Habilitation on the experimental methodologies of distributed systems [13]. This concludes 10 years of research on this topic (including the elements presented in this section), and paves the road of future research.

6.3.1. Simulation and dynamic verification

6.3.1.1. MPI simulation

Participants: Martin Quinson, Paul Bédaride, Marion Guthmuller.

We continued our long-term effort toward the simulation of HPC application within SimGrid. We slightly increased the API coverage of our reimplementation of MPI on top of SimGrid, and proposed a new model of the network performance for MPI applications on top of Ethernet TCP networks. This model combines the advantages of flow-based networks for large data transfers as previous SimGrid network models, but also leverage algorithmic performance models extending the classical LogP models. As shown in [16], these models greatly improve the realism of MPI simulations, enabling the prediction of the performance of a non-trivial application in great details.

6.3.1.2. Dynamic verification and SimGrid

Participants: Marion Guthmuller, Martin Quinson, Gabriel Corona.

This year, our work toward the verification of liveness properties within SimGrid became fully functional thanks to the PhD work of M. Guthmuller. This relies on a system-level introspection mechanism allowing the model checker to finely explore the state of the verified programs. This is mandatory to detect the execution cycles that constitute the counter examples to liveness properties. This introspection mechanism is also used to implement a new reduction mechanism that can mitigate the state space explosion problem. A publication presenting these results is currently under review.

6.3.1.3. SimGrid framework improvement

Participants: Paul Bédaride, Martin Quinson, Gabriel Corona.

We rolled out a new major version of the SimGrid framework to our users. It contains both the HPC network models used to improve the prediction of MPI applications and all of our developments toward the dynamic verification of distributed applications. We also improved further the usability of our framework, that is now properly integrated within the Debian Linux distribution.

The next release is already underway, with a proper integration of the work from our partners on virtual machines and with a full reimplementation of the simulation kernel in C++ for a better modularity.

6.3.1.4. Formal Verification of Distributed Algorithms

Participants: Esteban Campostrini, Martin Quinson, Stephan Merz.

M. Quinson co-advised an internship with S. Merz (project-team Veridis) on the formal verification of distributed algorithm. The goal was to push further the PlusCal algorithmic language and its compiler to TLA$^+$ on which we are working since several years within the Veridis team.

We wanted to explore some hard problem raised by the verification of distributed protocol, such as how to represent timeout errors in verification settings where the time is not present. We think that this could be modeled somehow similarly to fairness properties, but more work is needed in this topic for a definitive answer.

6.3.2. Experimentation on testbeds and production facilities, emulation

6.3.2.1. Distem improvements: scalability and matrix-based inter-nodes latencies

Participants: Ahmed Bessifi, Emmanuel Jeanvoine, Lucas Nussbaum.

(For context, see sections 3.3 and 5.3 .)

Following our PDP'13 publication[18], we focused on improving Distem's scalability. First, on the Distem engine side, we parallelized the startup of physical nodes and virtual nodes, and added support for BTRFS snapshots to enable starting a very large number of virtual nodes with their own filesystems. Second, during the internship of Ahmed Bessifi we investigated several networking issues causing problems with large-scale experiments (over 4000 virtual nodes). The resulting improvements to ARP parameters tunings were integrated in Distem 0.8, and enabled network-intensive experiments with up to 8000 virtual nodes. We plan to publish those results in early 2014.

In the context of the AEN HEMERA project, we worked with Trong-Tuan Vu (EPI DOLPHIN, Inria Lille Nord Europe) to add support for specifying inter-nodes latencies using a matrix. This is especially useful for experiments on load-balancing and locality.

6.3.2.2. Evaluating load balancing HPC runtimes with Distem

Participants: Joseph Emeras, Emmanuel Jeanvoine, Lucas Nussbaum.

(For context, see sections 3.3 and 5.3 .)

We aim at demonstrating the suitability of Distem to evaluate Exascale and Cloud runtime environments providing load balancing and fault tolerance features. In that context, we reproduced some experiments published at CCGrid'2013 on Charm++ load balancers. Preliminary results are promising, and we hope that this will lead to collaborations with runtime developers.

A publication presenting how Distem to test HPC runtimes (scalability, fault tolerance and load balancing capabilities) is in the works.

6.3.2.3. Further improvements to XPFlow

Participants: Tomasz Buchert, Lucas Nussbaum, Jens Gustedt.

(For context, see sections 3.3 and 5.6 .)

We strengthened our XPFlow experiment control system using several sets of experiments, including experiments on the OpenStack IaaS Cloud stack on hundreds of Grid'5000 nodes.

A publication describing XPFlow was submitted to CCGrid'2014[21].

6.3.2.4. Further improvements to Kadeploy

Participants: Luc Sarzyniec, Emmanuel Jeanvoine, Lucas Nussbaum.

(For context, see sections 3.3 and 5.5 .)

We continued the development of Kadeploy:

- The support for multi-partition images was added;
- The communication interface between the Kadeploy server and the Kadeploy client was completely rewritten to use a REST API;
- A test framework, integrated with Inria's Continuous Integration facility, was added.

Two new Kadeploy releases were published during 2013, including those changes.

6.3.2.5. Grid'5000

Participants: Sébastien Badia, Luc Sarzyniec, Émile Morel, Lucas Nussbaum.

(For context, see sections 3.3 and 5.7 .)

The team continued to support Grid’5000. Highlights of 2013 include:

- Lucas Nussbaum is now a member of the *Bureau* and *Comité d’Architectes* of GIS Grid’5000. In the context of the *Comité d’Architectes*, he led the writing on several internal documents (on possible evolutions of the testbed).
- An article describing Grid’5000’s support for experiments on IaaS Clouds[15] was published at the *Testing The Cloud* workshop.
- A new cluster, *graphite*, was installed in Nancy.

6.3.3. Convergence and co-design of experimental methodologies

6.3.3.1. Practical study on combining experimental methodologies

Participants: Maximiliano Geier, Lucas Nussbaum, Martin Quinson.

During an internship, we explored how simulation, emulation and experimentation on Grid’5000 could be combined in practice. Starting with a simple question on a particular system, we used a representative framework for each methodology: SimGrid for simulation, Distem for emulation and Grid’5000 for experimentation, and described our experiments using the workflow logic provided by the XPFlow tool. We identified a set of pitfalls in each paradigm that experimenters may encounter regarding models, platform descriptions and others. We proposed a set of general guidelines to avoid these pitfalls. We showed these guidelines may lead to accurate simulation results. Finally, we provided some insight to framework developers in order to improve the tools and thus facilitate this convergence.

The results of this work were published at the *WATERS* workshop[17].

6.3.3.2. Organization of an event on reproducible research

Participant: Lucas Nussbaum.

We organized Realis, an event aiming at testing the experimental reproducibility of papers submitted to Compas’2013. Associated to the Compas’13 conference, this workshop aimed at providing a place to discuss the reproducibility of the experiments underlying the publications submitted to the main conference. We hope that this kind of venue will motivate the researchers to further detail their experimental methodology, ultimately allowing others to reproduce their experiments.

ALICE Project-Team

5. New Results

5.1. Geometry Processing

5.1.1. Fitting Polynomial Volumes to Surface Meshes with Voronoi Squared Distance Minimization

Participants: Gilles-Philippe Paillé, Bruno Lévy.

We propose a method for mapping polynomial volumes. Given a closed surface and an initial template volume grid, our method deforms the template grid by fitting its boundary to the input surface while minimizing a volume distortion criterion. The result is a point-to-point map distorting linear cells into curved ones. Our method is based on several extensions of Voronoi Squared Distance Minimization (VSDM) combined with a higher-order finite element formulation of the deformation energy. This allows us to globally optimize the mapping without prior parameterization. The anisotropic VSDM formulation allows for sharp and semi-sharp features to be implicitly preserved without tagging. We use a hierarchical finite element function basis that selectively adapts to the geometric details. This makes both the method more efficient and the representation more compact. We apply our method to geometric modeling applications in computer-aided design and computer graphics, including mixed-element meshing, mesh optimization, subdivision volume fitting, and shell meshing.

This work was presented at the "ACM Symposium on Geometry Processing" and published in the "Computer Graphics Forum" journal [16].

5.1.2. Particle-Based Anisotropic Surface Meshing

Participant: Bruno Lévy.

This paper introduces a particle-based approach for anisotropic surface meshing. Given an input polygonal mesh endowed with a Riemannian metric and a specified number of vertices, the method generates a metric-adapted mesh. The main idea consists of mapping the anisotropic space into a higher dimensional isotropic one, called "embedding space". The vertices of the mesh are generated by uniformly sampling the surface in this higher dimensional embedding space, and the sampling is further regularized by optimizing an energy function with a quasi-Newton algorithm. All the computations can be re-expressed in terms of the dot product in the embedding space, and the Jacobian matrices of the mappings that connect different spaces. This transform makes it unnecessary to explicitly represent the coordinates in the embedding space, and also provides all necessary expressions of energy and forces for efficient computations. Through energy optimization, it naturally leads to the desired anisotropic particle distributions in the original space. The triangles are then generated by computing the Restricted Anisotropic Voronoi Diagram and its dual Delaunay triangulation. We compare our results qualitatively and quantitatively with the state-of-the-art in anisotropic surface meshing on several examples, using the standard measurement criteria. This work was published in the "ACM Transactions on Graphics" journal (SIGGRAPH conference proceedings) [19].

5.1.3. Approximating Functions on a Mesh with Restricted Voronoi Diagrams

Participant: Bruno Lévy.

../../../../projets/alice/IMG/volumeFitting.png

Figure 2. Fitting Polynomial Volumes to Surface Meshes with Voronoi Squared Distance Minimization

We propose a method that computes a piecewise constant approximation of a function defined on a mesh. The approximation is associated with the cells of a restricted Voronoi diagram. Our method optimizes an objective function measuring the quality of the approximation. This objective function depends on the placement of the samples that define the restricted Voronoi diagram and their associated function values. We study the continuity of the objective function, derive the closed-form expression of its derivatives and use them to design a numerical solution mechanism. The method can be applied to a function that has discontinuities, and the result aligns the boundaries of the Voronoi cells with the discontinuities. Some examples are shown, suggesting potential applications in image vectorization and compact representation of lighting. This work was presented at the "ACM Symposium on Geometry Processing" and published in the "Computer Graphics Forum" journal [15].

../../../../projets/alice/IMG/Approx-pegase.png

Figure 3. Approximating Functions on a Mesh with Restricted Voronoi Diagrams

5.1.4. Spectral Clustering of Plant Units From 3D Point Clouds

Participant: Dobrina Boltcheva.

High-resolution terrestrial Light Detection And Ranging (tLiDAR), a 3-D remote sensing technique, has recently been applied for measuring the 3-D characteristics of vegetation from grass to forest plant species. The resulting data are known as a point cloud which shows the 3-D position of all the hits by the laser beam giving a raw sketch of the spatial distribution of plant elements in 3-D, but without explicit information on their geometry and connectivity.

We have developed a new approach based on a delineation algorithm (Fig. 4) that clusters a point cloud into elementary plant units such as internodes, petioles and leaves. The algorithm creates a graph (points + edges) to recover plausible neighbouring relationships between the points and embeds this graph in a spectral space in order to segment the point-cloud into meaningful elementary plant units.

We have presented this work at the 7th International Conference on Functional—Structural Plant Models (FSPM) which took place in Finland this summer [21].

5.1.5. Fixing Normal Constraints for Generation of Polycubes

Participants: Nicolas Ray, Dmitry Sokolov.

A polycube is a piecewise linearly defined surface where all faces are squares that are perpendicular to an axis of a global basis. Deforming triangulated surfaces to polycubes provides maps (form the original surface to the polycube) that can be used for a number of applications including hex-meshing. To define such a deformation, it is necessary to determine, for each point of the original surface, what will be its orientation (global axis) in the polycube.

../../../../projets/alice/IMG/plant-pipeline.jpg

Figure 4. From a point cloud scan of a plant to a segmentation of its leaves.

This problem is actually tackled by heuristics that basically affect the closest global axis to the surface normal. Coupled with a mesh deformation as pre-processing and some fixing rules as a post-processing, it is able to provide nice results for a number of surfaces. However, nothing ensures that the surface can be deformed to a polycube having these desired face orientations.

We have worked on a method able to determine if there exists a deformation of the surface that respects a given orientation constraint on each point. We have also designed an automatic solution that can fix constraints that would prevent the existence of a deformation into a polycube (Figure 5).

This study has highlighted that the constraints on desired orientation are global and requires constrained optimization methods to be solved. Our current solution is able to manage many cases where previous works would fail, but we can still produce some complex cases where interactions between dimension may lead to deadlocks.

../../../../projets/alice/IMG/polycube.png

*Figure 5. **Upper row:** the surface is deformed to make its normals closer to major axis, but to reach an equality, we need to have a coherent "wished orientation" of the faces. **Middle row:** we define a valid deformation into a polycube by editing the "wished orientation". **Lower row:** the resolution is performed a dimension at a time.*

5.1.6. Some Basic Geometric Considerations in Variational Multiview Stereo

Participant: Rhaleb Zayer.

We developed a technique for processing correspondences originating from dense variational matching in the context of multiview stereo. Such data tends to be very large and can easily encompass tens or hundreds of millions of points, these figures keep growing as high resolution images are becoming mainstream. Inspired by Lambert's cosine law, we regard the matching as sequences of planar maps across neighboring views, and show how to take advantage of geometric properties of such maps to favor image areas where the cosine angle between the surface normal and the line of sight is maximal. As the approach operates in the planar domain on smaller subsets of neighboring views, it is computationally efficient and has a low memory footprint. A preprint is in preparation.

5.1.7. *Multi-frontal Propagation Based Matching*

Participants: Rhaleb Zayer, Patricio Galindo.

We consider the propagation-based matching problem, which deals with expanding a limited set of correspondences towards a quasi-dense map across two views. Two issues which have not received much interest in earlier work are raised here. The traversal of weakly textured regions is shown to negatively impact the quality of subsequent correspondences. Analysis of the propagation results using the commonly adopted global best-fit strategy reveals that only a small subset of the input seeds contributes effectively to the propagation, which is probably not optimal since the quality of the matches may deteriorate as the propagation region becomes significantly large as shown in figure 7 -bottom. This research extends existing propagation techniques in two ways: (i) The selection of reliable expansion regions is automatized and adapted to the propagation by categorizing the image into three regions, no-propagation regions, safe-propagation regions and buffer-regions where seeds can propagate but cannot generate new seeds. (ii) A multi-frontal propagation approach is proposed with emphasis on the balance between the greedy nature of the original algorithm and the contribution of the seeds. A preprint is in preparation.

5.1.8. *Large Deformations of Slender Objects*

Participant: Rhaleb Zayer.

We studied the problem of large spatial deformation in the context of interactive editing of slender curve-like objects. The deformation is analyzed in the local frame of the individual curve segments (beams) and the rigid motion of the local frame is updated using a total Lagrangian approach. Analysis of the virtual work in the light of this decoupling allows formulating the Hessian of the deformation in a simple but principled manner. The resulting representation is sparser than existing derivations and can handle the simultaneous action of torques, and forces, efficiently, so as to reproduce a natural behavior in such path dependent situations. The proposed approach is conceptually simple, easy to implement, and suitable for object editing. The numerical solution is carried out using an efficient iterative scheme which allows stable convergence. A preprint is in preparation.

5.2. Computer Graphics

5.2.1. *By-example Synthesis of Curvilinear Structured Patterns*

Participants: Anass Lasram, Sylvain Lefebvre.

Many algorithms in Computer Graphics require to synthesize a pattern along a curve. This is for instance the case with line stylization, to decorate objects with elaborate patterns (chains, laces, scratches), or to synthesize curvilinear features such as mountain ridges, rivers or roads. We describe a simple yet effective method for this problem. Our method addresses the main challenge of maintaining the continuity of the pattern while following the curve. It allows some freedom to the synthesized pattern: It may locally diverge from the curve so as to allow for a more natural global result. This also lets the pattern escape areas of overlaps or fold-overs. This makes our method particularly well suited to structured, detailed patterns following complex curves. Our synthesizer copies tilted pieces of the exemplar along the curve, following its orientation. The result is optimized through a shortest path search, with dynamic programming. We speed up the process by an efficient parallel implementation. Finally, since discontinuities may always remain we propose an optional post-processing step optimally deforming neighboring pieces to smooth the transitions.

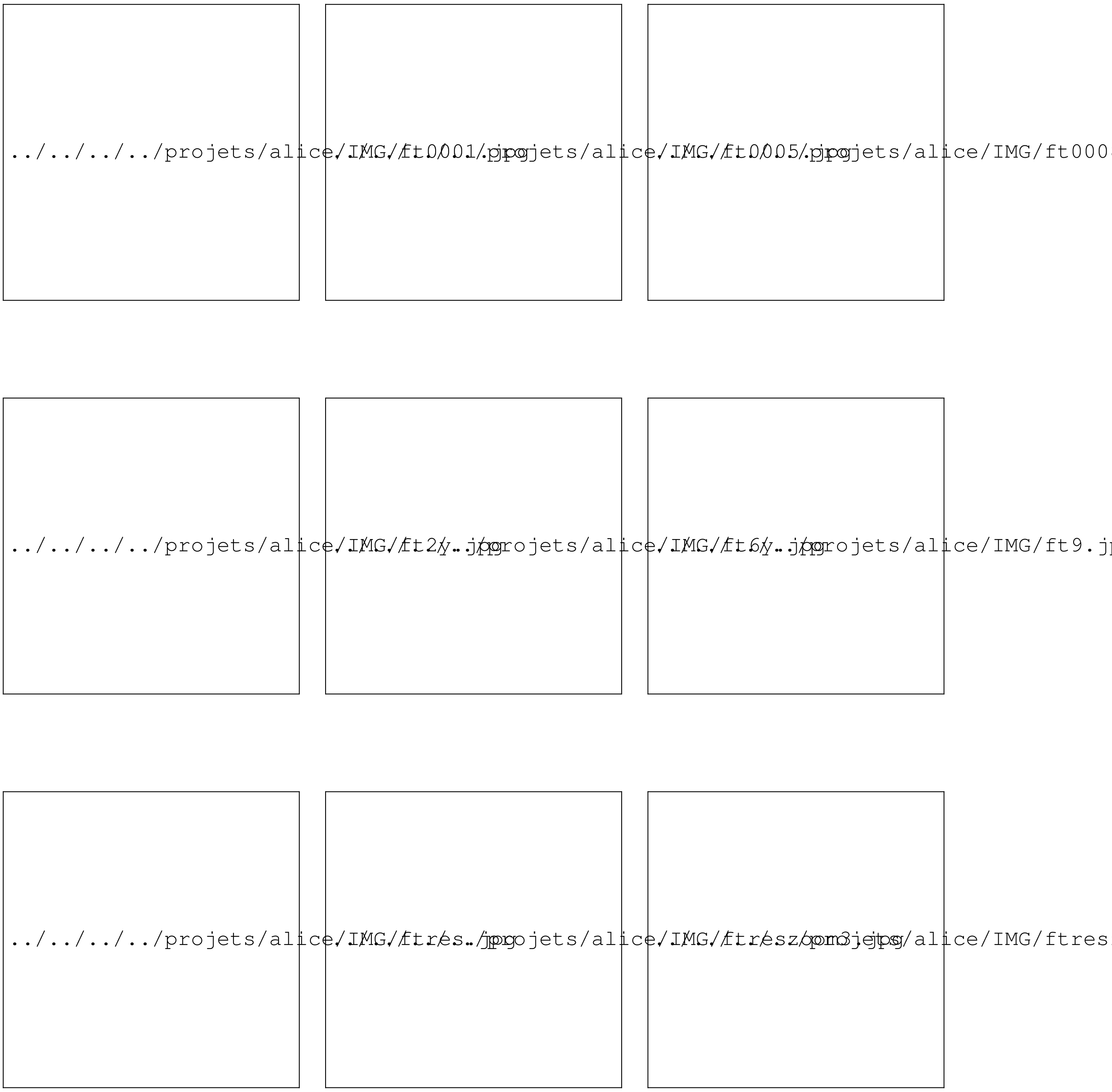


Figure 6. Processing best viewed regions in the Fountain data set (top). Each view represents the central image of a triplet (other two images not shown). The red-colored regions (middle) represent areas best viewed in the triplet. Yellow-marked regions represent regions which are only visible in the triplet and therefore are included even if they do not comply with the best view requirement. The resulting reconstruction (bottom) shows an almost outlier free point cloud.

../../../../projets/alice/IMG/final11s.jpg ../../../../projets/alice/IMG/final22s.jpg

../../../../projets/alice/IMG/nice1originals.jpg ../../../../projets/alice/IMG/nice2originals.jpg

Figure 7. Typical result of our approach (top), compared to a best-first strategy (bottom). In both experiments, the same initial seeds were used (≈ 40 seeds). The descendants of each initial seed are uniquely colored. Our approach clearly allows all seeds to contribute, whereas the greedy approach marginalize a majority of them.

Figure 8. Typical editing examples of slender objects under various constraint, the faded snapshots shows initial or intermediate configurations.

This work was presented at the Eurographics conference and published in the “Computer Graphics Forum” journal [20].

5.2.2. Game Level Layout

Participant: Sylvain Lefebvre.

This work is a collaboration with the University of British Columbia. We consider a long standing problem in the video game industry: How to automatically generate game levels. Most procedural game levels tend to exhibit a random organization, reducing their interest. Instead, our approach lets a professional game designer describe the global organization of the level through a planar graph, capturing the connectivity and sequencing of different level ’rooms’. Our approach then automatically generates multiple level geometries that correspond to this high-level description.

The work will be presented at *Eurographics 2014* [13].

5.2.3. Dynamic Element Textures

Participant: Sylvain Lefebvre.

This work is a collaboration with Microsoft Research Asia. We consider the problem of synthesizing animated details from an example. We first define the notion of a ’textured’ animation and extract details from the example animation. Intuitively, these are small scale repetitive motions found for instance for leaves in the wind or in swarms. We then propagate these motions to a coarse scale animation. Our techniques work on 1D, 2D and 3D objects.

We published this work in *ACM Transactions on Graphics (SIGGRAPH proceedings)* [14].

5.2.4. Make It Stand: Balancing Shapes for 3D Fabrication

Participant: Sylvain Lefebvre.

This work is a collaboration with ETH Zurich. We consider the problem of balancing 3D models so that they stand in static equilibrium on their base of support after printing. We formulate the problem as the joint optimization of a voxel selection inside the model and a continuous detail preserving deformation of the outter surface.

The work has been published in *ACM Transactions on Graphics (SIGGRAPH proceedings)* [18].

5.2.5. Clean Colors

Participants: Jean Hergel, Sylvain Lefebvre.

../../../../projets/alice/IMG/Game-level-layout.png

Figure 9. Game Level Layout

../../../../projets/alice/IMG/Dynamic-element-texture.jpg

Figure 10. Dynamic Element Textures

In this work we consider the problem of tool path planning for low-cost FDM (Fused Deposition Modeling) printers when using multiple filaments. Our method is based on three components which together reduce most of the defects found in such prints. Our algorithm first optimizes the orientation (azimuth angle) of the print so as to minimize defects. It then builds a rampart in close proximity of the model. This captures most of the strings of plastic oozing from idle extruders. Finally, we optimize for navigation paths minimizing the apparition of defects.

The work will be presented at *Eurographics 2014* [22].

5.2.6. Fast Fragment Sorting on the GPU

Participants: Sylvain Lefebvre, Samuel Hornus.

In this work, we build upon our result on "hashing on the GPU" from 2011 [1] to develop new techniques for sorting per-pixel lists of fragments as the latter are rasterized. We can then obtain, for each pixel, the list of surface elements visible through that pixel, sorted according to their distance to the viewpoint. The lists are obtained in a single rasterization pass instead of two for some earlier work; this is a clear win for bandwidth usage and processing time. Two important applications are the possibility to correctly visualize transparent objects and to directly display constructive-solid-geometry models without having to compute their boundary first (the boolean operations are performed on the fly, per pixel).

Our initial work has been published as a research report [25]. It has then been extended into a book chapter [24].

The techniques developed in this work are extensively used in our 3D printing sotfware IceSL (see section 4.2).

5.2.7. Techniques for Shooting Highly Coherent Rays

Participant: Samuel Hornus.

This work explores novels ways to exploit the coherence of some set of rays used in the ray-tracing and other realistic image synthesis techniques. We propose new ways to traverse the usual data-structure for 3D indexes and leverage optimized and exact geometric predicates. Our first results give a faster ray shooting technique for pinhole camera rays and exhibit a remarkable increase in efficiency as the number of rays rises. A manuscript was submitted but not accepted to Eurographics.

5.3. Algorithms and analysis

Participant: Laurent Alonso.

5.3.1. The Majority Problem

Given a set of n elements each of which is either red or blue, Boyer and Moore's algorithm uses pairwise equal/not equal color comparisons to determine the majority color. We analyze the average behavior of their algorithm, proving that if all 2^n possible inputs are equally likely, the average number of color comparisons used is $n - \sqrt{2n/\pi} + O(1)$ and has variance $\frac{\pi-2}{\pi}n - \frac{\sqrt{2\,n}}{\sqrt{\pi}} + O(1)$. This joint work with Edward M. Reingold was published in the IPL journal [8].

5.3.2. The $X + Y$ Sorting Problem

Some combinatorial approaches were taken to try to find bounds on the $X + Y$ problem: Given two lists: $X = (x_1, ..., x_n)$, $Y = (y_1, ..., y_m)$, determine the ordering of the values $x_i + y_j$ for $i \in [1, n], j \in [1, m]$.

5.4. Fractal Geometry

Participant: Dmitry Sokolov.

Fractal geometry is a relatively new branch of mathematics that studies complex objects of non-integer dimensions. It finds applications in many branches of science as objects of such complex structure often exhibit interesting properties.

In 1988 Barnsley presented the Iterative Function System (IFS) model that allows modelling complex fractal shapes with only a limited set of contractive transformations. Later many other models were based on the IFS model such as Language-Restricted IFS, Projective IFS, Controlled IFS and Boundary Controlled IFS. The latter allows modeling complex shapes with control points and specific topology. These models cover classical geometric models such as B-splines and subdivision surfaces as well as fractal shapes.

This year we focused on the analysis of the differential behaviour of the shapes described with Controlled IFS and Boundary Controlled IFS. We derive the necessary and sufficient conditions for differentiability for everywhere dense sets of points. Our study is based on the study of the eigenvalues and eigenvectors of the transformations composing the IFS.

We apply the obtained conditions to modeling curves in surfaces. We describe different examples of differential behaviour presented in shapes modeled with Controlled IFS and Boundary Controlled IFS. We also use the Boundary Controlled IFS to solve the problem of connecting different subdivision schemes. We construct a junction between Doo-Sabin and Catmull-Clark subdivision surfaces and analyse the differential behaviour of the intermediate surface.

An article about this work is in the publication process in LNCS.

5.5. Scientific Computing for Linear and Nonlinear Wave Problems

Participant: Xavier Antoine.

We consider the Backward Euler SPectral (BESP) scheme that was proposed for computing the stationary states of Bose-Einstein Condensates (BECs) through the Gross-Pitaevskii equation. We show that the fixed point approach introduced earlier fails to converge for fast rotating BECs. A simple alternative approach based on Krylov subspace solvers with a Laplace or Thomas-Fermi preconditioner is given. Numerical simulations (obtained with the associated freely available Matlab toolbox GPELab) for complex configurations show that the method is accurate, fast and robust for 2D/3D problems and multi-components BECs.

This work was published in the journal “Journal of Computational Physics” [9].

5.6. Accelerating Structural Biology Software

Participant: Xavier Cavin.

This work is a collaboration with Dave Ritchie (team ORPAILLEUR, Nancy). The aim of this project is to leverage parallelism, multi-core computing and GPU in order to speed-up costly computations in cryo-electron microscopy. Several tools have been developped. Two of those “gEM tools” have been the subject of two articles were published in 2013 in “Journal of Structural Biology” [10] and “BMC Structural Biology” [11].

ALPAGE Project-Team

6. New Results

6.1. Unsupervised segmentation of Mandarin Chinese

Participants: Pierre Magistry, Benoît Sagot.

In Chinese script, very few symbols can be considered as word boundary markers. The only easily identifiable boundaries are sentence beginnings and endings, as well as positions before and after punctuation marks. Although the script doesn't rely on typography to define (orthographic) "words", a word-level segmentation is often re- quired for further natural language processing, which is a highly non-trivial task.

A great variety of methods have been proposed in the literature, mostly in supervised machine learning settings. Our work addresses the question of unsupervised segmentation, i.e., without any manually segmented training data. Although supervised learning typically performs better than unsupervised learning, we believe that unsuper- vised systems are worth investigating as they require less human labour and are likely to be more easily adaptable to various genres, domains and time periods. They can also provide more valuable insight for linguistic studies.

Amongst the unsupervised segmentation systems described in the literature, two paradigms are often used: Branching Entropy (BE) and Minimum Description Length (MDL). The system we have developed relies on both. We have introduced a new algorithm [22] which searches in a larger hypothesis space using the MDL criterion, thus leading to lower Description Lengths than other previously published systems. Still, this improvement concerning the Description Length does not come with better results on the Chinese word segmentation task, which raises interesting issues. However, it turns out that it is possible to add very simple constraints to our algorithm in order to adapt it to the specificities of Mandarin Chinese in a way that leads to results better than the state-of-the-art on the Chinese word segmentation task.

Moreover, an important part of discrepancies between the various segmentation guidelines concerns the so-called "factoids." This term covers a variety of language phenomena that include: numbers, dates, addresses, email addresses, proper names, and others. We have shown that specific treatment of a subset of such expressions is both sound (as factoids to not resort to general language, which we try and capture with our segmentation model, both rather to conventions that are easy to encode as rules). By augmenting the local grammars of SxPipe to deal with the aforementioned expressionsin Chinese, and use them as a pre-processing for our task, we can discard the matched expressions from the training data and segment them accordingly to the guidelines as a post-processing step. Our results show a significant improvement over previous results.

6.2. Dynamic extension of a French lexical resources based a text stream

Participants: Damien Nouvel, Benoît Sagot, Rosa Stern, Virginie Mouilleron, Marion Baranes.

Lexical incompleteness is a recurring problem when dealing with natural language and its variability. It seems indeed necessary today to regularly validate and extend lexica used by tools processing large amounts of textual data. This is even more true when processing real-time text flows. In this context, we have introduced two series of techniques for addressing words unknown to lexical resources, and applied them to French within the context of the EDyLex ANR project:

- **Extending a morphological lexicon** We have studied neology (from a theoretic and corpus-based point of view) and developed modules for detecting neologisms in AFP news wires in real time and inferring information about them (lemma, category, inflectional class) [24]. We have shown that we are able, using among others modules for analyzing derived and compound neologisms, to generate lexical entries candidates in real time and with a good precision, to be added in the Le*fff* lexicon.
- **Extending an entity database** We have also extended our previous work on named entities detection and linking in order to be able to extract new named entities from AFP news wires and create candidate entries for the *Aleda* entity database.

6.3. Transferring lexical knowledge from a resourced language to a closely-related resource-free language

Participants: Yves Scherrer, Benoît Sagot.

We have developed a generic approach for the transfer of part-of-speech (POS) annotations from a resourced language (RL) towards an etymologically closely related non-resourced language (NRL), without using any bilingual (i.e., parallel) data. We rely on two hypotheses. First, on the lexical level, the two languages share a lot of cognates, i.e., word pairs that are formally similar and that are translations of each other. Second, on the structural level, we admit that the word order of both languages is similar, and that the set of POS tags is identical. Thus, we suppose that the POS tag of one word can be transferred to its translational equivalent in the other language.

The proposed approach consists of two main steps. In the first step, we induce a translation lexicon from monolingual corpora. This step relies on several methods, including a character-based statistical machine translation model to infer cognate pairs, and 3-gram and 4-gram contexts to infer additional word pairs on the basis of their contextual similarity. This step yields a list of $< w_{\mathrm{NRL}}, w_{\mathrm{RL}} >$ pairs. In the second step, the RL lexicon entries are annotated with POS tags with the help of an existing resource, and these annotations are transferred onto the corresponding NRL lexicon entries. We complete the resulting tag dictionary with heuristics based on suffix analogy. This results in a list of $< w_{\mathrm{NRL}}, t >$ pairs, covering the whole NRL corpus.

We have evaluated our methods on several language pairs. We have worked among others onfive language pairs of the Iberic peninsula, where Spanish and Portuguese play the role of RLs: Aragonese–Spanish, Asturian–Spanish, Catalan–Spanish, Galician– Spanish and Galician–Portuguese [27]. We have also conducted experiments on germanic [28] and slavic languages. We have also applied it in a slightly different context, in collaboration with Tomaž Erjavec (IJS, Slovenia), namely that of inducing resources for historical Slovene based on existing resources for contemporary Slovene [26]. Although no direct comparison can be performed, because of the novelty of the task, our results are very satisfying in so far that they are almost as high as published result on a related but simpler task, that of unsupervized part-of-speech tagging — which, contrarily to our work, relies on an existing morphological lexicon for the language at hand.

6.4. Building a large-scale translation graph

Participants: Valérie Hanoka, Benoît Sagot.

Large-scale general-purpose multilingual translation databases are useful in a wide range of Natural Languages Processing (NLP) tasks. This is especially true concerning researches tackling problems specific to under-resourced languages, as translation databases can be used for adapting existing resources in other languages. This has been applied for example for the development of wordnets in languages other than English. There is thus a real need in NLP for *open-source* multilingual lexical databases that compiles as many translations as can be found on any freely available resource in any language.

We have developed, and are about to release, a new open-source heavily multilingual (over 590 languages) translation database built using several sources, namely various wiktionaries and the OPUS parallel corpora.

Our graph was built in several steps. We first extracted a preliminary set of translation and synonym pairs, which we stored in a large translation and synonym graph. We then applied filtering techniques for increasing the accuracy of this graph. We have evaluated the accuracy of our graph as being as high as 98% for translations extracted from wiktionaries.

6.5. Computational morphology

Participant: Benoît Sagot.

In 2013, following previous collaborative work [92], [105], we have designed and developed Alexina$_{\text{PARSLI}}$ in collaboration with Géraldine Walther (LLF and DDL), a formalism for encoding inflectional descriptions (lexicon and grammar) that aims at filling the gap between morphologically and typologically motivated approaches on the one hand and implemented approaches on the other hand, as will be discussed in the remainder of this section. Indeed, Alexina$_{\text{PARSLI}}$ is both:

- an **implementation formalism for PARSLI**, a formal model of inflectional morphology [106] that accounts for concepts underlying the canonical approach of morphological typology;
- an **extension of the Alexina lexical framework** developed at Alpage for modeling lexical information and developing lexical resources. The Alexina framework now supports both morphological grammars that use the original Alexina morphological formalism as well as new grammars developed in Alexina$_{\text{PARSLI}}$.

The Alexina$_{\text{PARSLI}}$ formalism and tools have been proven greatly beneficial to works both in descriptive and formal morphology, in particular in studies about Latin passivisation and Maltese verbal inflection [106] and in studies comparing the compacity of morphological descriptions [106], [92], [105], as well as in NLP, for the efficient development of a large-scale and linguistically sound morphological lexicon for German (a paper describing this new lexicon is to be presented at the LREC 2014 conference).

In collaboration with Géraldine Walther and Guillaume Jacques (CRLAO, CNRS), within operation LR4.11 from strand 6 of the LabEx EFL, we have also developed two Alexina$_{\text{PARSLI}}$ descriptions of (part of) the Khaling (Kiranti, Sino-Tibetan) verbal inflectional system, together with a medium-scale lexicon. Our study shows that an explicit account for the so-called direct-inverse marking, based on concepts developed within PARSLI, allows for a more compact account of this inflectional system [42].

6.6. Extracting Derivational Relations from an Inflectional Lexicon

Participants: Marion Baranes, Benoît Sagot.

Derivational morphological can provide useful information for natural language processing tasks. Indeed, it can improve any application which has to deal with unknown words such as information extraction, spell-checking and others.

We define a morphological family as a set of semantically related lexical entries which differ by their prefix and/or suffix, thus limiting ourselves to concatenative derivational morphology. We shall denote as derivationally related two morphological lexical entries that belong to the same morphological family.

We have developed a system which performs an analogy-based unsupervized extraction of weighted transformation rules that relate derivationally related lexical entries, and use these rules for extracting derivational relations within an existing inflectional lexicon. Our transformation rules can also be used to infer morphological information (both inflectional and derivational) for wordforms unknown to the inflectional lexicon. Our system is language-independent, although restricted to concatenative derivational morphology. We have evaluated it on four languages: English, French, German and Spanish. Our results will be published at the LREC 2014 conference.

6.7. Improving post-OCR correction with shallow linguistic processing

Participants: Kata Gábor, Benoît Sagot.

Providing wider access to national cultural heritage by massive digitalization confronts the actors of the field to a set of new challenges. State of the art optical character recognition (OCR) software currently achieve an error rate of around 1 to 10% depending on the age and the layout of the text. While this quality may be adequate for indexing, documents inteded for reading need to meet higher standards. A reduction of the error rate by a factor of 10 to 100 becomes necessary for the diffusion of digitalized books and journals through emerging technologies such as e-books.

Within the PACTE project, an "Investissements d'avenir" project led by the Numen company, we have worked on the automatic post-processing of digitalized documents in the aim of reducing the OCR error rate by using contextual information and linguistic processing, by and large absent from current OCR engines. At the current stage of the project, we are focusing on French texts coming from the archives of the French National Library (Bibliothèque Nationale de France).

We adopted a hybrid approach, making use of both statistical classification techniques and linguistically motivated modules to detect OCR errors and generate correction candidates. The technology is based on the noisy chanel model, widely used in the field of machine translation and spelling correction and subsequently in OCR post-correction. As to linguistically enhanced models, POS tagging was succesfully applied to spelling correction. However, to our knowledge, little work has been done to exploit linguistic analysis for post-OCR correction.

We have proposed to integrate a shallow processing module to detect certain types of named entities, and a POS tagger trained specifically to deal with NE-tagged input. Our studies demonstrate that linguistically informed processing can efficiently contribute to reduce the error rate by 1) detecting false corrections proposed by the statistical correction module, 2) detecting a certain amount of OCR errors not detected by the statistical correction module.

6.8. Named Entity Linking

Participants: Rosa Stern, Benoît Sagot.

The Ph.D. research work started in 2009 lead in 2013 to the development of a joint entity recognition and linking system for the processing of textual data at the Agence France Presse (AFP).

This system, Nomos, allows to use any existing named entity recognition system, as well as combinations of such systems; their results are passed to a linking module in charge of the association between each detected mention and a unique reference within an existing data inventory. The two tasks (recognition and linking) are jointly operated: the recognition module presents a set of possible detections, which are further disambiguated by the linking module concurrently to the search for the best linking solution to each mention. This joint approach is justified by the need to limit the error propagation between two such modules in a pipeline system.

Experiments were achieved in order to evaluate the performance of Nomos over AFP news wires. They showed that the joint approach, relatively to a purely sequential one, improves the system's global precision, i.e. the linking accuracy as well as the named entity recognition task itself. A gain of 3 points (87,6) is observed for the recognition precision with a low recall loss, while a gain of 8 points (92,9) is observed when several recognition systems are combined - although with a more significant loss of recall.

The Nomos system also allows to anchor of the AFP's textual production in the Linked Data network and the Semantic Web paradigm, since the annotations derived from the entity linking associate each entity to an identified resource in repositories such as Wikipedia, DBPedia, Geonames or the New York Times Linked Data.

6.9. Treebanking at Alpage

Participants: Djamé Seddah, Benoît Sagot, Marie-Hélène Candito, Corentin Ribeyre, Benoît Crabbé, Éric Villemonte de La Clergerie, Virginie Mouilleron, Vanessa Combet.

Since the advents of supervized methods for building accurate statistical parsing models, treebank engineering has become of crucial importance. In fact building a treebank, namely a set of carefully annotated syntactic parses with possibly different annotation layers and covering potentially different text domains, can be seen as providing a parser with both a grammar and a set of probabilities used for disambiguation. The main problem of such approaches lies in the nature of the lexical probabilities: they force the parsing model to be extremely sensitive to its training data and hence limit its performance to some low upper-bound when applied in out-of-domain scenario.

6.9.1. Written French Treebanks

Originating from the merging of two NLP teams specialized in grammar engineering and in which the creation of the first treebank for French was initiated [46], it is no wonder that we decided to increase the coverage of our French Treebank-based parsers by building out-of-domain treebanks: the Sequoia Corpus, [55], [18], made from Europarl, biomedical and wikipedia data, and the French Social Media Bank (outside English, the first data set covering Facebook, Twitter and other social media noisy text data) [95], [96]. We built those two corpus for two purposes: first, we wanted to evaluate the performance of our nlp chains (tokenization, tagging, parsing) on out-of-domain data, being noisy or not ; then we increased the coverage of our French treebank based models by simply adding those new data set to the canonical training set (using of-course many lexical variation, morphological clustering, brown clustering, etc.). We're also on the process of finalizing a new 2600 sentence data set, made essentially of questions, which are strikingly absent from all the treebanks we've been using and developing. So far, only one such data set exist and only for English: the Question-Bank [66]. Our very preliminary results show that simply adding a third of that corpus to the French Treebank greatly improve our parser performance.

Finally, Alpage is leading, in collaboration with the Nancy-based team Calligrame, a project to annotate the Sequoia corpus and the French Treebank with a richer, "deeper" syntactic layer, at the interface between syntax and semantics. A paper describing this effort is to appear at the LREC 2014 conference.

6.9.2. Spoken French Treebank

In collaboration with Anne Abeillé (LLF, CNRS), we have also contributed to the deign of a spoken treebank for French based on data produced in the ANR ETAPE. Contrary to other languages such as English, where spoken treebanks such as the Switchboard corpus treebank (Meteer, 1995), there is no sizable spoken corpus for French annotated for syntactic constituents and grammatical functions. Our project is to build such a resource which will be a natural extension of the Paris 7 treebank (Abeillé et al. 2003) for written French, in order to be able to compare with similar annotations written and spoken French. We have reused and adapted the parser (Petrov et al., 2006) which has been trained on the written treebank, with manual correction and validation. The first results are promising [32].

6.10. Linear time constituent parser

Participant: Benoît Crabbé.

We have designed an efficient and accurate lexicalized LR inspired discriminative parsing algorithm that recasts some current advances in dependency parsing to the constituency setting. We specifically designed and evaluated a Graph Structured Stack-based parser (Huang et al. 2010) using some additional specific approximate inference techniques such as the max violation update for the perceptron (Huang et al. 2012) . By contrast with dependency parsing however, lexicalized constituent parsing raises some additional correctness issues that motivate the explicit use of an LR automata instead of a simpler shift reduce framework.

The parsing model is linear in time and has been evaluated on French data, where it turns out to be state of the art on SPMRL 2013 datasets [29] both in time and in accuracy. The parsing framework has been designed to be further extended with compositional semantic representations and allows in principle an easy integration of ressources — such as those developped in the team — considered to be important for parsing morphologically rich languages.

6.11. Improving FRMG through partially supervised learning

Participant: Éric Villemonte de La Clergerie.

Since the emergence of several statistical parsers for French developed on the French TreeBank (FTB), including those developed at Alpage, it was important to be able to compare the symbolic meta-grammar-based parser FRMG with these statistical parsers on their native treebank, but also possibly to extend the comparison for other treebanks.

A first necessary step in this direction was a conversion from FRMG's native dependency scheme into FTB's dependency scheme, a tedious task highlighting the differences in design at all levels (segmentation, parts of speech, representation of the syntactic phenomena, etc.). A preliminary evaluation has shown that accuracy is good, but largely below the scores reached by the statistical parsers.

A challenge was then to explore if training on the FTB could be used to improve the accuracy of a symbolic parser like FRMG. However, the main difficulty arises from the fact that FTB's dependency scheme has little in common with FRMG's underlying grammar, and that no reverse conversion from FTB to FRMG structures is available. Such a conversion could be investigated but would surely be difficult to develop. Instead, we tried to exploit directly FTB data, using only very minimal assumptions, nevertheless leading to important gains and results close to those obtained by the statistical parsers [31]: it was possible to tune the disambiguisation process of FRMG and strongly increase its accuracy, from 83% up to 87.17% (in terms of CONLL Labeled Attachment Score), a level comparable to those reached by statistical parsers trained on the FTB. Preliminary experiments show that (a) disambiguisation tuning also improve the performances on other corpora and (b) that FRMG seems to be more stable than statistical parsers on corpora other than the FTB. Finer-grained comparison of FRMG wrt statistical parsers have been done that provide some insight for further improvements of FRMG.

The interest is that the technique should be easily adaptable for training data with different annotation schemes. Furthermore, our motivation was not just to improve the performances on the FTB and for the annotation scheme of FTB, for instance by training a reranker (as often done for domain adaptation), but to exploit the FTB to achieve global improvement over all kinds of corpora and for FRMG native annotation scheme.

6.12. Statistical parsing of Morphologically Rich Languages

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

6.12.1. The SPMRL shared task

Since several years, Djamé Seddah, together with Marie-Hélène Candito and more generally the whole Alpage team, has played a major role in setting up and animating an international network of researchers focusing on parsing morphologically rich languages (MRLs).

In 2013, Djamé Seddah led the organization of the first shared task on parsing MRLs, hosted by the fourth SPMRL workshop and described in a 36-page overview paper that constitutes an in-depth state-of-the-art analysis and review of the domain [29]. The primary goal of this shared task was to bring forward work on parsing morphologically ambiguous input in both dependency and constituency parsing, and to show the state of the art for MRLs. We compiled data for as many as 9 languages, which represents an immense scientific and technical challenge.

6.12.2. DyALog-SR

The SPMRL 2013 shared task was the opportunity to develop and test, with promising results, a simple beam-based shift-reduce dependency parser on top of the tabular logic programming system DYALOG. We used (Huang and Sagae, 2010) as the starting point for this work, in particular using the same simple arc-standard strategy for building projective dependency trees. The parser was also extended to handle ambiguous word lattices, with almost no loss w.r.t. disambiguated input, thanks to specific training, use of oracle segmentation, and large beams. We believe that this result is an interesting new one for shift-reduce parsing.

The current implementation scales correctly w.r.t. sentence length and, to a lesser extent, beam size. Nevertheless, for efficiency reasons, we plan to implement a simple C module for beam management to avoid the manipulation in DYALOG of sorted lists. Interestingly, such a module, plus the already implemented model manager, should also be usable to speed up the disambiguation process of DYALOG-based TAG parser FRMG (de La Clergerie, 2005a). Actually, these components could be integrated in a slow but on-going effort to add first-class probabilities (or weights) in DYALOG, following the ideas of (Eisner and Filardo, 2011) or (Sato, 2008).

6.12.3. The Alpage-LIGM French parser

The second Alpage system that participated to the SPMRL shared task, although on French language only, was developed in collaboration with Mathieu Constant (LIGM), based on the Bonsai architecture. This system is made of several single statistical dependency parsing systems whose outputs are combined into a reparser. We use two types of single parsing architecture: (a) pipeline systems; (b)"joint" systems.

The pipeline systems first perform multi-word expression (MWE) analysis before parsing. The MWE analyzer merges recognized MWEs into single tokens and the parser is then applied on the sentences with this new tokenization. The parsing model is learned on a gold training set where all marked MWEs have been merged into single tokens. For evaluation, the merged MWEs appearing in the resulting parses are expanded, so that the tokens are exactly the same in gold and predicted parses.

The "joint" systems directly output dependency trees whose structure comply with the French dataset annotation scheme. Such trees contain not only syntactic dependencies, but also the grouping of tokens into MWEs, since the first component of an MWE bears dependencies to the subsequent components of the MWE with a specific label. At that stage, the only missing information is the POS of the MWEs, which we predict by applying a MWE tagger in a post-processing step.

This parsing system obtains the best results for French, both for overall parsing and for MWE recognition, using a reparsing architecture that combines several parsers, with both pipeline architecture (MWE recognition followed by parsing), and joint architecture (MWE recognition performed by the parser).

6.13. Towards a French FrameNet

Participants: Marie-Hélène Candito, Marianne Djemaa, Benoît Sagot, Éric Villemonte de La Clergerie, Laurence Danlos.

The ASFALDA project [1] is a three-year project which started in October 2012, with the objective of building semantic resources (generalizations over predicates and over the semantic arguments of predicates) and a corresponding semantic analyzer for French. We chose to build on the work resulting from the FrameNet project [47], [2] which provides a structured set of prototypical situations, called *frames*, along with a semantic characterization of the participants of these situations (called *frame elements*, FEs). The resulting resources will consist of :

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

In the first year of the project, we focused on the first of these objectives. A team of 10 active members, from Alpage, the Laboratoire de Linguistique Formelle (LLF), the MELODI team (IRIT - Toulouse) and the CEA-List partners achieved :

- the delimitation and adaptation to French of a set of FrameNet frames, in order to cover a set of specific notional domains (commercial transaction, communication, cognitive positions, judgment/evaluation, temporal relations, spatial position, causality).
- and the semi-automatic construction of a French lexicon in which French lexical units are associated with frames

The current resource contains 110 frames, and roughly 2500 lexical units / frame pairs. The next phase consists in automatic pre-annotation of semantic annotations, that will serve as basis for the manual validation phase.

Note that a publication describing the project and these first achievements shall be presented at the LREC 2014 conference.

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

6.14. Modelisation of discourse structures with DSTAG

Participant: Laurence Danlos.

This work was done within the ANR Polymnie, in collaboration with Sylvain Pogodalla and Philippe de Groote from LORIA.

Neg-Raising (NR) verbs form a class of verbs with a clausal complement that show the following behavior: when a negation syntactically attaches to the matrix predicate, it can semantically attach to the embedded predicate. Such an implication does not always hold. Some contexts make it impossible to consider the negation as having scope over the embedded predicate only. This corresponds to the non-NR reading of the predicate.

We have developed and published [20] an account of NR predicates within Tree Adjoining Grammars (TAG) that relies on a Montague-like semantics for TAG. The different properties of NR predicates are rendered at different levels: the ambiguity of the readings is modeled by lexical ambiguity; the scoping and cyclicity properties are modeled through the lexical semantics and the higher-order interpretation of adjunction nodes; spurious am- biguities are avoided using fine-grained types for terms representing derivation trees. This provides us with a base layer where to account for interactions with discourse connectives and discourse representation represented in DSTAG.

6.15. Annotation of discourse structures on the FTB

Participants: Laurence Danlos, Margot Colinet.

With the aim of annotating the French TreeBank (FTB, already annotated for syntax) with discourse information, we have been working on the first step of the project, namely identify all the occurrences of discourse connectives in the FTB. This raises problems for lexemes which are ambiguous with a discourse usage and other uses. In collaboration with Mathilde Dargnat (ATILF) and Grégoire Winterstein, we have been working on the preposition *pour* (around 1500 occurrences) and the adverb *alors* (300 occurrences). This work is the basis for a future annotation manual.

In parallel, we have been working on adverbial discourse connectives and published on the topic [17]. This paper focuses on the following question: does the only syntactic argument of an adverbial discourse connective correspond to its second semantic argument? It shows that this is not always the case, which is a problem for the syntax-semantics interface. This interface brings us to distinguish two classes of adverbial connectives we sketch the study of.

6.16. Pairwise coreference models

Participant: Emmanuel Lassalle.

In collaboration with Pascal Denis (Magnet, Inria), we have proposed a new method for significantly improving the performance of pairwise coreference models [34]. Given a set of indicators, our method learns how to best separate types of mention pairs into equivalence classes for which we construct distinct classification models. In effect, our approach finds an optimal fea- ture space (derived from a base feature set and indicator set) for discriminating coreferential mention pairs. Although our approach explores a very large space of possible feature spaces, it remains tractable by exploiting the structure of the hierarchies built from the indicators.

In the framework of decision trees, this method can be seen as a pruning procedure and thus can be combined with different methods for expanding a decision tree. It can also be compared to polynomial kernels, but has the advantage of a lower computational complexity [21]. Our experiments on the CoNLL-2012 Shared Task English datasets (gold mentions) indicate that our method is robust relative to different clustering strategies and evaluation metrics, showing large and consistent improvements over a single pairwise model using the same base features. Our best system obtains a competitive 67.2 of average F1 over MUC, B3, and CEAF which, despite its simplicity, places it above the mean score of other systems on these datasets.

6.17. Identification of implicit discourse relations

Participant: Chloé Braud.

In collaboration with Pascal Denis (Magnet, Inria), we have developed a system for identifying "implicit" discourse relations (that is, relations that are not marked by a discourse connective) [33]. Given the little amount of available annotated data for this task, our system also resorts to additional automatically labeled data wherein unambiguous connectives have been suppressed and used as relation labels, a method introduced by Marcu and Echihabi (2002). As shown by Sporleder and Lascarides (2008) for English, this approach doesn't generalize well to implicit relations as annotated by humans. We have shown that the same conclusion applies to French due to important distribution differences between the two types of data. In consequence, we propose various simple methods, all inspired from work on domain adaptation, with the aim of better combining annotated data and artificial data. We have evaluated these methods through various experiments carried out on the ANNODIS corpus: our best system reaches a labeling accuracy of 45.6%, corresponding to a 5.9% significant gain over a system solely trained on manually labeled data.

ALPINES Team

6. New Results

6.1. Integral equations on multi-screens

We developed a new functional framework for the study of scalar wave scattering by objects, called multi-screens, that are arbitrary arrangements of thin panels of impenetrable materials. From a geometric point of view, multi-screens are a priori non-orientable non-Lipschitz surfaces. We use our new framework to study boundary integral formulations of the scattering by such objects.

6.2. Second-kind Galerkin boundary element method for scattering at composite objects

In the context of scattering of time-harmonic acoustic waves at objects composed of several homogeneous parts with different material properties, a novel second-kind boundary integral formulation for this scattering problem was proposed in [X. Claeys, A single trace integral formulation of the second kind for acoustic scattering, Report 2011-14, SAM, ETH Zürich]. We recasted it into a variational problem set in L2 and investigated its Galerkin boundary element discretization from a theoretical and algorithmic point of view. Empiric studies demonstrate the competitive accuracy and superior conditioning of the new approach compared to a widely used Galerkin boundary element approach based on a first-kind boundary integral formulation.

6.3. Instability phenomenon for a rounded corner in presence of a negative material

We studied a 2D transmission problem between a positive material and a negative material. In electromagnetics, this negative material can be a metal at optical frequencies or a negative metamaterial. We highlighted an unusual instability phenomenon in some configurations: when the interface between the two materials presents a rounded corner, it can happen that the solution depends critically on the value of the rounding parameter. To prove this result, we provided an asymptotic expansion of the solution, when it is well-defined, in the geometry with a rounded corner. Then, we demonstrated that the asymptotic expansion is not stable with respect to the rounding parameter. We also conducted obtained numerical experiments with finite element methods to validate these results.

6.4. Parallel design and performance of direction preserving preconditioners

In the context of preconditioned iterative methods, our work has focused on so called direction preserving preconditioners. In [9] we consider the parallel design and performance of nested filtering factorization (NFF), a multilevel parallel preconditioning technique for solving large sparse linear systems of equations by using iterative methods. NFF is based on a recursive decomposition that requires first to permute the input matrix, which can have an arbitrary sparsity structure, into a matrix with a nested block arrow structure. This recursive factorization is a key feature in allowing NFF to have limited memory requirements and also to be very well suited for hierarchical parallel machines. NFF is also able to preserve some directions of interest of the input matrix A. Given a set of vectors T which represent the directions to be preserved, the preconditioner M satisfies a right filtering property $MT = AT$. This is a property which has been exploited in different contexts, as multigrid methods [Brandt et al., 2011, SIAM J. Sci. Comput.], semiseparable matrices [Gu et al, 2010, SIAM J. Matrix Anal. Appl.], incomplete factorizations [Wagner, 1997, Numer. Math] , or nested factorization [Appleyard and Cheshire, 1983, SPE Symposium on Reservoir Simulation]. It is well known that for difficult problems with heterogeneities or multiscale physics, the iterative methods can converge very slowly, and this is often due to the presence of several low frequency modes. By preserving the directions corresponding to these low frequency modes in the preconditioner, their effect on the convergence is alleviated and a much faster convergence is often observed. NFF can be seen as an extension of nested factorization that can be used for matrices with arbitrary sparsity structure and for which the computation can be performed in parallel. While the algebra of NFF has been introduced previously [Grigori et al, 2010, Inria tech. report], we relate the arithmetic complexity of NFF to the depth of recursion of its decomposition, and with our data distribution and implementation, we estimate its arithmetic and communication complexity. We also discuss the convergence of NFF on a set of matrices arising from the discretization of a boundary value problem with highly heterogeneous coefficients on three-dimensional grids. Our results show that on a $400 \times 400 \times 400$ regular grid, the number of iterations with NFF increases slightly while increasing the number of subdomains up to 2048. In terms of runtime performance on Curie, a Bullx system formed by nodes of two eight-core Intel Sandy Bridge processors, NFF scales well up to 2048 cores and it is 2.6 times faster than the domain decomposition preconditioner Restricted Additive Schwarz (RAS) as implemented in PETSc http://www.mcs.anl.gov/petsc/. The choice of the filtering vectors plays an important role in direction preserving preconditioners. There are problems for which we have prior knowledge of the near kernel of the input matrix, and this is indeed the case for the problems tested in this paper. They can also be approximated by using techniques similar to the ones used in deflation, however we do not discuss further this option here.

6.5. New resuls in communication avoiding algorithms for sparse linear algebra

In the context of sparse linear algebra algorithms, our recent results focus on two operations, incomplete LU factorization preconditioners and sparse matrix-matrix multiplication.

In [12] we present a communication avoiding ILU0 preconditioner for solving large linear systems of equations by using iterative Krylov subspace methods. Recent research has focused on communication avoiding Krylov subspace methods based on so called s-step methods. However there was no communication avoiding preconditioner available yet, and this represents a serious limitation of these methods. Our preconditioner allows to perform s iterations of the iterative method with no communication, through ghosting some of the input data and performing redundant computation. It thus reduces data movement by a factor of 3s between different levels of the memory hierarchy in a serial computation and between different processors in a parallel computation. To avoid communication, an alternating reordering algorithm is introduced for structured and unstructured matrices, that requires the input matrix to be ordered by using a graph partitioning technique such as k-way or nested dissection. We show that the reordering does not affect the convergence rate of the ILU0 preconditioned system as compared to k-way or nested dissection ordering, while it reduces data movement and should improve the expected time needed for convergence. In addition to communication avoiding Krylov subspace methods, our preconditioner can be used with classical methods such as GMRES or s-step methods to reduce communication.

In [6] we consider a fundamental problem in combinatorial and scientific computing, the sparse matrix-matrix multiplication problem. Obtaining scalable algorithms for this operations is difficult, since this operation has a poor surface to volume ratio, that is a poor data re-use. We consider that the input matrices are random, corresponding to Erdos-Renyi random graphs. We determine new lower bounds on communication for this case, in which we assume that the algorithm is sparsity independent, where the computation is statically partitioned to processors independent of the sparsity structure of the input matrices. We show in this paper that existing algorithms for sparse matrix-matrix multiplication are sub-optimal in their communication costs, and we obtain new algorithms which are communication optimal, communicating less than the previous algorithms and matching new lower bounds.

6.6. New resuls in communication avoiding algorithms for dense linear algebra

In the context of dense linear algebra algorithms, we have focused on two operations, LU factorization and rank revealing QR factorization.

In [4] we present block LU factorization with panel rank revealing pivoting (block LU_PRRP), a decomposition algorithm based on strong rank revealing QR panel factorization. Block LU_PRRP is more stable than Gaussian elimination with partial pivoting (GEPP), with a theoretical upper bound of the growth factor of $(1+\tau b)^{(n/b)-1}$, where b is the size of the panel used during the block factorization, τ is a parameter of the strong rank revealing QR factorization, n is the number of columns of the matrix, and for simplicity we assume that n is a multiple of b. We also assume throughout all the paper that $2 \leq b \leq n$. For example, if the size of the panel is $b = 64$, and $\tau = 2$, then $(1+2b)^{(n/b)-1} = (1.079)^{n-64} \ll 2^{n-1}$, where 2^{n-1} is the upper bound of the growth factor of GEPP. Our extensive numerical experiments show that the new factorization scheme is as numerically stable as GEPP in practice, but it is more resistant to pathological cases. The block LU_PRRP factorization does only $O(n^2 b)$ additional floating point operations compared to GEPP.

We also present block CALU_PRRP, a version of block LU_PRRP that minimizes communication, and is based on tournament pivoting, with the selection of the pivots at each step of the tournament being performed via strong rank revealing QR factorization. Block CALU_PRRP is more stable than CALU, the communication avoiding version of GEPP, with a theoretical upper bound of the growth factor of $(1+\tau b)^{\frac{n}{b}(H+1)-1}$, where H is the height of the reduction tree used during tournament pivoting. The upper bound of the growth factor of CALU is $2^{n(H+1)-1}$. Block CALU_PRRP is also more stable in practice and is resistant to pathological cases on which GEPP and CALU fail.

We have also introduced CARRQR (paper submitted to SIAM Journal on Matrix Analysis and Applications), a communication avoiding rank revealing QR factorization with tournament pivoting. We show that CARRQR reveals the numerical rank of a matrix in an analogous way to QR factorization with column pivoting (QRCP). Although the upper bound of a quantity involved in the characterization of a rank revealing factorization is worse for CARRQR than for QRCP, our numerical experiments on a set of challenging matrices show that this upper bound is very pessimistic, and CARRQR is an effective tool in revealing the rank in practical problems. Our main motivation for introducing CARRQR is that it minimizes data transfer, modulo polylogarithmic factors, on both sequential and parallel machines, while previous factorizations as QRCP are communication sub-optimal and require asymptotically more communication than CARRQR. Hence CARRQR is expected to have a better performance on current and future computers, where commmunication is a major bottleneck that highly impacts the performance of an algorithm.

6.7. Scalable Schwarz domain decomposition methods

Domain decomposition methods are, alongside multigrid methods, one of the dominant paradigms in contemporary large-scale partial differential equation simulation. A lightweight implementation [8] of a theoretically and numerically scalable preconditioner was developped in the context of overlapping methods. The performance of this work is assessed by numerical simulations executed on thousands of cores, for solving various highly heterogeneous elliptic problems in both 2D and 3D with billions of degrees of freedom. Such problems arise in computational science and engineering, in solid and fluid mechanics.
For example, in 3D, the initial highly heterogeneous problem of 74 million unknowns is solved in 200 seconds on 512 threads. Using 16384 threads, the problem is now made of approximately 2.3 billions unknowns, and it is solved in 215 seconds, which yields an efficiency of $\approx 90\%$. In 2D, the initial problem of 695 million unknowns is solved in 175 seconds on 512 threads. Using 16384 threads, the problem is now made of approximately 22.3 billions unknowns, and it is solved in 187 seconds, which yields an efficiency of $\approx 96\%$.

6.8. Schur domain decomposition methods

We have introduced spectral coarse spaces for the BDD and FETI methods in [5]. These coarse spaces are specifically designed for the two-level methods to be scalable and robust with respect to the coefficients in the equation and the choice of the decomposition. We achieve this by solving generalized eigenvalue problems on the interfaces between subdomains to identify the modes which slow down convergence. Theoretical bounds for the condition numbers of the preconditioned operators which depend only on a chosen threshold and the maximal number of neighbours of a subdomain were proved. For FETI there are two versions of the two-level method: one based on the full Dirichlet preconditioner and the other on the, cheaper, lumped preconditioner. Some numerical tests confirm these results.

6.9. Non conforming domain decomposition methods

We have designed and analyzed a new non-conforming domain decomposition method, named the NICEM method, based on Schwarz-type approaches that allows for the use of Robin interface conditions on non-conforming grids. The method is proven to be well posed. The error analysis is performed in 2D and in 3D for P1 elements. Numerical results in 2D illustrate the new method. This work is in collaboration with C. Japhet and Y. Maday.

6.10. Quadratic finite elements with non-matching grids for the unilateral boundary contact

We analyze in [3] a numerical model for the Signorini unilateral contact, based on the mortar blue method, in the quadratic finite element context. The mortar frame enables one to use non-matching grids and brings facilities in the mesh generation of different components of a complex system. The convergence rates we state here are similar to those already obtained for the Signorini problem when discretized on conforming meshes. The matching for the unilateral contact driven by mortars preserves then the proper accuracy of the quadratic finite elements. This approach has already been used and proved to be reliable for the unilateral contact problems even for large deformations. We provide however some numerical examples to support the theoretical predictions with FreeFem++ (http://www.freefem.org/ff++).

AMIB Project-Team

5. New Results

5.1. RNA

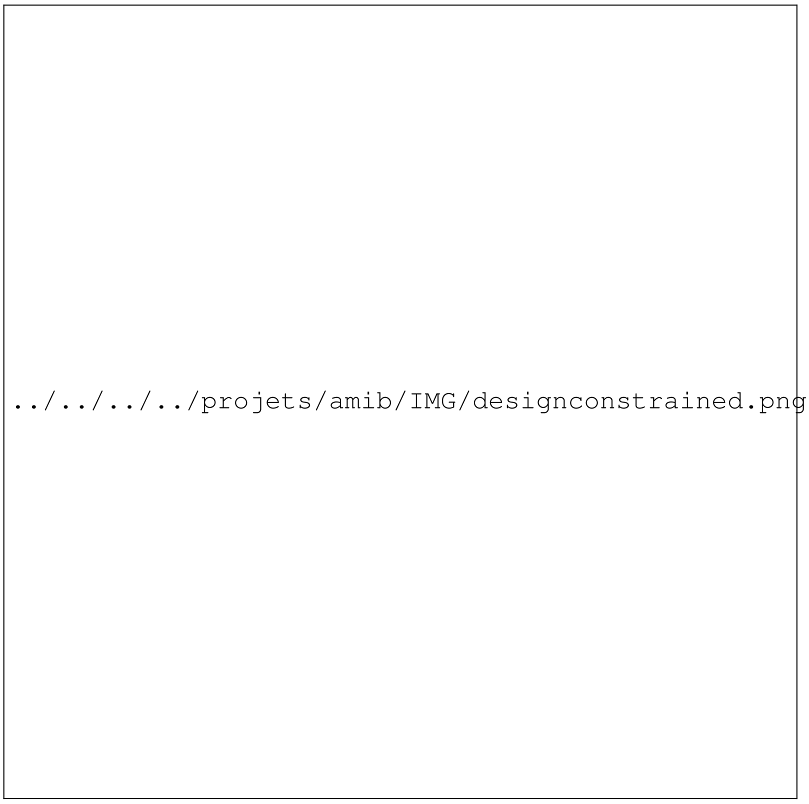


Figure 3. Language-theoretical constructs for the constrained design of RNAs. Starting from a secondary structure, the language of sequences compatible with base-pairing constraints is modeled as a context-free grammar (Left), while forced and forbidden motifs (here, $\{\mathsf{AA}\}$ is forbidden, and $\{\mathsf{AGC}, \mathsf{GG}\}$ are forced) can be modeled by a dedicated automaton (Right).

5.1.1. RNA design through random generation

Extensive experiments revealed a drift of existing software towards sequences with a high G+C-content. Relying on our random generation methods, we showed how to control this distributional bias in sequences using a multidimensional Boltzmann sampling [30], [22]. We also explored the combination of random generation (global sampling) and local search into a novel category of *glocal* approaches, yielding promising results.

Finally, we explored language-theoretic constructs, namely products of finite-state automata and context-free languages, to force or forbid the presence of identified functional motifs within designed sequences [33].

5.1.2. Towards 3D modeling of large molecules

Ab initio research benefited from our works on research and classification of RNA structural motifs [63]. Significant progress towards the *ab initio* prediction of the 3D structure of large RNAs were achieved. This problem is beyond the scope of current approaches and we proposed a promising coarse-grained approach based on game theory [13] that scales up to several hundreds of bases.

5.1.3. Fast-fourier transform for riboswitch

In the field of RNA computational biology, many algorithms use dynamic programming to partition the folding landscape according to a set of structural parameters. More precisely, the goal is to compute the number (resp. cumulated Boltzmann weight) $c_{p_1,p_2,p_3...}$ of secondary structures having p_i occurrences of some structural parameter P_i, where P_i may denote the distance to a reference structure, the number of # helices, base-pairs...The resulting algorithms, although polynomial in theory, are usually unusable in practice, particularly due to their unreasonable complexities (typically $\Theta(n^{3+2k})/\Theta(n^{2+k})$ time/memory for k parameters) and the intrinsic difficulties one encounters while trying to distribute their computation over multiple processors (highly connected dependency graph).

In collaboration with P. Clote's group (Boston College), we have described generic algorithmic principles to dramatically decrease these complexities, and make this class of algorithms practical. The main idea is to capture the partitioned space within a large polynomial, which can typically be efficiently evaluated (typically in $\Theta(n^3)$) as soon as the parameters are additive. One can then perform (possibly in parallel) $\Theta(n^k)$ independent evaluations of the polynomial, and use the Discrete Fourier Transform to recover the coefficients in $\Theta(k \cdot n^k \cdot \log(n))$ time. Applying these principles to the `RNAbor` algorithm, whose complexities were in $\Theta(n^5)/\Theta(n^3)$, we obtained an novel $\Theta(n^4)/\Theta(n^2)$ (parallelizable in $\Theta(n^3)/\Theta(n^2)$ time/memory on $m \to \infty$ processors), we obtained a novel algorithm to detect bistable thermodynamic structures, such as riboswitches, which we presented at `Recomb`'13 [32].

5.2. Sequences

5.2.1. Random generation

The random generation of decomposable combinatorial structures, pioneered by P. Flajolet in the 80s, provides an elegant, yet powerful, framework to model and sample the objects which appear in computational biology. Random samples can then be used to assert the significance of a given observable when closed form formulae are difficult to obtain.

Messenger RNAs (mRNAs) encode proteins, but may also independently feature structured motifs which are crucial to recoding and alternative splicing mechanisms. In order to predict such motifs, the stability of smaller regions within a given mRNA must be compared to that of sequences generated with respect to a **background model** which, at the same time, preserves the encoded amino-acid sequence and the capacity of the overall sequence to form a stable fold (proxy-ed by the dinucleotide composition). Using multidimensional Boltzmann sampling, we have revisited the underlying – well-defined, yet never solved exactly – random generation problem, and provided the first unbiased and practical algorithm for the problem [27]. The algorithm, developed in collaboration with McGill and Université de Montréal (Canada), has linear time complexity as soon as a small tolerance (typically $\Theta(1/\sqrt{n})$) on the composition is allowed.

Some other biological objects, such as RNA secondary structures, naturally appear with probabilities which are poorly modeled by the uniform distribution. To better model such objects, Denise *et al* [3] have introduced the **weighted distribution**, and adapted classic random generation algorithms such that each object within a given combinatorial family can be generated with respect to it. However, the exponentially increasing probability ratio between the most and least probable object sometimes leads to a large degree of redundancy within generated sets . To work around this issue, and generate non-redundant sets of objects, we have proposed a sequential algorithm with deterministically avoids any previously generated word, without introducing any bias in the generation [17].

../../../../projets/amib/IMG/sparcs.png

Figure 4. Workflow of our NASP *pipeline [27]: An assessment of significantly (un)-structured regions in protein-coding RNAs can be achieved through a dinucleotide-preserving random generation of sequences encoding the same protein.*

../../../../projets/amib/IMG/nonredgen.png

Figure 5. A uniform random generation of words avoiding a predefined set of words can be achieved using a dedicated data structure, leading to a careful correction of the emission probabilities. Enriching the set of forbidden words after each generation, one obtains a non-redundant generation algorithm [17].

Besides, in collaboration with the Fortesse group at LRI, we developed a new divide and conquer algorithm for the random generation of words of regular languages, and we performed a complete benchmarking of all state-of-the-art methods dedicates to this problem [56].

../../../../projets/amib/IMG/RNAPyro.png

Figure 6. While simultaneously sequencing the genome of a (microbial) community, Next-Generation Sequencing techniques produce small genomic fragments, whose diversity arises from a combination of genetic variants and sequencing errors. We used knowledge of the RNA secondary structure to develop a pre-filter that detects and corrects post-mapping sequencing errors.

5.2.2. Next Generation Sequencing (NGS)

As a side-product of our previous collaborative studies with J. Waldispühl (McGill, Canada), focusing on sequence/structure relationship in RNA, we revisited the problem of detecting and correcting RNA sequences obtained using pyrrosequencing techniques. Indeed, ribosomal RNAs are often used to estimate the population diversity within a microbiome, and sequencing errors may lead to biased estimates. In this context, we investigated whether a complete knowledge of the RNA secondary structure could be exploited to detect and correct errors in NGS reads.

To that end, we introduced a probabilistic model, defined over all sequences at maximal distance d from the input read and their respective folding. This model captures both the stability of the induced fold and its compatibility with a reference multiple sequence alignment. We designed a linear-time inside/outside algorithm to compute exactly the probability that a given position is mutated in the ensemble. Our initial implementation, presented at RECOMB'13 [29] and published an extended version in *Journal of Computational Biology* [23], revealed encouraging results, and we plan to combine it with a population diversity estimator to test its potential in a metagenomics context.

5.2.3. Combinatorics of motifs

An algorithm for pvalue computation has been proposed in [44] that takes into account a Hiddden Markov Model and an implementation, SUFPREF, has been realized (http://server2.lpm.org.ru/bio).

Combinatorics of clumps have been extensively studied, leading to the definition of the so-called *canonic clumps*. It is shown in [28] that they contain the necessary information needed to calculate, approximate, and study probabilities of occurrences and asymptotics. This motivates the development of a *clump automaton*. It allows for a derivation of pvalues, decreasing the space and time complexity of the generating function approach or previous weighted automata.

Large deviations approximations are needed for very rare events, e.g. very small pvalues, as Gaussian approximations are known not to be applicable. In [21], combinatorial properties of words allow to provide an explicit and tractable formula for the tail distribution with a low space and time complexity and a guaranteed tightness. Double strands counting problem is addressed where dependencies between a sequence and its complement plays a fundamental role. A large deviation result is also provided for a set of small sequences, with non-identical distributions. Possible applications are the search of cis-acting elements in regulatory sequences that may be known, for example from ChIP-chip or ChipSeq experiments, as being under a similar regulatory control. In a recent internship at LIX, F. Pirot detected a Chi-like motif in Archae genome.

In a collaboration with AlFarabi University, where M. Régnier acts as a foreign co-advisor), word statistics were used to identify mRNA targets for miRNAs involved in various cancers [8], [9].

5.3. 3D Modelling and Interactions

Transmembrane beta-barrel proteins (TMB) account for 20 to 30% of identified proteins in a genome but, due to difficulties with standard experimental techniques, they are only 2% of the RCSB Protein Data Bank. Therefore, we study and design algorithmic solutions addressing the secondary structure, an abstraction of the 3D conformation of a molecule, that only retains the contacts between its residues. Although this representation may disregard some of the fine details of the molecule conformation, it still retains the general architecture of molecules, and is especially useful in the study of RiboNucleic Acids (RNAs) and transmembrane beta-barrel proteins (TMB). The latter class of proteins accounts for 20 to 30% of identified proteins in a genome but, due to difficulties with standard experimental techniques, they constitute only 2As TMB perform many vital functions, the prediction of their structure is a challenge for life sciences, while the small number of known structures prohibits knowledge-based methods for structure prediction. As TMBs are strongly structured objects, model based methodologies [26], [25] are an interesting alternative to these conventional methods. The efficiently obtained 3D structures provide a good model for further 3D and interaction analyses.

In a recent work [34], we focused on the identification of protein-protein complexes based on the putative interaction between pairs of proteins as the sole source of information. From the results obtained on *E. coli*, we started working on the prediction of multi-body protein complexes from sequence information alone.

In our protein-RNA project, we managed to obtain the first learning results. We optimized the RosettaDock scores and showed that such an optimization cannot be done efficiently without expert knowledge. The first results are to be presented at EGC in 2014 [61].

5.3.1. *Large scale cross-docking study of the specificity of protein-protein interactions*

The year 2013 saw the conclusion of a long-term collaboration, involving A. Carbone (UPMC) and A. Lopes (IGM, Paris XI). In a recent paper published in the prestigious *Plos Computational Biology* [16] journal, we showed that combining coarse-grain molecular cross-docking simulations and binding site predictions based on evolutionary sequence analysis is a viable route to identify true interacting partners for hundreds of proteins with a variate set of protein structures and interfaces. Also, we realized a large-scale analysis of protein binding promiscuity and provided a numerical characterization of partner competition and level of interaction strength for about 28000 false-partner interactions. Finally, we demonstrated that binding site prediction is useful to discriminate native partners, but also to scale up the approach to thousands of protein interactions. This study was based on a large computational effort made by thousands of internet users helping the World Community Grid over a period of 7 months.

5.4. Data Integration

Work performed in the Data Integration axis this year has been dedicated to the design and implementation of a new approach to reduce the complexity of scientific workflow structures. More precisely, we focused on the presence of “anti-patterns” in the workflow structures, idiomatic structures that lead to over-complicated design. We have then proposed the *DistilFlow* method and a tool for automatically detecting such anti-patterns and replacing them with different patterns which result in a reduction in the workflow’s overall structural complexity [10] (BMC Journal paper accepted, published early 2014). This work has been performed in close collaboration with the Taverna group from the University of Manchester.

DistilFlow is part of J. Chen’s thesis who has defended his PhD on October 11th, 2013 [7] and is now back to China as a research assistant in Lanzhou University.

5.5. Systems Biology

Systems Biology includes the study of interaction networks such as gene regulatory, metabolic, or signaling networks. It involves both designing the topology of the networks and predicting their dynamic and spatiotemporal aspects. It requires the import of concepts from across various disciplines and crosstalk between theory, benchwork, modelling and simulation.

5.5.1. *Topological analysis of metabolic networks*

In [73] we have developed a biclustering algorithm for elementary flux modes that is based on the Agglomeration of Common Motifs (ACoM). This allows a drastic diminution of the number of less significant fluxes and a kind of factorization of most important fluxes, yielding an algorithm running in quadratic time in the number of elementary flux modes.

We applied this algorithm to describe the decomposition into elementary flux modes of the central carbon metabolism in *Bacillus subtilis* and of the yeast mitochondrial energy metabolism. For *Bacillus subtilis*, a specific inhibition on the second domain of the lipoamide dehydrogenase (pdhD) component of pyruvate dehydrogenase complex that leads to the loss of all fluxes was exhibited [20]. Such a conclusion is not predictable in the classical approach.

5.5.2. *Evolution of metabolic networks*

A collaboration with IGM on the evolution of metabolic networks is ungoing. We aim at understanding how such networks would emerge over time among the variety of species, and how these changes could be responsible for characteristic life traits. Our methodology to characterize the evolutionary origin of the enzymatic repertoire of different fungal groups relies on machine learning. Preliminary results were presented at JOBIM 2013 [35].

5.5.3. Signaling networks

Our goal is to help the understanding of signaling pathways involving (GPCR) and to provide means to semi-automatically construct the signaling networks. Our method takes into account various kinds of biological experiments and their origin and automatically builds and draws the inferred network. Comparing the automatically deduced network with an already known fragment of the FSHR network allowed us to obtain new interesting hypotheses that are currently experimentally tested by biologists, our collaborators from INRA-BIOSin Tours. In the next months, experimental data for some GPCR (FSH, 5HT2 et 5HT4) will be prepared by BIOS and IGF (Montpellier), in the context of a GPCRNET ANR project.

Besides, in collaboration with K. Inoue, through the NII International Internship Program, we have studied the System Biology Graphical Notation language, a standard for expressing molecular networks, especially signaling networks, and proposed a translation of SBGN-AF into a logical formalism [31].

5.5.3.1. Modelling with Hsim

In a collaboration of P. Amar with microbiologists, the group of Marie-Joëlle Virolle from the *Institut de Génétique et de Microbiologie*, a first explicative model was proposed for the sigmoidicity of the shape of the survival curve of bacteria (*S. lividans*) having a antibiotic resistance gene, expressed at different levels, in presence of a constant concentration of antibiotics [24], [6], [18], [41].

This is particularly important since this method of inclusion of an antibiotics resistance gene to report the activity of its promoter is widely used in the streptomyces community.

5.5.3.2. Cancer and metabolism

It is shown in M. Behzadi's PhD thesis that most systems have very stable behaviours and that even large variations of their chemical characteristics do not affect the nature of the equilibria. This very general situation has been discovered by simulation but in some cases it is even possible to prove it mathematically.

Our collaborators M. Israël and L. Schwartz have listed more than a hundred tentative such bifurcations that we intend to study systematically. A preliminary study of the mitotic cycle with L. Paulevé has also put in evidence the strong influence of the pH of the cell on its capacity to duplicate. The PhD thesis of E. Bigan, co-directed by S. Daoudi (Univ. Denis Diderot) and J.-M. Steyaert investigates the generic properties of such complex systems and confirms that the ones we have already studied are not exceptions [43]. Some prospective cases are studied in [14].

ANGE Team

6. New Results

6.1. Geophysical flows

6.1.1. A numerical scheme for the Saint-Venant equations

Participants: Emmanuel Audusse, Christophe Chalons [Univ. Versailles], Philippe Ung.

In order to improve the numerical simulations of the shallow-water equations, one has to face three important issues related to the well-balanced, positivity and entropy-preserving properties, as well as the ability to handle vacuum states. In that purpose, we propose a Godunov-type method based on the design of a three-wave Approximate Riemann Solver (ARS) which satisfies all aformentioned properties.

6.1.2. Two-phase flows

Participants: Frédéric Coquel [CNRS], Jean-Marc Hérard [EDF], Khaled Saleh [IRSN], Nicolas Seguin.

After having developed numerical schemes for models of compressible two-phase flows [17], [19], we have proven some fundamental properties of these systems: symmetrizability and (non strict) convexity of the entropy [18]. This enables us now to address the well-posedness of these models when the relaxation terms are included.

6.1.3. Non-hydrostatic models

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

The objective is to derive a model corresponding to a depth averaged version of the incompressible Euler equations with free surface. We have already contributed to this subject but the obtained results extend previous ones [29] in several directions:

- the derivation of the model is more rigorous and follows the entropy-based moment closures proposed in [28],
- the properties of the model and especially its connections with Green-Nagdhi model have been investigated,
- a family of analytical solutions for the proposed model have been obtained.

These analytical solutions emphasize the non-hydrostatic effects appearing for large slope variations.

6.1.4. Fluids with complex rheology

Participants: Anne Mangeney, Jacques Sainte-Marie.

We have been able

- to develop detection, characterization and localisation methods applicable to the seismic signals generated by rockfalls and thus to analyse ths spatio-temporal change of rockfall localisation and properties during several years, making it possible to show how rockfalls can be used as a precursor of volcanic activity,
- to propose an empirical “universal” law describing friction weakening in landslides over a broad range of volumes and geological contexts,
- to propose a new debris flow model with an energy balance,
- show the existence of a slow propagation phase in granular flows, playing a key role in their dynamics and in erosion processes.

6.1.5. Dynamics of sedimentary river beds with stochastic fluctuations

Participants: Emmanuel Audusse, Philippe Ung.

The Exner equation is a coarse model for the dynamics of sedimentary river beds, derived using both many heuristics and empirism. Though, it is also quite practical for hydraulic engineering applications, and efficient enough in numerous situations. Our goal in this work is to improve the model by including some effects that have been neglected so far in the heuristics. In particular, inline with other current research directions in the field, we study the possibility of introducing some stochasticity in the model. To this end, we suggest to numerically experiment some recently proposed variations of the Exner equation based on the introduction of stochastic fluctuations within the standard formulation.

This project has been the subject of a study during the 2013 session of the CEMRACS.

6.2. Ecology and sustainable energies

6.2.1. Hydrodynamic-biology coupling

Participants: Olivier Bernard [Inria BIOCORE], Anne-Céline Boulanger, Marie-Odile Bristeau, Raouf Hamouda, Jacques Sainte-Marie.

An important part of our research activity is built around a biological and industrial problem: the simulation of the coupling of hydrodynamics and biology in the context of industrial microalgae culture in outdoor raceways. The numerical modelling is adressed with the use of a multilayer vertical discretization of hydrostatic Navier-Stokes equations coupled with a light sensitive Droop model. Numerically, kinetic schemes allow for the development of efficient, positivity preserving, well balanced and entropy satisfying schemes. Simulations are carried out in 2D and 3D [1]. From a practical point of view, this model is capable of accounting for the utility of a paddlewheel and exhibits Lagrangian trajectories underwent by algae. Hence providing hints on the light history of algae in the pond, which is a key information to biologists, since it enables them to adapt their phytoplankton growth models to those particular, non natural conditions.

6.3. Coupling methods

6.3.1. Data assimilation for conservation laws associated with kinetic description

Participants: Anne-Céline Boulanger, Philippe Moireau [Inria M3DISIM], Jacques Sainte-Marie.

In order to take advantage of the kinetic description of conservation laws already used for the building of efficient schemes, an innovative data assimilation method for hyperbolic balance laws based in a Luenberger observer on the kinetic equation is developed. It provides a nice theoretical framework for scalar conservation laws, for which we study the cases of complete observations, partial observations in space, in time, and noisy observations. As far as systems are concerned, we focus on the Saint-Venant system, which is hyperbolic, nonlinear and has a topographic source term. We build an observer based only on water depths measurements. Numerical simulations are provided in the case of scalar laws and systems, in one and two dimensions, which validate the efficiency of the method [14].

6.3.2. Mach-parametrized flows

Participants: Stéphane Dellacherie [CEA], Bruno Després [UPMC Paris 6], Yohan Penel.

In order to enrich the modelling of fluid flows, we investigate in this paper a coupling between two models dedicated to distinct regimes. More precisely, we focus on the influence of the Mach number as the low Mach case is known to induce theoretical and numerical issues in a compressible framework. A moving interface is introduced to separate a compressible model (Euler with source term) and its low Mach counterpart through relevant transmission conditions. A global steady state for the coupled problem is exhibited. Numerical simulations are then performed to highlight the influence of the coupling by means of a robust numerical strategy [20].

6.3.3. Error analysis in a coupling strategy

Participants: Clément Cancès [UPMC Paris 6], Frédéric Coquel [CNRS], Edwige Godlewski, Hélène Mathis [Univ. Nantes], Nicolas Seguin.

We have proposed in a simplified framework an error analysis for an adaptive method which automatically selects the optimal model to use, the choice being between a reference model and an associated simplified one, see [15]. In particular, we are able to balance the thickness of the coupling buffer zone with the threshold on the modelling error which appears when introducing the coarse model.

6.4. Software development and assessments

6.4.1. Analytical solutions for the incompressible Euler system

Participants: Anne-Céline Boulanger, Marie-Odile Bristeau, Jacques Sainte-Marie.

We have proposed in [5] a large set of analytical solutions (FRESH-ASSESS) for the hydrostatic incompressible Euler system in 2d and 3d. These solutions mainly concern free surface flows but partially free surface flows are also considered. These analytical solutions can be especially useful for the validation of numerical schemes.

6.4.2. Software

Several tasks have been achieved in the FRESHKISS3D software (§ 5.1):

- First tests with a uniform pression before moving to the variable case;
- Rethinking of the C++ code with an object-oriented rewriting which provides a better memory management;
- Automatic boundary conditions handling in the case of a fluid/solid transition;
- New computations of the particule trajectories when they leave out the domain by means of directional interpolation procedures;
- Achievement of the 2nd-order space accuracy;
- Taking into account the wind.

AOSTE Project-Team

6. New Results

6.1. Process Networks with routing for parallel architectures

Participants: Robert de Simone, Emilien Kofman, Jean-Vivien Millo.

In the past we developed a dedicated Process Network (PN) formalism with explicit static switching/routing schemes for data flow. This year we considered the practical use of our formalism to model data-streams in specific applicative contexts.
In a first direction we considered the case of stencil algorithms, usually modeled with cellular automata (CA) (as in heat or gas propagation models for instance). In that case, the application itself is modeled in a way strongly similar to a physical architecture consisting of a regular mesh/array of parallel processors (MPPA). Mapping can seem to be straighforward then, *safe that* the neighborhood and connection topology may differ from the CA model to the MPPA. Our results consider efficient routing and propagation schemes on a given MPPA interconnect fabric, so as to match all-to-all broadcast paterns up to a given distance (on the CA topology). They are described in [20], and were implemented on Kalray MPPA256 prototype architecture.
A similar modeling effort was conducted, this time on FFT algorithm models (again described as parallel pipe-lined tasks). Again switching/routing schemes were provided in our formal PN model to map the virtual logical dependences onto concrete connection patterns in a MPPA256 model. This was the subject of Emilien Kofman internship, of which preliminary results were presented in a junior workshop [36].

6.2. Formal analysis of MARTE Time Model and CCSL

Participants: Frédéric Mallet, Robert de Simone, Yuliia Romenska, Jean-Vivien Millo, Ling Yin.

We have worked on building analysis methods and tools for running exhaustive analyses on MARTE/CCSL specifications. This was done by endowing CCSL with a State-Based semantics [51]. Each operator is described as a boolean state machine, some operators require an infinite number of states. When this is the case we rely on a lazy representation technique to capture symbolically the infinite number of states [45]. The semantics of a CCSL specification is then expressed as the synchronized product of the (infinite) state machines for each operator. Even though the operators are infinite, their composition can sometimes be bounded. When the synchronized product has only a finite number of reachable states, it is said to be safe. We have identified a set of representative and frequently used examples where this is the case [38]. When the product is not finite, our (semi-)algorithm to build the product does not terminate, therefore it is important to be able to know in advance whether or not the product is safe. We have thus proposed an algorithm to decide whether a CCSL specification is safe [37]. It relies on an intermediate representation called Clock Causality Graph and uses results from marked graph theory.

Building the product for a CCSL specification is exponential in the number of clocks and is not practical for large specifications. So, to avoid building explicitly the product we have proposed another technique to explore symbolically the state-space of a CCSL specification [49]. This relies on a liveness condition where no conflict may prevent an infinite clock from ticking infinitely often. Branches that may lead to states where an infinite clock dies are pruned by a fix-point algorithm.

These two solutions focus on the logical and discrete aspects of MARTE/CCSL, which was devised to unify logical and physical time constraints. An attempt to support verification of the physical time constraints of MARTE/CCSL was conducted through the use of UppAal timed automata and model-checker [46]. The proposed technique combines the logical clocks of CCSL with the real-valued clocks of timed automata. Synchronous/Polychronous aspects are solved with TimeSquare 5.1 while the UppAal model-checker is used to explore the space derived from the real-valued clocks.

6.3. Logical time in Model-Driven Engineering of embedded systems

Participants: Frédéric Mallet, Julien Deantoni, Robert de Simone, Marie-Agnès Peraldi Frati, Matias Vara-Larsen, Arda Goknil.

In the context of our approach based on logical time to specify causalities and synchronizations on models, 3.2 , we developed an extension of the OMG OCL Object Constraint Language. Named ECL (Event Constraint Language) it provides such specifications of causalitity and synchronization at syntactic language level, which enabled then automatic generation of semantic logical time constraints for any model that conforms the language.
This year, we extended to a new challenge, using logical time constraints to coordinate models of *several distinct* languages used jointly for a large heterogeneous system description. This work is reported in [25], [52].
It was illustrated in practice in the automotive domain by coordinating together the Timed Augmented Description Language (TADL2) and the EAST-ADL language [34], [32] (the formalisms are rather similar, but still with clear distinctions at places) .
Finally, we proposed a pattern to assemble the (possibly concurrent) semantics of a language associating our logical time constraints (based on pure clocks) with a syntactic action language (providing behavior content). By reifying events and constraints, this specification of the semantics is amenable to its composition [25]. Such approach has been, again, recently used for a first attempt to coordinate distinct behavioral models [47].

As part of our collaboration in the DAESD associated-team with ECNU Shone-SEI in Shanghai we studied the coupling of discrete-logical with continuous-physical time models, ending with a proposal of Hybrid MARTE statecharts [19] specified in a style much like a combinaison of MARTE state diagrams and timed automata.

In another setting we presented a new model of scenarios [21], dedicated to the specification and verification of system behaviours in the context of software product lines (SPL). The formalism uses the logical time modeling aproach, with a strong link to synchronous semantics. We draw our inspiration from some techniques that are mostly used in the hardware community, and we show how they could be applied to the verification of software components and product line variability. We point out the benefits of synchronous languages and models to bridge the gap between both worlds.

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

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

Power models for embedded architectures (where power consumption is highly constrained) provide an ideal example of a non-functional modeling framework with strong interactions with the functional and performance models: more speed in computation comes at the cost of larger energy consumption. There was also a demand for a framework allowing combinaison of models, each representing a distinct view of the system. We demonstrated as part of the HeLP ANR project 8.2.1.1 , followed by the newly started HOPE ANR project 8.2.1.2 , how such multiview modeling could be done, and how it could be connected down to more concrete simulation code or model, as in SystemC, Docea Power AcePlorer, or Scilab code. The multiview modeling applied to power intent and power managers was described in [35], and led to the PhD defense of Carlos Gomez Cardenas in December 2013 [16].

6.5. Performance variability analysis on manycore architectures

Participants: Sid Touati, Amin Oueslati, Franco Pestarini, Robert de Simone, Emilien Kofman.

In the context of the collaboration with Kalray (see 7.1.1), we conducted a systematic benchmarking campaign to test the stability (or low variability) of the performances of the MPPA256 prototype manycore processor. We first addressed issues of memory access and network latency, then programmed a distributed verson of the classical ALL_PAIRS_SHORTEST_PATH parallel algorithm with an hybrid OpenMP/MPI style. This was the objectif of Amin Oueslati Master2 internship. Results were encouraging, and showed stability of performance over a large set of runs.

This work is currently extended during the International Internship grant of Franco Pescarini. Specific on-chip communication modes offered by the MPPA256 processor (namely *portal* and *channel* communication modes) are being extensively benchmarked. Results show time predictability on the case of light on-chip communication traffic, but stability gets degraded as performance decreases in presence of heavy traffic and congestion (various runs show quite different execution time).
In another effort we conducted during the internship period of Emilien Kofman an experiment on MPPA256 quite similar to the work conducted as part of the collaboration with Kontron (see 7.1.3), exploring various mapping options of FFT algorithm variants, with the goal of figuring how to best map (in the future) several such algorithms onto the computation fabric of the many-cores available.

6.6. Off-line (static) mapping of real-time applications onto NoC-based many-cores

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

Modern computer architectures are increasingly relying on multi-processor systems-on-chips (MPSoCs, also called chip-multiprocessors), with data transfers between cores and RAM banks managed by on-chip networks (NoCs). This reflects in part a convergence between embedded, general-purpose PC, and high-performance computing (HPC) architecture designs. In past years we have identified and compared the hardware mechanisms supporting precise timing analysis and efficient resource allocation in existing NoCs. We determined that the NoC should ideally provide the means of enforcing a global communications schedule that is computed off-line and which is synchronized with the scheduling of computations on CPU cores (and we have built such a NoC).

This year we have focused on the problem of mapping applications onto NoC-based MPSoCs (discussed in this section) and on the associated problem of timing analysis of the resulting parallel implementations (discussed in section 6.7). On-chip networks used in MPSoCs pose significant challenges to both on-line and off-line real-time scheduling approaches. They have large numbers of potential contention points, have limited internal buffering capabilities, and network control operates at the scale of small data packets. Therefore, precise schedulability analysis requires scalable algorithms working on hardware models with a level of detail that is unprecedented in real-time scheduling.

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

Our method has been implemented within the Lopht tool presented in section 5.4 , and first results are presented in [54]. One of the objectives of the collaboration with Kalray SA is the evaluation of the possibility of porting Lopht onto the Kalray MPPA platform.

6.7. WCET estimation for parallel code

Participant: Dumitru Potop Butucaru.

This is joint work with Isabelle Puaut, Inria, EPI ALF.

Classical timing analysis techniques for parallel code isolate micro-architecture analysis from the analysis of synchronizations between cores by performing them in two separate analysis phases (WCET – worst-case execution time – and WCRT – worst-case response time analyses). This isolation has its advantages, such as a reduction of the complexity of each analysis phase, and a separation of concerns that facilitates the development of analysis tools. But isolation also has a major drawback: a loss in precision which can be significant. To consider only one aspect, to be safe the WCET analysis of each synchronization-free sequential code region has to consider an undetermined micro-architecture state. This may result in overestimated WCETs, and consequently on pessimistic execution time bounds for the whole parallel application.

The contribution of this work [56], [44] is an *integrated* WCET analysis approach that considers at the same time micro-architectural information and the synchronizations between cores. This is achieved by extending a state-of-the-art WCET estimation technique and tool to manage synchronizations and communications between the sequential threads running on the different cores. The benefits of the proposed method are twofold. On the one hand, the micro-architectural state is not lost between synchronization-free code regions running on the same core, which results in tighter execution time estimates. On the other hand, only one tool is required for the temporal validation of the parallel application, which reduces the complexity of the timing validation toolchain.

Such a holistic approach is made possible by the use of deterministic and composable software and hardware architectures (many-cores with no cache sharing and time-predictable interconnect, static assignment of the code and data to the memory banks). Such code can be written by hand or automatically synthesized using the Lopht tool 5.4 or other automatic parallelization techniques.

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

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

We have continued this year the work on real-time scheduling and code generation for time-triggered platforms. This work was mainly carried out as part of a bilateral collaboration with Astrium Space Transportation (now part of Airbus Defence and Space), which co-funded with the CNES the post-doctorate of Raul Gorcitz (started in September).

The work focused this year on the improvement of the real-time scheduling and code generation (the PhD work of T. Carle), and on determining their adequacy to Astrium's industrial needs (the post-doc of Raul Gorcitz). We have improved our specification, mapping, and code generation technique at all levels. We have extended the Lopht tool to allow automatic mapping and code generation for single-processor and multi-processor partitioned targets (using an ARINC 653-compliant OS).

6.9. Uniprocessor Real-Time Scheduling

Participants: Yves Sorel, Falou Ndoye, Daniel de Rauglaudre.

6.9.1. Formal Proofs of Uniprocessor Real-Time Scheduling Theorems

We continued writing a monograph about three formal proofs, done in 2011/2012, in Coq on scheduling of fixed priority real-time preemptive tasks: one about the scheduling conditions of strict periodicity and two about the worst response time in the case of preemptive deadline monotonic scheduling. This document contains about 120 pages for the moment.

6.9.2. Real-Time Scheduling with Exact Preemption Cost

We proposed a new schedulability condition for dependent tasks executed on a uniprocessor which takes into account the exact preemption cost. Unlike the work presented in [10] which achieves that goal only for fixed priority tasks, our schedulability condition considers fixed as well as dynamic priorities tasks. Thus, we can overcome priority inversions involved by data dependent tasks. The schedulability analysis based on this schedulability condition led to an off-line scheduler [42] described by a scheduling table. Therefore, we have proposed an on-line time-trigger scheduler which implements this scheduling table. Compared to classical on-line schedulers, the proposed approach has two benefits. On the one hand the cost of the task selection amounts only to read the task to be executed in the scheduling table built off-line, rather than using on-line a scheduling algorithm like RM, DM, EDF, etc. On the other hand this cost is fixed since it does not depend on the number of ready tasks. In addition, with our on-line scheduler we do not need to synchronize, on-line, the utilization of the shared memory data, due to dependences, because this synchronization is performed during the off-line schedulability analysis.

6.10. Multiprocessor Real-Time Scheduling

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

6.10.1. Multiprocessor Partitioned Scheduling with Exact Preemption Cost

We finalized the work started in previous years on multiprocessor scheduling of preemptive independent real-time tasks with exact preemption cost [43].

This year we proposed a heuristic for the multiprocessor scheduling of preemptive dependent real-time tasks with exact preemption cost. We chose the partitioned approach that avoids migration of tasks and allows the utilization of the uniprocessor schedulability condition, previously proposed, that takes into account the exact preemption cost. In addition, this schedulability condition takes into account the inter-processor communications and guarantees that no data is lost. The result of such an off-line scheduling provided by the heuristic, is a scheduling table for every processor which includes also inter-processor communication tasks. We compared our multiprocessor scheduling heuristic with a Branch & Bound exact algorithm using the same schedulability condition. Our heuristic provides similar results and is very much faster.

6.10.2. Multiprocessor Semi-Partitioned Mixed Criticality Scheduling

We mainly focused on the mixed criticality scheduling problem applied to semi-partitioned scheduling considering a static pattern of migration for jobs. We have studied this problem in the context of Mixed Criticality (MC) scheduling, a promising approach that can be used to take into account applications of different criticality levels on the same platform. The goal of MC approach is to better utilize computing resources by allowing low criticality tasks to execute in conjunction with high criticality tasks when the system criticality is not high.

6.10.3. Gateway with Modeling Languages for Certified Code Generation

This work was carried out in the P FUI project 8.2.2 . We defined a SynDEx UML profile for functional specifications. We developed a gateway between the P pivot formalism and SynDEx. This gateway deals with the data-flow modeling part of the P formalism which is compliant with the Simulink subset blocks supported by the P project, except for the IF, FOR, MERGE and MUX blocks. Presently, we enhance the gatetway to include these blocks and we colloborate with the other partners to define the architectural part of the P formalism. This part is intended to replace the non functional specifications, presently described with the UML profile MARTE (Modeling and Analysis of Real-Time Embedded Systems).

6.10.4. SynDEx updates with new results

We released an alpha version of SynDEx V8. This version is based on a new textual language whose compiler may be launched with commandes-lines featuring various options. In Syndex V8, the adequation heuristic which performs the multiprocessor real-time schedulability analysis on multi-periodic applications, is based on the theorems and algorithms provided in the Mohamed Marouf's thesis defended last year in the team. These algorithms have been deeply improved for better consideration of data dependencies in the case of multiprocessor architectures. On the other hand, the new heuristic generates a scheduling table composed of, in addition to the usual permanent phase, a transient phase that takes into account the distribution constraints defined by the user in the multi-periodic applications as well as in the mono-periodic applications.

6.11. Probabilistic Real-Time Systems

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

The adventof complex hardware, in response to the increasing demand for computing power in next generation systems, exacerbates some of the limitations of static timing analysis for the estimation of the worst-case execution time (WCET) estimation. In particular, the effort of acquiring (1) detail information on the hardware to develop an accurate model of its execution latency as well as (2) knowledge of the timing behaviour of the program in the presence of varying hardware conditions, such as those dependent on the history of previously executed instructions. These problems are also known as the timing analysis walls. The probabilistic timing analysis, a novel approach to the analysis of the timing behaviour of next-generation real-time embedded systems, provides answers to timing analysis walls. In [17], [48], [31] timing analysis attacks the timing analysis walls. We have also presented experimental evidence that shows how probabilistic timing analysis reduces the extent of knowledge about the execution platform required to produce probabilistically-safe and tight WCET estimations.

Based on existing estimations of WCET or minimal inter-arrival time, one may propose different probabilistic schedulability analyses [39]. These results were reported in the (PhD thesis of Dorin Maxim, mostly conducted in the Inria TRIO team (before its completion and the move to Aoste in Sept 2013).

2013 was also the year when through several invited talks [26], [28], [27], we had the opportunity to underline historical misunderstandings on probabilistic real-time systems. The most common is related to the notion of independence that is used with a wrong meaning by different papers.

APICS Project-Team

6. New Results

6.1. Source recovery problems

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

This section is concerned with inverse problems for 3-D Poisson-Laplace equations. Though the geometrical settings differ in the 2 sections below, the characterization of silent sources (that give rise to a vanishing potential at measurement points) is one of the common problems to both which has been recently achieved in the magnetization setup, see [14].

6.1.1. Application to EEG

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

In 3-D, functional or clinical active regions in the cortex are often modeled by point-wise sources that have to be localized from measurements on the scalp of a potential satisfying a Laplace equation (EEG, electroencephalography). In the work [3] it was shown how to proceed via best rational approximation on a sequence of 2-D disks cut along the inner sphere, for the case where there are at most 2 sources. Last year, a milestone was reached in the research on the behavior of poles in best rational approximants of fixed degree to functions with branch points [6], to the effect that the technique carries over to finitely many sources (see Section 4.2).

In this connection, a dedicated software "FindSources3D" is being developed, in collaboration with the team Athena and the CMA. We took on this year algorithmic developments, prompted by recent and promising contacts with the firm BESA (see Section 5.6), namely automatic detection of the number of sources (which is left to the user at the moment) and simultaneous processing of data from several time instants. It appears that in the rational approximation step, *multiple* poles possess a nice behavior with respect to branched singularities. This is due to the very physical assumptions on the model (for EEG data, one should consider *triple* poles). Though numerically observed in [8], there is no mathematical justification so far why multiple poles generate such strong accumulation of the poles of the approximants. This intriguing property, however, is definitely helping source recovery. It is used in order to automatically estimate the "most plausible" number of sources (numerically: up to 2, at the moment).

In connection with the work [14] related to inverse magnetization issues (see Section 6.1.2), the characterization of silent sources for EEG has been carried out [42]. These are sums of (distributional) derivatives of Sobolev functions vanishing on the boundary.

In a near future, magnetic data from MEG (magneto-encephalography) will become available along with EEG data; indeed, it is now possible to use simultaneously corresponding measurement devices, in order to measure both electrical and magnetic fields. This should enhance the accuracy of our source recovery algorithms.

Let us mention that discretization issues in geophysics can also be approached by such techniques. Namely, in geodesy or for GPS computations, one is led to seek a discrete approximation of the gravitational potential on the Earth's surface, from partial data collected there. This is the topic of a beginning collaboration with physicist colleagues (IGN, LAREG, geodesy). Related geometrical issues (finding out the geoid, level surface of the gravitational potential) are worthy of consideration as well.

6.1.2. Magnetization issues

This work is carried out in the framework of the "équipe associée Inria" IMPINGE, comprising Eduardo Andrade Lima and Benjamin Weiss from the Earth Sciences department at MIT (Boston, USA) and Douglas Hardin and Edward Saff from the Mathematics department at Vanderbilt University (Nashville, USA), Localizing magnetic sources from measurements of the magnetic field away from the support of the magnetization is the fundamental issue under investigation by IMPINGE The goal is to determine magnetic properties of rock samples (*e.g.* meteorites or stalactites) from fine field measurements close to the sample that can nowadays be obtained using SQUIDs (supraconducting coil devices). Currently, rock samples are cut into thin slabs and the magnetization distribution is considered to lie in a plane, which makes for a somewhat less indeterminate framework than EEG as regards inverse problems because "less" magnetizations can produce the same field (for the slab has no inner volume).

The magnetization operator is the Riesz potential of the divergence of the magnetization, see (1). Last year, the problem of recovering a thin plate magnetization distribution from measurements of the field in a plane above the sample led us to an analysis of the kernel of this operator, which we characterized in various functional and distributional spaces [14]. Using a generalization of the Hodge decomposition, we were able to describe all magnetizations equivalent to a given one. Here, equivalent means that the magnetizations generate the same field from above and from below if, say, the slab is horizontal. When magnetizations have bounded support, which is the case for rock samples, we proved that magnetizations equivalent from above are also equivalent from below, but this is no longer true for unbounded supports. In fact, even for unidirectional magnetizations, uniqueness of a magnetization generating a given field depends on the boundedness of the support, as we proved that *any* magnetization is equivalent from above to a unidirectional one (with infinite support in general). This helps explaining why methods in the Fourier domain (which essentially loose track of the support information) do encounter problems. It also shows that information on the support must be used in a crucial way to solve the problem.

This year, we produced a fast inversion scheme for magnetic field maps of unidirectional planar geological magnetization with discrete support located on a regular grid, based on discrete Fourier transform [18]. Figures 5 , 6 , 7 and 8 show an example of reconstruction. As the just mentioned article shows, the Fourier approach is computationally attractive but undergoes aliasing phenomena that tend to offset its efficiency. In particular, estimating the total moment of the magnetization sample seems to require data extrapolation techniques which are to take place in the space domain. This is why we have started to study regularization schemes based on truncation of the support in connection with singular values analysis of the discretized problem.

../../../../projets/apics/IMG/inriaVisuel.png

Figure 5. Inria's logo were printed on a piece of paper. The ink of the letters "In" were magnetized along a direction D_1. The ink of the letters "ria" were magnetized along another direction D_2 (almost orthogonal to D_1).

In a joint effort by all members of IMPINGE, we set up a heuristics to recover dipolar magnetizations, using a discrete least square criterion. At the moment, it is solved by a singular value decomposition procedure of the magnetization-to-field operator, along with a regularization technique based on truncation of the support. Preliminary experiments on synthetic data give quite accurate results to recover the net moment of a sample, see the preliminary document http://www-sop.inria.fr/apics/IMPINGE/Documents/NotesSyntheticExample.pdf. We also ran the procedure on real data (measurements of the field generated by Lunar spherules) for which the net moment can be estimated by other methods. The net moment thus recovered matches well the expected moment.

This shows that the technique we use to reduce the support, which is based on thresholding contributions of dipoles to the observations, is capable of eliminating some nearly silent dipole distributions which flaw the

../../../../projets/apics/IMG/inriaMesureBZ.png

Figure 6. The Z-component of the magnetic field generated by the sample is measured by a SQUID microscope. The measure is performed 200µm above the sample.

../../../../projets/apics/IMG/inriaRecover1.png

Figure 7. The field measured in Figure 6 is inversed, assuming that the sample is uni-dimensionally magnetized along the direction D_1. The letters “In” are fairly well recovered while the rest of the letters is blurred (because the hypothesis about the direction of magnetization is false for “ria”).

../../../../projets/apics/IMG/inriaRecover2.png

Figure 8. The field measured in Figure 6 is inversed, assuming that the sample is uni-dimensionally magnetized along the direction D_2. The letters “ria” are fairly well recovered while the rest of the letters is blurred (because the hypothesis about the direction of magnetization is false for “In”).

singular value analysis. In order to better understand the geometric nature of such distributions, and thus affirm theoretical bases to the above mentioned heuristics, we raised the question of determining an eigenbasis for the positive self adjoint operator mapping a L^2 magnetization on a rectangle to the field it generates on a rectangle parallel to the initial one. Once ordered according to decreasing eigenvalues, such a basis should retain "as much information as possible" granted the order of truncation.

This is not such an easy problem and currently, in the framework of the PhD thesis of D. Ponomarev, we investigate a simplified two-dimensional analog, defined via convolution of a function on a segment with the Poisson kernel of the upper half-plane and then restriction to a parallel segment in that half-plane. Surprisingly perhaps, this issue was apparently not considered in spite of its natural character and the fact that it makes contact with classical spectral theory. Specifically, it amounts to spectral representation of certain compressed Toeplitz operators with exponential-of-modulus symbols. Beyond the bibliographical research needed to understand the status of this question, only preliminary results have been attained so far.

6.2. Boundary value problems

Participants: Laurent Baratchart, Slah Chaabi, Sylvain Chevillard, Juliette Leblond, Dmitry Ponomarev, Elodie Pozzi.

This work was the occasion of collaborations with Alexander Borichev (Aix-Marseille University), Jonathan Partington (Univ. Leeds, UK), and Emmanuel Russ (Univ. Grenoble, IJF).

6.2.1. Generalized Hardy classes

As we mentioned in Section 4.4 2-D diffusion equations of the form $\mathrm{div}(\sigma\nabla u) = 0$ with real non-negative valued conductivity σ can be viewed as compatibility relations for the so-called conjugate Beltrami equation: $\overline{\partial} f = \nu\overline{\partial f}$ with $\nu = (1-\sigma)/(1+\sigma)$ [4]. Thus, the conjugate Beltrami equation is a means to replace the initial second order diffusion equation by a first order system of two real equations, merged into a single complex one. Hardy spaces under study here are those of this conjugate Beltrami equation: they are comprised of solutions to that equation in the considered domain whose L^p means over curves tending to the boundary of the domain remain bounded. They will for example replace holomorphic Hardy spaces in Problem (P) when dealing with non-constant (isotropic) conductivity. Their traces merely lie in L^p $(1 < p < \infty)$, which is suitable for identification from point-wise measurements, and turn out to be dense on strict subsets of the boundary. This allows one to state Cauchy problems as bounded extremal issues in L^p classes of generalized analytic functions, in a reminiscent manner of what was done for analytic functions as discussed in Section 3.3.1 .

The study of such Hardy spaces for Lipschitz σ was reduced in [4] to that of spaces of pseudo-holomorphic functions with bounded coefficients, which were apparently first considered on the disk by S. Klimentov. Typical results here are that solution factorize as $e^s F$, where F is a holomorphic Hardy function while s is in the Sobolev space $W^{1,r}$ for all $r < \infty$ (Bers factorization), and the analog to the M. Riesz theorem which amounts to solvability of the Dirichlet problem for the initial conductivity equation with L^p boundary data for all $p \in (1,\infty)$. Over the last two years, the case of $W^{1,q}$ conductivities over finitely connected domains, $q > 2$, has been carried out in [13] [61].

In 2013, completing a study begun last year in the framework of the PhD of S. Chaabi, we established similar results in the case where $\log\sigma$ lies in $W^{1,2}$, which corresponds to the critical exponent in Vekua's theory of pseudo-holomorphic functions. This is completely new, and apparently the first example of a solvable Dirichlet problem with L^p boundary data where the conductivity can be both unbounded an vanishing at some places. Accordingly, solutions may also be unbounded inside the domain of the equation, that is, the maximum principle no longer holds. The proof develops a refinement of the Bers factorization based on Muckenhoupt weights and on an original multiplier theorem for $\log W^{1,2}$ functions. A paper on this topic has been submitted [28].

The PhD work of S. Chaabi (defended December 2) contains further work on the Weinstein equation and certain generalizations thereof. This equation results from 2-D projection of Laplace's equation in the presence of rotation symmetry in 3-D. In particular, it is the equation governing the free boundary problem of plasma confinement in the plane section of a tokamak. A method dwelling on Fokas's approach to elliptic boundary value problems has been developed which uses Lax pairs and solves for a Riemann-Hilbert problem on a Riemann surface. It was used to devise semi-explicit forms of solutions to Dirichlet and Neumann problems for the conductivity equation satisfied by the poloidal flux.

In another connection, the conductivity equation can also be regarded as a static Schrödinger equation for smooth coefficients. In particular, a description of laser beam propagation in photopolymers can be crudely approximated by a stationary two-dimensional model of wave propagation in a medium with negligible change of refractive index. In this setting, Helmholtz equation is approximated by a linear Schrödinger equation with one spatial coordinate as evolutionary variable. This phenomenon can be described by a non-stationary model that relies on a spatial nonlinear Schrödinger (NLS) equation with time-dependent refractive index. A model problem has been considered in [20], when the rate of change of refractive index is proportional to the squared amplitude of the electric field and the spatial domain is a plane.

We have also studied composition operators on generalized Hardy spaces in the framework of [13]. In the work [32] submitted for publication, we provide necessary and/or sufficient conditions on the composition map, depending on the geometry of the domains, ensuring that these operators are bounded, invertible, isometric or compact.

6.2.2. Best constrained analytic approximation

Several questions about the behavior of solutions to the bounded extremal problem (P) of Section 3.3.1 have been considered. For instance, truncated Toeplitz operators have been studied in [17], that can be used to quantify robustness properties of our resolution schemes in H^2 and to establish error estimates. Moreover we considered additional interpolation constraints on the disk in Problem (P), and derived new stability estimates for the solution [46]. Such interpolation constraints arise naturally in inverse boundary problems like plasma shaping in last generation tokamaks, where some measurements are performed inside the chamber 4.4 . Of course the version studied so far is much simplified, as it must be carried over to non-constant conductivities and annular geometries.

6.3. Synthesis of compact multiplexers and de-embedding of multiplexers

Participants: Martine Olivi, Sanda Lefteriu, Fabien Seyfert.

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

6.3.1. Synthesis of compact multiplexers

We focused our research on multiplexer with a star topology. These are comprised of a central N-port junction, and of filters plugged on all but common ports (see Figure 9). A possible approach to synthesis of the multiplexer's response is to postulate that each filter channel has to match the multiplexer at n_k frequencies (n_k being the order of the filter) while rejecting the energy at m_k other frequencies (m_k being the order the transmission polynomial of the filter). The desired synthesis can then be cast into computing of a collection of filter's responses matching the energy as prescribed and rejecting it at specified frequencies when plugged simultaneously on the junction. Whether such a collection exists is one of the main open issues facing co-integration of systems in electronics. Investigating the latter led us to consider the simpler problem of matching a filter, on a frequency-varying load, while rejecting energy at fixed specified frequencies. If the order of the filter is n this amounts to fix a given transmission polynomial r and to solve for a unitary polynomial p meeting interpolation conditions of the form:

$$j = 1 \cdots n, \qquad \frac{p}{q}(w_j) = \gamma_j, \quad |\gamma_j| < 1$$

where q is the unique monic Hurwitz polynomial satisfying the Feldtkeller equation

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

This problem can be seen as an extended Nevanlinna-Pick interpolation problem, which was considered in [65] when the interpolation frequencies lie in the *open* left half-plane. Last year we conjectured the existence and uniqueness of a solution, which were eventually proved true this year when r has no roots on the imaginary axis. We already communicated on the subject (9.1), and a scientific report as well as an article are being written on this result [30]. The proof relies on the local invertibility of an evaluation map that is established using a differential argument and the structure of particular Pick matrices. The case where r has zeros on the imaginary axis is of great interest, and though existence then holds again uniqueness is still not well-understood: it is conjectured that under minor restrictions on the localization of the $\gamma_k' s$ (typically off an algebraic subvariety) the main results still hold.

This research lies at heart of our collaboration with CNES on multiplexer synthesis and the core of the starting ANR project COCORAM on co-integration of filters and antennas (see Section 8.1.1).

6.3.2. De-embedding of multiplexers

Let S be the external scattering parameters of a multiplexer composed of a N-port junction with response T and $N-1$ filters with responses $F_1, \cdots F_{N-1}$ as plotted on Figure 9 . The de-embbeding problem concerns the recovery of the F_k and can be considered under different hypotheses. Last year we studied the de-embedding problem where S and T are known [76] but no particular structure on the F_k is assumed. It was shown that for a generic junction T and for $N > 3$ the de-embedding problem has a unique solution. It was however observed that in practice the junction's response is far from being generic (as it is usually obtained by assembly of smaller T-junctions) which renders the problem extremely sensitive to measurement noise. It was also noticed that in practical applications, scattering measurements of the junction are hardly available.

It was therefore natural to consider following de-embedding problem. Given S the external scattering measurement of the multiplexer, and under the assumptions:

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

what can be said about the filter's responses ? It was shown that under the above hypotheses, in particular with no a priori knowledge of T, the filter's responses are identifiable up to a constant chain matrix chained at their second port (nearest to the junction) [24]. It was also shown that this uncertainty bears only on the resonant frequency of the last cavity of each filter, as well as on their output coupling. Most of the filters' important parameters can therefore be recovered. The approach is constructive and relies on rational approximation of certain external scattering parameters, and on an extraction procedure similar to Darlington's synthesis for filters. Software developments have been pursued to implement the latter and practical studies are under way with data furnished by Thales Alenia Space and by Siae Microelettronica. A medium term objective is to extend the Presto-HF (5.3) software to de-embedding problems for multiplexers and more general multi-ports.

This work is pursued in collaboration with Thales Alenia Space, Siae Microelettronica, XLIM and CNES in particular under contract with CNES on compact N-port synthesis (see Section 7.1).

6.4. Detection of the instability of amplifiers

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

This work is conducted in collaboration with Jean-Baptiste Pomet from the McTao team. It is a continuation of a collaboration with CNES and the University of Bilbao.The goal is to help developing amplifiers, in particular to detect instability at an early stage of the design.

Figure 9. Multiplexer made of a junction T and filtering devices $F_1, F_2 \cdots F_N$

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

In non-degenerate cases, non-linear electrical components can be replaced by their first order approximation when studying stability to small perturbations. Using this approximation, diodes appear as perfect negative resistors and transistors as perfect current sources controlled by the voltages at certain points of the circuit.

In previous years, we had proved that the class of transfer functions which can be realized with such ideal components and standard passive components (resistors, selfs, capacitors and transmission lines) is rather large since it contains all rational functions in the variable and in the exponentials thereof. This makes possible to design circuits that are unstable, although they have no pole in the right half-plane. This remains true even if a high resistor is put in parallel of the circuit, which is rather unusual. These pathological examples are unrealistic, though, because they assume that non-linear elements continue to provide gain even at very high frequencies. In practice, small capacitive and inductive effects (negligible at moderate frequencies) make these components passive for very high frequencies.

In 2013, we showed that under this simple assumption that there are small inductive and capacitive effects in active components, the class of transfer functions of realistic circuits is much smaller than in previous situation. Our main result is that a realistic circuit is unstable if and only if it has poles in the right half-plane. Moreover, there can only be finitely many of them. Besides this result, we also generalized our description of the class of transfer functions achievable with ideal components, to include the case of transmission lines with loss. An article is currently being written on this subject.

6.5. Rational and meromorphic approximation

Participants: Laurent Baratchart, Sylvain Chevillard.

This work has been done in collaboration with Herbert Stahl (Beuth-Hochsch.), Maxim Yattselev (Purdue Univ. at Indianapolis, USA), Tao Qian (Univ. Macao).

We published last year an important result in approximation theory, namely the counting measure of poles of best H^2 approximants of degree n to a function analytically continuable, except over finitely many branchpoints lying outside the unit disk, converges to the Green equilibrium distribution of the compact set of minimal Green capacity outside of which the function is single valued [6] (see also [21]). This result warrants source recovery techniques used in Section 6.1.1 . We considered this year a similar problem for best uniform meromorphic approximants on the unit circle (so-called AAK approximants after Adamjan, Arov and Krein), in the case where the function may have poles and essential singularities. The technical difficulties are considerable, and though a line of attack has been adopted we presently struggle with the proof.

We also studied partial realizations, or equivalently Padé approximants to transfer functions with branchpoints. Identification techniques based on partial realizations of a stable infinite-dimensional transfer function are known to often provide unstable models, but the question as to whether this is due to noise or to intrinsic instability was not clear. This year, we published a paper showing that, in the case of 4 branchpoints, the pole behavior generically has deterministic chaos to it [15].

We also considered the issue of lower bounds in rational approximation. Prompted by renewed interest for linearizing techniques such as vector fitting in the identification community, we studied linearized errors in light of the topological approach in [51], to find that, when properly normalized, they give rise to lower bounds in L^2 rational approximation. Moreover, these make contact with AAK theory which furnishes more, easily computable lower bounds. This is an interesting finding, for lower bounds are usually difficult to get in approximation and though quite helpful to get an appraisal of what can be hoped for in modeling. Dwelling on this, we established for the first time lower bounds in L^2 rational approximation to some badly L^∞ approximable functions (Blaschke products) and showed equivalence, up to a constant, of best L^2 and L^∞ approximation to functions with branchpoints (such as those appearing in inverse source problems for

EEG, see Section 6.1.1). An article on this subject is currently submitted for publication in the Journal of Approximation Theory [29].

6.6. Tools for numerically guaranteed computations

Participant: Sylvain Chevillard.

The overall and long-term goal is to enhance the quality of numerical computations. The progress made during year 2013 is the following:

- Publication of a work with Marc Mezzarobba (who was with Aric project-team at that time, and who is now with LIP6) about the efficient evaluation of the Airy $\mathrm{Ai}(x)$ function when x is moderately large [22]. The Taylor series of the Airy Ai function (as many others such as, e.g., Bessel functions or erf) is ill-conditioned when x is not small. To overcome this difficulty, we extend a method by Gawronski, Müller and Reinhard, known to solve the issue in the case of the error function erf. We rewrite $\mathrm{Ai}(x)$ as $G(x)/F(x)$ where F and G are two functions with well-conditioned series. However, the coefficients of G turn out to obey a three-terms ill-conditioned recurrence. We evaluate this recurrence using Miller's backward algorithm with a rigorous error analysis. Function Ai is an example, but ideally the process could be automated to handle some appropriate class of functions in a future work.
- A more general endeavor is to develop a tool that helps developers of libms in their task. This is performed by the software Sollya [3], developed in collaboration with C. Lauter (Université Pierre et Marie Curie) and M. Joldeş (LAAS). In 2013, we released version version 4.0 (in May) and 4.1 (in November) of Sollya. Among other things these releases make available to the user all features of Sollya as a C library. They also introduce the possibility of computing Chebyshev models, and a generalization of Remez algorithm allowing the user to compute a L^∞ best approximation of a real-valued function on a bounded real interval by any linear combination of given functions.

[3] http://sollya.gforge.inria.fr/

ARAMIS Team

6. New Results

6.1. Spatial and anatomical regularization of SVM

Participants: Rémi Cuingnet, Joan Glaunès, Marie Chupin, Habib Benali, Olivier Colliot [Correspondant].

We developed a general framework to introduce spatial and anatomical priors in SVM for brain image analysis based on regularization operators. A notion of proximity based on prior anatomical knowledge between the image points is defined by a graph (e.g. brain connectivity graph) or a metric (e.g. Fisher metric on statistical manifolds). A regularization operator is then defined from the graph Laplacian, in the discrete case, or from the Laplace-Beltrami operator, in the continuous case. The regularization operator is then introduced into the SVM, which exponentially penalizes high frequency components with respect to the graph or to the metric and thus constrains the classification function to be smooth with respect to the prior. It yields a new SVM optimization problem whose kernel is a heat kernel on graphs or on manifolds. We then present different types of priors and provide efficient computations of the Gram matrix. The proposed framework is finally applied to the classification of brain magnetic resonance (MR) images (based on gray matter concentration maps and cortical thickness measures) from 137 patients with Alzheimer's disease and 162 elderly controls. The results demonstrate that the proposed classifier generates less-noisy and consequently more interpretable feature maps (Figure 1) with high classification performances.

More details in [4].

6.2. Segmentation of the hippocampus in neurodegenerative dementias

Participants: Leonardo Cruz de Souza, Marie Chupin, Maxime Bertoux, Stéphane Lehéricy, Bruno Dubois, Foudil Lamari, Isabelle Le Ber, Michel Bottlaender, Olivier Colliot [Correspondant], Marie Sarazin.

Our team develops various applications of our automatic segmentation method SACHA to neurological disorders, in particular in neurodegenerative dementias. This research is done in close collaboration with IM2A (Institut de la Mémoire et de la Maladie d'Alzheimer, Bruno Dubois and Marie Sarazin) at Pitié-Salpêtrière hospital.

We previously showed that automatic hippocampal segmentation can discriminate patients with Alzheimer's disease (AD) from elderly control subjects, with high sensitivity and specificity. In patients with Alzheimer's disease, we further studied the relationship between hippocampal atrophy and memory deficits. We also showed that hippocampal volume loss is correlated to tau and hyperphosphorylated tau levels measured in the cerebro-spinal fluid (CSF) but not with $A\beta_4 2$ levels.

Here, our objective was to study the ability of hippocampal volumetry (HV) to differentiate between two neurodegenerative dementias: Alzheimer's disease (AD) and fronto-temporal dementia (FTD). Seventy-two participants were included: 31 AD patients with predominant and progressive episodic memory deficits associated with typical AD cerebrospinal fluid (CSF) profile and/or positive amyloid imaging (PET with 11C-labeled Pittsburgh Compound B [PiB]), 26 patients with behavioral variant FTD (bvFTD) diagnosed according to consensual clinical criteria and with no AD CSF profile, and 15 healthy controls without amyloid retention on PiB-PET exam. HV were segmented with our automated method and were normalized to total intracranial volume (nHV). Significant reductions in HV were found in both AD and bvFTD patients compared with controls, but there were no significant difference between AD and bvFTD patients. Mean nHV distinguished normal controls from either AD or bvFTD with high sensitivity (80.6% and 76.9%, respectively) and specificity (93.3% for both), but it was inefficient in differentiating AD from bvFTD (9.7% specificity). There was no difference in the clinical and neuropsychological profiles according to HV in bvFTD and AD patients. In conclusion, when considered alone, measures of HV are not good markers to differentiate AD from bvFTD. Hippocampal sclerosis associated with FTD may explain the high degree of overlap in nHV between both groups.

../../../../projets/aramis/IMG/fig20eps.png

Figure 1. Anatomical regularization of support vector machines for automatic classification of patients with Alzheimer's disease. The figure displays the normalized vector orthogonal to the optimal margin hyperplane, for increasing levels of regularization.

More details in [5].

6.3. Diffeomorphic Iterative Centroids for Template Estimation on Large Datasets

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

A common approach for analysis of anatomical variability relies on the estimation of a template representative of the population. The Large Deformation Diffeomorphic Metric Mapping is an attractive framework for that purpose. However, template estimation using LDDMM is computationally expensive, which is a limitation for the study of large datasets. We proposed an iterative method which quickly provides a centroid of the population in the shape space. This centroid can be used as a rough template estimate or as initialization of a template estimation method. The approach was evaluated on datasets of real and synthetic hippocampi segmented from brain MRI. The results showed that the centroid is correctly centered within the population and is stable for different orderings of subjects. When used as an initialization, the approach allows to substantially reduce the computation time of template estimation.

More details in [30].

6.4. Sparse Adaptive Parameterization of Variability in Image Ensembles

Participants: Stanley Durrleman [Correspondant], Sarang Joshi, Stéphanie Allassonnière.

We introduce a new parameterization of diffeomorphic deformations for the characterization of the variability in image ensembles. Dense diffeomorphic deformations are built by interpolating the motion of a finite set of control points that forms a Hamiltonian flow of self-interacting particles. The proposed approach estimates a template image representative of a given image set, an optimal set of control points that focuses on the most variable parts of the image, and template-to-image registrations that quantify the variability within the image set. The method automatically selects the most relevant control points for the characterization of the image variability and estimates their optimal positions in the template domain. The optimization in position is done during the estimation of the deformations without adding any computational cost at each step of the gradient descent. The selection of the control points is done by adding a L^1 prior to the objective function, which is optimized using the FISTA algorithm.

Related publication: [12]

6.5. Toward a comprehensive framework for the spatiotemporal statistical analysis of longitudinal shape data

Participants: Stanley Durrleman [Correspondant], Xavier Pennec, Alain Trouvé, José Braga, Guido Gerig, Nicholas Ayache.

We introduce a comprehensive framework for the statistical analysis of longitudinal shape data. The proposed method allows the characterization of typical growth patterns and subject-specific shape changes in repeated time-series observations of several subjects. This can be seen as the extension of usual longitudinal statistics of scalar measurements to high-dimensional shape or image data.

The method is based on the estimation of continuous subject-specific growth trajectories and the comparison of such temporal shape changes across subjects. Differences between growth trajectories are decomposed into morphological deformations, which account for shape changes independent of time, and time warps, which account for different rates of shape changes over time.

Given a longitudinal shape data set, we estimate a mean growth scenario representative of the population, and the variations of this scenario both in terms of shape changes and in terms of change in growth speed. Then, intrinsic statistics are derived in the space of spatiotemporal deformations, which characterize the typical variations in shape and in growth speed within the studied population. They can be used to detect systematic developmental delays across subjects.

../../../../projets/aramis/IMG/IJCV_Digit2.jpg

Figure 2. Left: template image estimated from 20 images of the US postal database. Momentum vectors are placed at the most variable places and paramterize mappings from the template to each image in the data set. Right: sample images from the data set (top) and template image deformed to match the corresponding sample image (bottom)

In the context of neuroscience, we apply this method to analyze the differences in the growth of the hippocampus in children diagnosed with autism, developmental delays and in controls. Result suggest that group differences may be better characterized by a different speed of maturation rather than shape differences at a given age. In the context of anthropology, we assess the differences in the typical growth of the endocranium between chimpanzees and bonobos. We take advantage of this study to show the robustness of the method with respect to change of parameters and perturbation of the age estimates.

Related publication: [13]

6.6. Bayesian Atlas Estimation for the Variability Analysis of Shape Complexes

Participants: Pietro Gori [Correspondant], Olivier Colliot, Yulia Worbe, Linda Marrakchi-Kacem, Sophie Lecomte, Cyril Poupon, Andreas Hartmann, Nicholas Ayache, Stanley Durrleman.

We propose a Bayesian framework for multi-object atlas estimation based on the metric of currents which permits to deal with both curves and surfaces without relying on point correspondence. This approach aims to study brain morphometry as a whole and not as a set of different components, focusing mainly on the shape and relative position of different anatomical structures which is fundamental in neuro-anatomical studies. We propose a generic algorithm to estimate templates of sets of curves (fiber bundles) and closed surfaces (sub-cortical structures) which have the same "form" (topology) of the shapes present in the population. This atlas construction method is based on a Bayesian framework which brings to two main improvements with respect to previous shape based methods. First, it allows to estimate from the data set a parameter specific to each object which was previously fixed by the user: the trade-off between data-term and regularity of deformations. In a multi-object analysis these parameters balance the contributions of the different objects and the need for an automatic estimation is even more crucial. Second, the covariance matrix of the deformation parameters is estimated during the atlas construction in a way which is less sensitive to the outliers of the population.

Related publication: [33]

6.7. Geodesic regression of shape and image data

Participants: James Fishbaugh [Correspondant], Marcel Prastawa, Guido Gerig, Stanley Durrleman.

Shape regression is emerging as an important tool for the statistical analysis of time dependent shapes. We develop a new generative model which describes shape change over time, by extending simple linear regression to the space of shapes represented as currents in the large deformation diffeomorphic metric mapping (LDDMM) framework. By analogy with linear regression, we estimate a baseline shape (intercept) and initial momenta (slope) which fully parameterize the geodesic shape evolution. This is in contrast to previous shape regression methods which assume the baseline shape is fixed. We further leverage a control point formulation, which provides a discrete and low dimensional parameterization of large diffeomorphic transformations. This flexible system decouples the parameterization of deformations from the specific shape representation, allowing the user to define the dimensionality of the deformation parameters. We present an optimization scheme that estimates the baseline shape, location of the control points, and initial momenta simultaneously via a single gradient descent algorithm.

Shapes can be given as 3D meshes (as in [32]) or as 3D images (as in [31]).

Related publications: [32], [31].

6.8. Discriminating brain microbleeds using phase contrast MRI in a multicentre clinical dataset

Participants: Takoua Kaaouana [Correspondant], Marie Chupin, Didier Dormont, Ludovic de Rochefort, Thomas Samaille.

../../../../projets/aramis/IMG/IJCV_4D_Regression.jpg

a- Temporal regression of endocasts of bonobos (top) and chimpanzees (bottom)

../../../../projets/aramis/IMG/IJCV_4D_Registration.jpg

b - morphological deformation

../../../../projets/aramis/IMG/Pietro_Final_template.png

Figure 4. Atlas construction from a data set of left caudate nucleus and its associated fiber bundle that were segmented in images of patients with Gilles de la Tourette syndrome and controls. An initial template complex determines the topology of the model. Its shape is optimized given the patients data or the controls data only, thus resulting in two atlases showing different distributions of the fibers on the surface of the nucleus.

../../../../projets/aramis/IMG/IPMI_James.png

Figure 5. Average development of genu fiber tract from 2 to 24 months. Top row shows observed data for all subjects, which is clustered around 2, 12, and 24 months. Bottom row shows genu fiber tracts estimated from geodesic regression at several time points with velocity of fiber development displayed on the estimated fibers.

Brain microbleeds (BMBs) have emerged as a new imaging marker of small vessel diseases and they may play a crucial role in degenerative pathology such as Alzheimer's disease. Composed of hemosiderin, BMBs can be efficiently detected with MRI sequences sensitive to magnetic susceptibility (e.g. gradient recalled echo T2*W images). Nevertheless, that identification remains challenging because of confounding structures and lesions. Most T2*-weighted hyposignals result from local magnetic field inhomogeneity and can be identified either as BMBs, veins or brain micro-calcifications (BMCs). Differential diagnosis of BMBs and BMCs usually requires an additional CT scan. Quantitative susceptibility mapping techniques were proposed to discriminate between diamagnetic and paramagnetic structures, but they require a full 3D dataset and complex post-processing. We introduced a fast 2D phase processing technique including unwrapping and harmonic filtering thus yielding the internal field map, namely the field map generated only by sources within the volume of interest. We demonstrate its applicability and robustness on multicenter data acquired in standardized clinical settingand and its ability to discriminate between paramagnetic BMBs and diamagnetic BMCs through the use of the orientation of the dipolar pattern.

Related publications: [36].

6.9. Network symmetries and functional modules in the brain

Participants: Vincenzo Nicosia, Miguel Valencia, Mario Chavez [Correspondant], Albert Diaz-Guilera, Vito Latora.

We study the classical Kuramoto model in which the oscillators are associated to the nodes of a network and the interactions include a phase frustration, thus preventing full synchronization. The system organizes into a regime of remote synchronization where pairs of nodes with the same network symmetry are fully synchronized, despite their distance on the graph. We provide analytical arguments to explain this result and we show how the frustration parameter affects the distribution of phases. An application to brain networks suggests that anatomical symmetry plays a role in neural synchronization by determining correlated functional modules across distant locations.

Related publication: [19]

6.10. Accessibility of cortical networks during motor tasks

Participants: Mario Chavez [Correspondant], Fabrizio de Vico Fallani, Miguel Valencia, Mario Chavez, Julio Artieda, Vito Latora, Donatella Mattia, Fabio Babiloni.

Recent findings suggest that the preparation and execution of voluntary self-paced movements are accompanied by the coordination of the oscillatory activities of distributed brain regions. We used electroencephalographic source imaging methods to estimate the cortical movement-related oscillatory activity during finger ex- tension movements. We applied network theory to investigate changes (expressed as differences from the baseline) in the connectivity structure of cortical networks related to the preparation and execution of the movement. We computed the topological accessibility of different cortical areas, measuring how well an area can be reached by the rest of the network. Analysis of cortical networks revealed specific agglomerates of cortical sources that become less accessible during the preparation and the execution of the finger movements. The observed changes neither could be explained by other measures based on geodesics or on multiple paths, nor by power changes in the cortical oscillations.

Related publication: [3]

6.11. Abnormal functional connectivity between motor cortex and pedunculopontine nucleus following chronic dopamine depletion

Participants: Miguel Valencia, Mario Chavez [Correspondant], Julio Artieda, J. Paul Bolam, Juan Mena-Segovia.

../../../../projets/aramis/IMG/ImageTakoua.png

Figure 6. Philips (a,b,c,d) and Siemens (e,f,g,h) sample cases. Magnitude image (a,e), native phase image (b,f), and internal field map; axial (c,g) and sagittal (d,h). Zoom in white rectangle showing a dipolar pattern BMB (white arrow) and a physiologic calcification of the choroid plexus (black arrow).

../../../../projets/aramis/IMG/synchSymmetry.png

Figure 7. a) Brain areas with similar and dissimilar phases of the frustrated Kuramoto model are colored and superimposed onto an anatomical image. b) Examples of functional data from one subject recorded at the brain areas indicated in panel a). Colors are the same as those used in the anatomical image. c) Functional correlation (normalised values) Z between pairs of nodes as a function of their phase differences $\Delta\theta$ according to the simulated Kuramoto dynamics. The black solid curve corresponds to the average value over all the subjects, while the gray area covers the 5^{th} and the 95^{th} percentiles of the distribution. The dashed horizontal line indicates the threshold for statistical significant correlations ($p < 0.05$, corrected for multiple comparisons).

../../../../projets/aramis/IMG/generalScheme4eps.png

Figure 8. a) Averaged EMG and EEG (recorded at the postcentral region) signals of a subject during the execution of finger movements. Boxes define the three temporal epochs of EEG activity studied here: baseline (BASE), preparation (PRE) and execution period (EXE). Vertical dotted line indicates the movement onset. Examples of scalp and source-level networks obtained from one subject, at the frequency band Beta1, during the epoch EXE are shown in panels b) and c), respectively. Color map codes the number of connections.

The activity of the basal ganglia is altered in Parkinson's disease (PD) as a consequence of the degeneration of dopamine neurons in the substantia nigra pars compacta. This results in aberrant discharge patterns and expression of exaggerated oscillatory activity across the basal ganglia circuit. Altered activity has also been reported in some of the targets of the basal ganglia, including the pedunculopontine nucleus (PPN), possibly due to its close interconnectivity with most regions of the basal ganglia. However, the nature of the involvement of the PPN in the pathophysiology of PD has not been fully elucidated. We recorded local field potentials in the motor cortex and the PPN in the 6-hydroxydopamine (6-OHDA)-lesioned rat model of PD under urethane anesthesia. By means of linear and nonlinear statistics, we analyzed the synchrony between the motor cortex and the PPN, and the delay in the interaction between these two structures. We observed the presence of coherent activity between the cortex and the PPN in low- (5-15 Hz) and high-frequency bands (25-35 Hz) during episodes of cortical activation. In each case the cortex led the PPN. Dopamine depletion strengthened the interaction of the low-frequency activities by increasing the coherence specifically in the theta and alpha ranges and reduced the delay of the interaction in the gamma band. Our data show that cortical inputs play a determinant role in leading the coherent activity with the PPN, and support the involvement of the PPN in the pathophysiology of PD.

Related publication: [25]

6.12. Subthalamic Nucleus High-Frequency Stimulation Restores Altered Electrophysiological Properties of Cortical Neurons in Parkinsonian Rat

Participants: Bertrand Degos, Jean Michel Deniau, Mario Chavez [Correspondant], Nicolas Maurice.

Electrophysiological recordings performed in parkinsonian patients and animal models have confirmed the occurrence of alterations in firing rate and pattern of basal ganglia neurons, but the outcome of these changes in thalamo-cortical networks remains unclear. Using rats rendered parkinsonian, we investigated, at a cellular level in vivo, the electrophysiological changes induced in the pyramidal cells of the motor cortex by the dopaminergic transmission interruption and further characterized the impact of high-frequency electrical stimulation of the subthalamic nucleus, a procedure alleviating parkinsonian symptoms. We provided evidence that a lesion restricted to the substantia nigra pars compacta resulted in a marked increase in the mean firing rate and bursting pattern of pyramidal neurons of the motor cortex. These alterations were underlain by changes of the electrical membranes properties of pyramidal cells including depolarized resting membrane potential and increased input resistance. The modifications induced by the dopaminergic loss were more pronounced in cortico-striatal than in cortico-subthalamic neurons. Furthermore, subthalamic nucleus high- frequency stimulation applied at parameters alleviating parkinsonian signs regularized the firing pattern of pyramidal cells and restored their electrical membrane properties.

Related publication: [7]

6.13. Non-parametric resampling of random walks for spectral networks clustering

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

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

Related publication: [43]

6.14. Multiscale topological properties of functional brain networks during motor imagery after stroke

Participants: Fabrizio de Vico Fallani [Correspondant], Floriana Pichiorri, Giovani Morone, Marco Molinari, Fabio Babiloni, Febo Cincotti, Donatella Mattia.

In recent years, network analyses have been used to evaluate brain reorganization following stroke. However, many studies have often focused on single topological scales, leading to an incomplete model of how focal brain lesions affect multiple network properties simultaneously and how changes on smaller scales influence those on larger scales. In an EEG-based experiment on the performance of hand motor imagery (MI) in 20 patients with unilateral stroke, we observed that the anatomic lesion affects the functional brain network on multiple levels. In the beta (13–30 Hz) frequency band, the MI of the affected hand (Ahand) elicited a significantly lower smallworldness and local efficiency (Eloc) versus the unaffected hand (Uhand). Notably, the abnormal reduction in Eloc significantly depended on the increase in interhemispheric connectivity, which was in turn determined primarily by the rise of regional connectivity in the parieto-occipital sites of the affected hemisphere. Further, in contrast to the Uhand MI, in which significantly high connectivity was observed for the contralateral sensorimotor regions of the unaffected hemisphere, the regions with increased connectivity during the Ahand MI lay in the frontal and parietal regions of the contralaterally affected hemisphere. Finally, the overall sensorimotor function of our patients, as measured by Fugl–Meyer Assessment (FMA) index, was significantly predicted by the connectivity of their affected hemisphere. These results improve on our understanding of stroke-induced alterations in functional brain networks.

Related publication: [6]

6.15. Wavelet analysis in ecology and epidemiology: impact of statistical tests

Participants: Bernard Cazelles, Kevin Cazelles, Mario Chavez [Correspondant].

Wavelet analysis is now frequently used to extract information from ecological and epidemiological time series. Statistical hypothesis tests are conducted on associated wavelet quantities to assess the likelihood that they are due to a random process. Such random processes represent null models and are generally based on synthetic data that share some statistical characteristics with the original time series. This allows the comparison of null statistics with those obtained from original time series. When creating synthetic datasets, different techniques of resampling result in different characteristics shared by the synthetic time series. Therefore, it becomes crucial to consider the impact of the resampling method on the results. We have addressed this point by comparing seven different statistical testing methods applied with different real and simulated data. Our results showed that statistical assessment of periodic patterns is strongly affected by the choice of the resampling method, so two different resampling techniques could lead to two different conclusions about the same time series. Moreover, we showed the inadequacy of resampling series generated by white noise and red noise that are nevertheless the methods currently used in the wide majority of wavelets applications in epidemiology. Our results highlight that the characteristics of a time series, namely its Fourier spectrum and autocorrelation, are important to consider when choosing the resampling technique. Results suggest that data-driven resampling methods should be used such as the hidden Markov model algorithm and the 'beta-surrogate' method.

Related publication: [2]

../../../../projets/aramis/IMG/figureFabrizio.jpg

Figure 9. Grand average of brain networks in the Beta band during the MI of the unaffected Uhand and affected Ahand hand. Top plots: Scalp representation relative to Uhand (panel A) and Ahand (panel B) condition. Nodes are positioned according the actual EEG montage scheme. Blue and red lines denote the links within the unaffected (Uhemi) and the affected (Ahemi) hemisphere, respectively. Gray lines denote the inter-hemispheric links. The intensity of the color and the thickness of the lines vary as function of the number of patients exhibiting that significant link. Bottom part: graph representation of the brain networks relative to Uhand (panel A) and Ahand (panel B) condition. In this representation nodes are spatially repositioned through a force-based algorithm so that all the links are approximately of equal length with as few crossing edges as possible. Only links that were in common to more than 4 patients (20% of the sample) are illustrated here. Blue and red nodes indicate scalp electrodes placed over the undamaged (Uhemi) and damaged (Ahemi) hemisphere, respectively. The midline scalp electrodes (from Fpz to Oz) are illustrated as white nodes

ARIC Project-Team

6. New Results

6.1. Cryptography and lattices

6.1.1. Group signatures

Group signatures are cryptographic primitives where users can anonymously sign messages in the name of a population they belong to. Gordon et al. (Asiacrypt 2010) suggested the first realization of group signatures based on lattice assumptions in the random oracle model. A significant drawback of their scheme is its linear signature size in the cardinality N of the group. A recent extension proposed by Camenisch et al. (SCN 2012) suffers from the same overhead.

F. Laguillaumie, A. Langlois, B. Libert (Technicolor), and D. Stehlé described in [24] the first lattice-based group signature schemes where the signature and public key sizes are essentially logarithmic in N (for any fixed security level). Their basic construction only satisfies a relaxed definition of anonymity (just like the Gordon et al. system) but readily extends into a fully anonymous group signature (i.e., that resists adversaries equipped with a signature opening oracle). They proved the security of their schemes in the random oracle model under the SIS and LWE assumptions.

6.1.2. Classical hardness of learning with errors

Z. Brakerski (Stanford U.), A. Langlois, C. Peikert (Georgia Institute of Technology), O. Regev (Courant Institute, New York U.), and D. Stehlé showed in [16] that the Learning with Errors (LWE) problem is classically at least as hard as standard worst-case lattice problems, even with polynomial modulus. Previously this was only known under quantum reductions. Their techniques capture the tradeoff between the dimension and the modulus of LWE instances, leading to a much better understanding of the landscape of the problem. The proof is inspired by techniques from several recent cryptographic constructions, most notably fully homomorphic encryption schemes.

6.1.3. Improved Zero-knowledge Proofs of Knowledge for the ISIS Problem, and Applications

In all existing efficient proofs of knowledge of a solution to the infinity norm Inhomogeneous Small Integer Solution ISIS_∞ problem, the knowledge extractor outputs a solution vector that is only guaranteed to be $\widetilde{O}(n)$ times longer than the witness possessed by the prover. As a consequence, in many cryptographic schemes that use these proof systems as building blocks, there exists a gap between the hardness of solving the underlying ISIS_∞ problem and the hardness underlying the security reductions. Together with S. Ling, K. Nguyen, and H. Wang (Nanyang Technological University, Singapore), D. Stehlé generalized in [26] Stern's protocol to obtain two statistical zero-knowledge proofs of knowledge for the ISIS_∞ problem that remove this gap. Their result yields the potential of relying on weaker security assumptions for various lattice-based cryptographic constructions. As applications of their proof system, they introduced a concurrently secure identity-based identification scheme based on the worst-case hardness of the $\mathrm{SIVP}_{\widetilde{O}(n^{1.5})}$ problem (in the L2 norm) in general lattices in the random oracle model, and an efficient statistical zero-knowledge proof of plaintext knowledge with small constant gap factor for Regev's encryption scheme.

6.1.4. Decoding by Embedding: Correct Decoding Radius and DMT Optimality

In lattice-coded multiple-input multiple-output (MIMO) systems, optimal decoding amounts to solving the closest vector problem (CVP). Embedding is a powerful technique for the approximate CVP, yet its remarkable performance is not well understood. In [8], C. Ling (Imperial College, London), L. Luzzi (ENSEA, U. Cergy Pontoise), and D. Stehlé analyzed the embedding technique from a bounded distance decoding (BDD) viewpoint. They proved that the Lenstra, Lenstra and Lovász (LLL) algorithm can achieve $1/(2\gamma)$-BDD for $\gamma \approx O(2^{n/4})$, yielding a polynomial-complexity decoding algorithm performing exponentially better than Babai's which achieves $\gamma = O(2^{n/2})$. This substantially improves the existing result $\gamma = O(2^n)$ for embedding decoding. They also proved that BDD of the regularized lattice is optimal in terms of the diversity-multiplexing gain tradeoff (DMT).

6.1.5. A New View on HJLS and PSLQ: Sums and Projections of Lattices

The HJLS and PSLQ algorithms are the de facto standards for discovering non-trivial integer relations between a given tuple of real numbers. In [19], J. Chen, D. Stehlé, and G. Villard provided a new interpretation of these algorithms, in a more general and powerful algebraic setup: they view them as special cases of algorithms that compute the intersection between a lattice and a vector subspace. Further, they extracted from them the first algorithm for manipulating finitely generated additive subgroups of a Euclidean space, including projections of lattices and finite sums of lattices. They adapted the analyses of HJLS and PSLQ to derive correctness and convergence guarantees. They also investigated another approach based on embedding the input in a higher dimensional lattice and calling the LLL lattice reduction algorithm.

6.2. Certified computing and computer algebra

6.2.1. Polynomial system solving

Polynomial system solving is a core topic of computer algebra. While the worst-case complexity of this problem is known to be hopelessly large, the practical complexity for large families of systems is much more reasonable. Progress has been made in assessing precise complexity estimates in this area.

First, M. Bardet (U. Rouen), J.-C. Faugère (PolSys team), and B. Salvy studied the complexity of Gröbner bases computations, in particular in the generic situation where the variables are in simultaneous Noether position with respect to the system. They gave a bound on the number of polynomials of each degree in a Gröbner basis computed by Faugère's F5 algorithm in this generic case for the grevlex ordering (which is also a bound on the number of polynomials for a reduced Gröbner basis), and used it to bound the exponent of the complexity of the F5 algorithm [35].

Next, a fundamental problem in computer science is to find all the common zeroes of m quadratic polynomials in n unknowns over 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 [1], M. Bardet (U. Rouen), J.-C. Faugère (PolSys team), B. Salvy, and P.-J. Spaenlehauer (CARAMEL team) gave 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, they showed that the deterministic variant of their 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 showed that the algebraic assumptions are satisfied with probability very close to 1. They have also given a rough estimate for the actual threshold between their method and exhaustive search, which is as low as 200, and thus very relevant for cryptographic applications.

6.2.2. Linear differential equations

Creative telescoping algorithms compute linear differential equations satisfied by multiple integrals with parameters. Together with A. Bostan and P. Lairez (SpecFun team), B. Salvy described a precise and elementary algorithmic version of the Griffiths–Dwork method for the creative telescoping of rational functions. This leads to bounds on the order and degree of the coefficients of the differential equation, and to the first complexity result which is simply exponential in the number of variables. One of the important features of the algorithm is that it does not need to compute certificates. The approach is vindicated by a prototype implementation [15].

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

With F. Johansson and M. Kauers (RISC, Linz, Austria), M. Mezzarobba presented in [23] a new algorithm for computing hyperexponential solutions of ordinary linear differential equations with polynomial coefficients. The algorithm relies on interpreting formal series solutions at the singular points as analytic functions and evaluating them numerically at some common ordinary point. The numerical data is used to determine a small number of combinations of the formal series that may give rise to hyperexponential solutions.

6.2.3. Exact linear algebra

Transforming a matrix over a field to echelon form, or decomposing the matrix as a product of simpler matrices that reveal the rank profile, is a fundamental building block of computational exact linear algebra. For such tasks the best previously available algorithms were either rank sensitive (i.e., of complexity expressed in terms of the exponent of matrix multiplication and the rank of the input matrix) or in place (i.e., using essentially no more memory that what is needed for matrix multiplication). In [6] C.-P. Jeannerod, C. Pernet, and A. Storjohann (U. Waterloo, Canada) have proposed algorithms that are both rank sensitive and in place. These algorithms required to introduce a matrix factorization of the form $A = CUP$ with C a column echelon form giving the row rank profile of the input matrix A, U a unit upper triangular matrix, and P a permutation matrix.

6.2.4. Certified multiple-precision evaluation of the Airy Ai function

The series expansion at the origin of the Airy function $Ai(x)$ is alternating and hence problematic to evaluate for $x > 0$ due to cancellation. S. Chevillard (APICS team) and M. Mezzarobba showed in [20] how an arbitrary and certified accuracy can be obtained in that case. Based on a method recently proposed by Gawronski, Müller, and Reinhard, they exhibited two functions F and G, both with nonnegative Taylor expansions at the origin, such that $Ai(x) = G(x)/F(x)$. The sums are now well-conditioned, but the Taylor coefficients of G turn out to obey an ill-conditioned three-term recurrence. They then used the classical Miller algorithm to overcome this issue. Finally, they bounded all errors and proposed an implementation which, by allowing an arbitrary and certified accuracy, can be used for example to provide correct rounding in arbitrary precision.

6.2.5. Standardization of interval arithmetic

The IEEE 1788 working group is devoted to the standardization of interval arithmetic. V. Lefèvre and N. Revol are very active in this group. This year is the last year granted by IEEE for the preparation of a draft text of the standard. 2014 will be devoted to a ballot on the whole text, first by the standardization working group and then by a group of experts appointed by IEEE. In 2013, the definition of interval literals, of constructors, and of input and output has been adopted. The work now concentrates on portions of the final text [42].

6.2.6. Parallel product of interval matrices

The problem considered here is the multiplication of two matrices with interval coefficients. Parallel implementations by N. Revol and Ph. Théveny [10] compute results that satisfy the inclusion property, which is the fundamental property of interval arithmetic, and offer good performances: the product of two interval matrices is not slower than 15 times the product of two floating-point matrices.

6.2.7. Numerical reproducibility

What is called *numerical reproducibility* is the problem of getting the same result when the scientific computation is run several times, either on the same machine or on different machines. In [43], the focus is on interval computations using floating-point arithmetic: N. Revol identifies implementation issues that may invalidate the inclusion property, and presents several ways to preserve this inclusion property. This work has also been presented at several conferences [30], [29], [31].

6.3. Floating-point arithmetic

6.3.1. Improved error bounds for complex floating-point arithmetic with a fused-multiply add

Assuming that a fused multiply-add (FMA) instruction is available, C.-P. Jeannerod, N. Louvet, and J.-M. Muller [22] obtained sharp error bounds for various alternatives to Kahan's FMA-based algorithm for 2 x

2 determinants (which they had analyzed in [5]). They showed how to combine such variants with Kahan's original scheme in order to derive componentwise-accurate algorithms for complex floating-point division. Finally, they established sharp or reasonably sharp error bounds for each of these division algorithms.

C.-P. Jeannerod, P. Kornerup (U. of Southern Denmark), N. Louvet, and J.-M. Muller [36] studied the impact of the FMA on the normwise relative accuracy of complex floating-point multiplication. They showed that the classical normwise relative error bound $\sqrt{5}\,u$ (with u the unit roundoff) can be decreased further to $2u$, and that this new constant is best possible for several FMA-based multiplication algorithms.

J.-M. Muller analyzed in [41] another 2 x 2 determinant algorithm, due to Cornea, Harrison, and Tang, and showed that for radix 2 it admits a sharp relative error bound of the form $2u + O(u^2)$.

6.3.2. Improved error bounds for numerical linear algebra

C.-P. Jeannerod and S. M. Rump (Hamburg University of Technology) [7] showed that when evaluating sums of n real numbers in standard floating-point arithmetic, the usual fraction $\gamma_n = nu/(1 - nu)$, which has the form $nu + O(u^2)$ and requires $nu < 1$, can be replaced by nu without any restriction on n. Applications include simpler and more general error bounds for inner products, matrix-vector multiplication, and classical matrix multiplication.

In [45] they extended these results to LU and Cholesky factorizations as well as to triangular linear system solving by showing that the constants γ_n that appear classically in the backward error bounds for such problems can all be replaced by $O(u^2)$-free and unconditional constants nu. To get these new bounds the main ingredient is a general framework for bounding expressions of the form $|\rho - s|$, where s is the exact sum of a floating-point number and $n - 1$ real numbers, and where ρ is a real number approximating the computed sum $\widehat{s}$.

6.3.3. On Ziv's rounding test

F. de Dinechin, J.-M. Muller and S. Torres studied with C. Lauter (Univ. Paris 6) the rounding test introduced by Ziv in its libultim software [4]. This test determines if an approximation to the value $f(x)$ of an elementary function at a given point x suffices to return the floating-point number nearest to $f(x)$. They showed that the same test may be used for efficient implementation of floating-point operations with input and output operands of different formats. That test depends on a "magic constant" e and they also showed how to choose that constant to make the test reliable and efficient. Various cases are considered, depending on the availability of an FMA instruction, and on the range of $f(x)$.

6.3.4. Various issues related to double roundings

Double rounding is a phenomenon that may occur when different floating-point precisions are available on the same system. Although double rounding is, in general, innocuous, it may change the behavior of some useful floating-point algorithms. G. Melquiond (Toccata team), E. Martin-Dorel (then in the Marelle team), and J.-M. Muller analyzed in [9] the potential influence of double rounding on the Fast2Sum and 2Sum algorithms, on some summation algorithms, and Veltkamp's splitting. When performing divisions using Newton-Raphson (or similar) iterations on a processor with a floating-point fused multiply-add instruction, one must sometimes scale the iterations, to avoid over/underflow and/or loss of accuracy. This may lead to double-roundings, resulting in output values that may not be correctly rounded when the quotient falls in the subnormal range. J.-M. Muller showed in [13] how to avoid this problem.

6.3.5. Comparison between binary and decimal floating-point numbers

The IEEE 754-2008 standard for floating-point arithmetic arithmetic specifies binary as well as decimal formats. N. Brisebarre, C. Lauter (Univ. Paris 6), M. Mezzarobba, and J.-M. Muller introduced in [17] an algorithm that allows one to quickly compare a binary64 floating-point number and a decimal64 floating-point number, assuming the "binary encoding" of the decimal formats specified by the IEEE-754 standard is used. It is a two-step algorithm: a first pass, based on the exponents only, makes it possible to quickly eliminate most cases; then, when the first pass does not suffice, a more accurate second pass is required. They provide an implementation of several variants of their algorithm, and compare them.

6.3.6. Conversions between binary and decimal floating-point numbers

Conversion between binary and decimal floating-point representations is ubiquitous. Floating-point radix conversion means converting both the exponent and the mantissa. O. Kupriianova and C. Lauter (Univ. Paris 6) and J.-M. Muller developed in [38] an atomic operation for floating-point radix conversion with simple straight-line algorithm, suitable for hardware design. Exponent conversion is performed with a small multiplication and a lookup table. It yields the correct result without error. Mantissa conversion uses a few multiplications and a small lookup table that is shared amongst all types of conversions. The accuracy changes by adjusting the computing precision.

6.3.7. Table-maker's dilemma

Computing hardest-to-round cases of elementary functions is a key issue when one wants to develop an efficient and reliable implementation of such a function. The algorithms developed until now required a large amount of computation and produced a simple yes/no answer. In [40], G. Hanrot developed together with E. Martin-Dorel (Toccata team), M. Mayero (IUT Villetaneuse, LIPN), and L. Théry (Marelle team) a certificate-based approach of the SLZ algorithm where the execution produces certificates which can then be validated using Coq. This allows one to validate a posteriori the fact that for a given function, a given input precision p and bound p', there is no pair (x, y) of floating-point representable numbers in precision p such that $2^{-e_p(f(x))}|f(x) - y| \leq 2^{-p'}$. This approach has been tested on the exponential function over $[1/2, 1]$, with an input precision of 53 bits and $p' = 300$.

6.4. Hardware and FPGA arithmetic

6.4.1. Reconfiguring arithmetic

With B. Pasca (Altera), F. de Dinechin contributed a book chapter about of the opportunities and challenges of computer arithmetic for reconfigurable/FPGA computing [32]. The main point of this chapter is to look beyond the heritage of processor arithmetic. Using many examples from the FloPoCo project and others, it shows the benefits of merging and fusing standard operators, it introduces an open-ended space of non-standard operators, and illustrates the power of machine-generation of such arithmetic cores.

6.4.2. The bit heap framework for fixed-point arithmetic

N. Brunie, F. de Dinechin, and M. Istoan, with students G. Sergent, K. Illyes, and B. Popa, extended FloPoCo with a versatile framework for manipulating sums of weighted bits [28], [18]. Such bit heaps may be used to express and optimize at the bit level a wide range of operators (from adders and multipliers to polynomials, filters, and other coarse arithmetic cores). A single piece of code can then be used to generate an architecture for any of these operators.

6.4.3. Elementary functions

F. de Dinechin, with P. Echeverria and M. Lopez-Vallejo (U. Madrid) and B. Pasca (Altera), published a hardware architecture for the floating-point pow and powr functions of the IEEE-754-2008 standard [3]. These functions compute x^y, and differ only in the specification of special cases. The implementation, distributed in FloPoCo, is parameterized in exponent and significand size. It combines suitably modified exponential and logarithm units.

F. de Dinechin and M. Istoan, with student G. Sergent, compared several hardware algorithms for the implementation of sine, cosine, and combined sine/cosine [21]: unrolled CORDIC in two variants with several minor improvements, polynomial approximation, and an ad-hoc architecture based on trigonometric identities. A surprising result is that the ad-hoc architecture betters CORDIC even when its multipliers and tables are synthesized as logic.

6.4.4. Contributions to processor architecture

S. Collange (ALF team) and N. Brunie with G. Diamos (Nvidia) suggested improvements for the architecture of general-purpose graphical processing units [11]. As threads take different paths across the control-flow graph, SIMD lockstep execution is partially lost, and must be regained whenever possible in order to maximize the occupancy of SIMD units. Two techniques are described to handle SIMT control divergence and identify reconvergence points. The most advanced one operates in constant space and handles indirect jumps and recursion. In terms of performance, this solution is at least as efficient as state-of-the-art techniques in use in current GPUs.

N. Brunie and F. de Dinechin studied with B. de Dinechin (Kalray) the integration of a tightly coupled reconfigurable accelerator in a massively parallel multiprocessor [27]. For this purpose, they described an architecture exploration framework that produces an architecture along with the relevant compilation software. This framework was demonstrated on AES, SHA2, and a FIR filter.

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 2013 have in particular accounted for the increasingly connected networking environment, as envisioned by the Future Internet, and further focused on one of its major components that is the Internet of Things, which allows connecting the physical with the digital word. In more detail, our research has 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);
- Revisiting service-oriented computing toward the Future Internet, in particular dealing with the composition of highly heterogeneous services while ensuring quality of service (§ 6.3);
- Service oriented middleware for the ultra large scale future mobile Internet of Things (§ 6.4);
- Abstractions for enabling domain experts to easily compose applications on the Internet of Things (§ 6.5);
- Lightweight streaming middleware for the Internet of Things (§ 6.6); and
- Dynamic decision networks for decision-making in self-adaptive systems (§ 6.7).

6.2. Emergent Middleware

Participants: Emil Andriescu, Amel Bennaceur, Valérie Issarny.

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. At the application layer, components may exhibit disparate data types and operations, and may have distinct business logics. At the middleware layer, they may rely on different communication standards, which define disparate data representation formats and induce different architectural constraints. Finally, at the network layer, data may be encapsulated differently according to the network technology in place.

A wide range of approaches have thus been proposed to address the interoperability challenge, as surveyed in [26]. However, solutions that require performing changes to the systems are usually not feasible since the systems to be integrated may be built by third parties (e.g., COTS —Commercial Off-The-Shelf— components or legacy systems); no more appropriate are 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; correct mediators guarantee that the components interact without errors (e.g., deadlocks) and reach their termination successfully. Therefore, formal approaches are used to synthesize mediators automatically.

We posit that interoperability should neither be achieved by defining yet another middleware nor yet another ontology but rather by exploiting existing middleware together with knowledge encoded in existing domain ontologies to synthesize and implement mediators automatically. In [2], we have introduced the notion of *emergent middleware* for realizing mediators, which was initiated as part of the FP7 FET IP CONNECT project. Our work during the year 2013 has more specifically focused on the further elaboration of a comprehensive approach to mediator synthesis, including dealing with interoperability across protocol layers.

Mediator synthesis for emergent middleware: We focus on functionally-compatible components, i.e., components that at some high level of abstraction require and provide compatible functionalities, but are unable to interact successfully due to mismatching interfaces and behaviors. To address these differences without changing the components, mediators that systematically enforce interoperability between functionally-compatible components by mapping their interfaces and coordinating their behaviors are required. Our approach for the automated synthesis of mediators is performed in several steps.

The first step is interface matching, which identifies the semantic correspondence between the actions required by one component and those provided by the other. We incorporate the use of ontology reasoning within constraint solvers, by defining an encoding of the ontology relations using arithmetic operators supported by widespread solvers, and use it to perform interface matching efficiently. For each identified correspondence, we generate an associated matching process that performs the necessary translations between the actions of the two components' interfaces. The second step is the synthesis of correct-by-construction mediators. To do so, we analyze the behaviors of components so as to generate the mediator that combines the matching processes in a way that guarantees that the two components progress and reach their final states without errors. The synthesised mediator is the most general component that ensures freedom of both communication mismatches and deadlock in the composition of the components [15]. The last step consists in making the synthesized mediator concrete by incorporating all the details about the interaction of components. To do so, we compute the translation functions necessary to reconcile the differences in the syntax of the input/output data used by each component and coordinate the different interaction patterns that can be used by middleware solutions.

We refer the interested reader to [7] for a complete description of the approach. Our contribution primarily lies in handling interoperability from the application to the middleware layer in an integrated way. The mediators we synthesize act as: (i) translators by ensuring the meaningful exchange of information between components, (ii) controllers by coordinating the behaviors of the components to ensure the absence of errors in their interaction, and (iii) middleware by enabling the interaction of components across the network so that each component receives the data it expects at the right moment and in the right format.

Automated mediation for cross-layer protocol interoperability: Existing approaches to interoperability are restricted to solving either application heterogeneity when the underlying middlewares are compatible, or solving middleware heterogeneity at each protocol layer separately. In real world scenarios, this does not suffice: application and middleware boundaries are ill-defined and solutions to interoperability must consider them in conjunction. We have been studying the case of cross-layer interoperability where protocol mediation is performed between protocol stacks, rather than between protocol layers separately. Such interoperability approaches are appropriate for systems that rely on complex protocol stacks, where application and middleware layers are tightly coupled.

Systems relying on tightly coupled protocol stacks exchange complex messages that consist of a composition of heterogeneous data formats. To enable interoperation, complex messages from one system must be translated into a different complex format that another system accepts such that the two can interact. While Off-The-Shelf and third party message parsers are widely available for simple message formats (i.e., message formats corresponding to a single protocol layer), complex message formats are typically unique since they are the result of a protocol binding. Protocol binding represents the connection between one protocol and another to create a new communication flow. Some middleware protocols recommend or restrict to certain types of default binding (e.g., HTTP provides an extensive set of rules for binding, such as Content-Encoding and Content-Type). However, real systems are often designed following a custom binding mechanism, restricting the application of automated mediation solutions. This problem occurs primarily because complex message formats cannot be easily interpreted.

Many solutions address this composition issue by introducing Domain Specific Languages that can be used by experts to specify parsers for complex message formats. Yet, whenever messages have a more complicated syntax, providing their DSL descriptions becomes difficult as well. Further, such approaches are not future proof as more protocols are expected to emerge, which will not be accounted for by DSLs that are defined according to known message formats. An alternative is to generate parsers based on the composition of third-party parsers that are usually included with protocol implementations. However, third-party parsers cannot be used unless the protocol binding rules are identified by an expert, further allowing to implement the bridge between one parser's output data and the other parser's input data. To this end, we designed an approach for generating composed parsers that can process complex messages, accompanied by a formal mechanism for defining complex message formats based on existing data formats. Our approach relies on user-provided parser composition rules, which reflect the binding requirements of complex message formats.

We posit that our method is more efficient than implementing complex parsers, defining them using DSLs, or directly implementing the binding of protocols. Furthermore, with this solution, we support the automated synthesis of mediators at the application layer using the mapping-based approach discussed above, by automatically generating an abstract representation of the application data exchanged by the interoperating components.

6.3. Service-oriented Computing in the Future Internet

Participants: Georgios Bouloukakis, Nikolaos Georgantas, Valérie Issarny, Ajay Kattepur.

With an increasing number of services and devices interacting in a decentralized manner, *choreographies* represent a scalable framework for the Future Internet. The service oriented architecture inherent to choreographies allows abstracting multiple devices as components, that interact through middleware connectors via standard protocols. However, the heterogeneous nature of devices leads to choreographies that not only include conventional services, but also sensor-actuator networks, databases and service feeds. We reason about their behavior through abstract middleware interaction paradigms, such as client-service (CS), publish-subscribe (PS) and tuple space (TS), made interoperable through the *eXtensible Service Bus* (XSB) connector.

Extensible Service Bus for the Future Internet: XSB is an abstract service bus that deals effectively with the cross-integration of heterogeneous interaction paradigms [17]. Inside the XSB, the CS, PS and TS paradigms are modeled as abstract base connectors. Their *space coupling* semantics are represented with programming interfaces used by applications (APIs) and corresponding application interface description languages (IDLs). Their behavioral semantics are formally specified in terms of LTS (Labeled Transition Systems). We formally verify the correctness of these behavioral specifications with respect to *time coupling* and *concurrency* properties expressed in LTL temporal logic. This allows stating the correctness of the connector models with respect to the semantics that they must have. This further enables identifying the behavioral semantics of the XSB connector derived from the interconnection of base connectors. More specifically, in order to identify the time coupling and concurrency semantics of XSB and construct a converter among the base connectors, we build upon the formal method of *protocol conversion via projections*[10]. According to this method, conversion between two different protocols is possible if both protocols can be projected (where projection is an abstraction defined as a set of transformations on the protocol LTS) to a *functionally sufficient* common *image protocol*. Then, the end-to-end protocol of the interconnection of the two protocols is this image protocol.

We have implemented our XSB solution into an extensible development and execution platform for application and middleware designers. Using this platform, they can easily develop composite applications: they only need to build descriptions for the constituent services and directives for data mapping among them. Our platform then deals with reconciling among the heterogeneous interaction paradigms and protocols of the services by employing *binding components (BCs)* that adapt between the native middleware of the services and the XSB bus protocol. The XSB itself is implemented on top of an existing ESB substrate. Support for new middleware platforms, new ESB substrates, or even new interaction paradigms can be incorporated in a facilitated way thanks to the provided XSB architectural framework.

[10] Lam, S.S.: Protocol Conversion. IEEE Trans. Softw. Eng. 14(3) (1988) 353–362.

QoS composition and analysis of heterogeneous choreographies: Leveraging on the functional interoperability across interaction paradigms offered by the XSB, we study the Quality of Service (QoS) performance of choreographies [21]. QoS dependency plays an important role in the service oriented system lifecycle, including discovery, runtime selection, replacement and contractual guarantees. Consequently, QoS composition among choreographed devices should tackle multi-dimensional probabilistic metrics combined with message passing constraints imposed at design-time. We make use of an algebraic QoS composition model that is applied at the interaction paradigm level to study the composition of QoS metrics, and the subsequent tradeoffs. While traditional QoS composition analysis has been done purely at the application level, analyzing the effect of middleware interactions allows us to study CS, PS and TS based device compositions. This produces interesting insights such as selection of a particular system and its middleware during design-time, or end-to-end QoS expectation/guarantees during runtime. Our formulation also allows for runtime reconfiguration, in order to optimally produce design time QoS expectations. Such flexible reconfiguration policies are crucial in the case of large scale choreographies with high variability in runtime performance of participating devices.

Further, we study the effect of time/space coupling on the latency of successful transactions across the XSB connector [20]. XSB models the message passing among peers through generic `post` and `get` operations, that represent peer behavior with both tight (CS) and loose (PS/TS) time/space coupling. The heterogeneous *lease* and *timeout* behaviors of these operations severely affect latency and success rates of messages passed either synchronously or through callbacks. By precisely studying the timing thresholds using timed automata models, we verify conditions for accurate message transactions with XSB connectors. This offers choreography designers the ability to set these timing thresholds (bottom-up) or select a particular interaction paradigm (top-down) for runtime enaction.

6.4. Service-oriented Middleware for the Mobile Internet of Things

Participants: Sara Hachem, Valérie Issarny, Georgios Mathioudakis, Animesh Pathak.

The Internet of Things (IoT) is characterized by an increasing number of Things embedding sensing, actuating, processing, and communication capacities. A considerable portion of those Things will be *mobile* Things, which come with several advantages yet lead to unprecedented challenges. The most critical challenges, that are directly inherited from, yet amplify, today's Internet issues, lie in handling i) the large scale of users and mobile Things, ii) providing interoperability across the heterogeneous Things, and iii) overcoming the unknown dynamic nature of the environment, due to the mobility of an ultra-large number of Things.

Service-Oriented Architecture (SOA) provides solid basis to address the above challenges as it allows the functionalities of sensors/actuators embedded in Things to be provided as services, while ensuring loose-coupling between those services and their hosts, thus abstracting their heterogeneous nature. In spite of its benefits, SOA has not been designed to address the ultra-large scale of the mobile IoT. Consequently, an alternative is provided within a novel Thing-based Service-Oriented Architecture, that revisits SOA interactions and functionalities, service discovery and composition in particular. The novel architecture is concretized within MobIoT, a middleware solution that is specifically designed to manage and control the ultra-large number of mobile Things in partaking in IoT-related tasks.

In accordance with SOA, MobIoT comprises *Discovery*, *Composition & Estimation*, and *Access* components, yet modifies their internal functionalities. In more detail, the Discovery component enables Thing-based service registration (for Things to advertise hosted services) and look-up (for Things to retrieve remote services of interest). In order to handle the ultra large number of mobile Things and their services in the IoT, the component revisits the Service-Oriented discovery and introduces *probabilistic discovery* to provide, not *all*, but only a sufficient *subset of services that can best approximate* the result that is being sought after [18], [11]. Furthermore, the Composition & Estimation component (C&E) provides automatic composition of Thing-based services. This capacity is of interest in the case where no service can perform a required measurement/action task directly (based on its atomic functionalities). Thing-based service composition executes in three phases: i) *expansion*, where composition specifications are automatically identified; ii) *mapping*, where actual service instances (running services) are selected based on their functionalities and the physical attributes of their hosts; and iii) *execution*, where the services are accessed and the composition specifications are executed.

Thing-based service composition revisits Service-Oriented composition by executing seamlessly with no involvement from developers or end users. Last but not least, the Access component provides an easy to use interface for developers to sample sensors/actuators while abstracting sensor/actuator hardware specifications. Additionally, it revisits Service-Oriented access by executing access to services transparently and wrapping access functionalities internally. Thus, it alleviates that burden from users, initially in charge of this task. The Access component supports access to remote services and to locally hosted services.

6.5. Composing Applications in the Internet of Things

Participants: Aness Bajia, Pankesh Patel, Animesh Pathak, Françoise Sailhan.

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. 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. The goal of our research is then to propose a suitable application development framework, where our work this year covered the two 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 [22]. 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 evaluation based on two realistic scenarios shows that the use of our techniques/framework succeeds in improving productivity in the IoT application development process. More details of our approach can be found in [8].

Integrating support for non-functional requirements while programming IoT applications: Given that devices and networks constituting the IoT are prone to failure and consequent loss of performance, it is natural that IoT applications are expected to encounter and tolerate several classes of faults - something that still largely remains within the purview of low-level-protocol designers. As part of our work on the MURPHY project (§ 7.1.1.1), we are addressing this issue by proposing: i) a set of abstractions that can be used during macroprogramming to express fault tolerance requirements, and ii) a runtime system that employs adaptive fault tolerance (AFT) to provide fault tolerance to the sensing application. Complementary to this, we have proposed task mapping algorithms to satisfy those requirements through a constraint programming approach [19]. Through evaluations on realistic application task graphs, we show that our constraint programming model can effectively capture the end-to-end requirements and efficiently solves the combinatorial problem introduced.

We have continually incorporating our research results in the above areas into *Srijan* (§ 5.6), 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. Lightweight Streaming Middleware for the Internet of Things

Participants: Benjamin Billet, Valérie Issarny.

The Internet of Things (IoT) is a promising concept toward pervasive computing as it may radically change the way people interact with the physical world. One of the challenges raised by the IoT is the in-network continuous processing of data streams presented by Things, which must be investigated urgently because it affects the future data models of the IoT. This cross-cutting concern has been previously studied in the context of Wireless Sensor and Actuator Networks (WSAN) given the focus on the acquisition and in-network processing of sensed data. However, proposed solutions feature heterogeneous technologies that are difficult to integrate and complex to use, which represents a hurdle to their wide deployment. In addition, new types of smart sensors are emerging due to technological advances (e.g., Oracle SunSpot), enabling the implementation of complex processing tasks directly into the network, without using proxies or sending every data to the cloud. There is thus a need for a distributed middleware solution for data stream management that leverages existing WSAN work, while integrating it with today's Web technologies in order to improve the flexibility and the interoperability of the future IoT. Toward that goal, we have been developing Dioptase, a Data Stream Management System for the IoT, which aims to integrate the Things and their streams into today's Web by presenting sensors and actuators as services. The middleware specifically provides a way to describe complex fully-distributed stream-based mashups and to deploy them dynamically, at any time, as task graphs, over available Things of the network, including resource-constrained ones. To this end, Dioptase enables task graphs to be composed of Thing-specific tasks (directly implemented on the Thing) and dynamic tasks that communicates using data streams. Dynamic tasks are then described in a lightweight DSL, which is directly interpreted by the middleware and provides specific primitives to manipulate data streams.

As part of the design of Dioptase, we have been investigating dedicated task mapping. Task mapping, which basically consists of mapping a set of tasks onto a set of nodes, is a well-known problem in distributed computing research. However, as a particular case of distributed systems, the Internet of Things (IoT) poses a set of renewed challenges, because of its scale, heterogeneity and properties traditionally associated with WSAN, shared sensing, continous processing of data streams and real time computing. To handle IoT features, we present a formalization of the task mapping problem that captures the varying consumption of resources and various constraints (location, capabilities, QoS) in order to compute a mapping that guarantees the lifetime of the concurrent tasks inside the network and the fair allocation of tasks among the nodes (load balancing). It results in a binary programming problem for which we provide an efficient heuristic that allows its resolution in polynomial time. Our experiments show that our heuristic: (i) gives solutions that are close to optimal and (ii) can be implemented on reasonably powerful Things and performed directly within the network, without requiring any centralized infrastructure.

6.7. Dynamic Decision Networks for Self-Adaptive Systems

Participants: Amel Belaggoun, Nelly Bencomo, Valérie Issarny, Peter Sawyer.

Different modeling techniques have been used to model requirements and decision-making of self-adaptive systems [25]. Important successful techniques based on goal models have been prolific in supporting decision-making according to partial and total fulfillment of functional (goals) and non-functional requirements (softgoals). The final decision about what strategy to use is based on a utility function that takes into account the weighted sum of the different effects of the non-functional requirements. Such solutions have been used both at design and run time including our own solutions using runtime goal models. Different modeling techniques have been used to model requirements and decision-making of self-adaptive systems [25]. Important successful techniques based on goal models have been prolific in supporting decision-making according to partial and total fulfillment of functional (goals) and non-functional requirements (softgoals). The final decision about what strategy to use is based on a utility function that takes into account the weighted sum of the different effects of the non-functional requirements. Such solutions have been used both at design- and run-time including our own solutions using runtime goal models.

We have enriched the decision-making supported by goal models with the use of Bayesian Dynamic Decision Networks (DDNs) [12]. Our novel approach supports reasoning about partial satisfaction of soft-goals using probabilities and uses machine learning. When using DDNs, we introduce new ways to tackle uncertainty based on probabilities that can be updated based on runtime evidence. We have reported the results of the application of the approach on two different cases, one of them being the case of dynamic reconfiguration of a remote data mirroring network that must spread data among servers while minimizing costs and loss of data. Our early results suggest the decision-making process of self-adaptive systems can be improved by using DDNs.

This work has been developed under the umbrella of the Marie Curie Project Requirements@run.time (§ 7.2.1.4). The main results achieved during the year 2013 are:

- A Bayesian-based technique to support the decision making of self-adaptive systems [14]. DDN-based approaches adopt probabilistic methods (i.e., Bayesian methods) and decision theory to assess the consequences of uncertainty. Using the approach, suitable choices to satisfice functional requirements of the system are identified from a range of alternative decisions and their expected utilities. Satisficement of NFRs is modeled using conditional probabilities given the design decisions. Preferences over decisions are modeled using weights associated with pairs of design alternatives and NFRs, and used when computing the expected utilities of the architectural design alternatives. The decision taken by the DDN is that with the highest expected utility. The approach offers the benefits of machine learning.
- A formal Bayesian definition of surprise as the basis for quantitative analysis to measure degrees of uncertainty and deviation of self-adaptive systems from normal behavior [13]. Specifically, a Bayesian surprise quantifies how new evidence affects assumptions of the world (properties in the models). A “surprising" event may provoke a large divergence between the beliefs distributions prior and posterior to that event. As such and depending on how big or small this divergence is, the running system may decide to either: (i) dynamically adapt accordingly, or (ii) temporarily avoid any action of adaptation and flag up the fact that a potential abnormal situation has been found. While doing (ii) we are offering a specific implementation of the RELAX language previously developed by Bencomo and her co-authors.

ASAP Project-Team

6. New Results

6.1. Models and abstractions for distributed systems

6.1.1. Randomized loose renaming in O(loglog n) time

Participant: George Giakkoupis.

Renaming is a classic distributed coordination task in which a set of processes must pick distinct identifiers from a small namespace. In [24], we consider the time complexity of this problem when the namespace is linear in the number of participants, a variant known as loose renaming. We give a non-adaptive algorithm with $O(\log\log n)$ (individual) step complexity, where n is a known upper bound on contention, and an adaptive algorithm with step complexity $O((\log\log k)^2)$, where k is the actual contention in the execution. We also present a variant of the adaptive algorithm which requires $O(k\log\log k)$*total* process steps. All upper bounds hold with high probability against a strong adaptive adversary. We complement the algorithms with an $\Omega(\log\log n)$ expected time lower bound on the complexity of randomized renaming using test-and-set operations and linear space. The result is based on a new coupling technique, and is the first to apply to non-adaptive randomized renaming. Since our algorithms use $O(n)$ test-and-set objects, our results provide matching bounds on the cost of loose renaming in this setting.

This work was done in collaboration with Dan Alistarh, James Aspnes, and Philipp Woelfel.

6.1.2. An O(sqrt n) space bound for obstruction-free leader election

Participant: George Giakkoupis.

In [32] we present a deterministic obstruction-free implementation of leader election from $O(\sqrt{n})$ atomic $O(\log n)$-bit registers in the standard asynchronous shared memory system with n processes. We provide also a technique to transform any deterministic obstruction-free algorithm, in which any process can finish if it runs for b steps without interference, into a randomized wait-free algorithm for the oblivious adversary, in which the expected step complexity is polynomial in n and b. This transformation allows us to combine our obstruction-free algorithm with the leader election algorithm by Giakkoupis and Woelfel (2012), to obtain a fast randomized leader election (and thus test-and-set) implementation from $O(\sqrt{n})O(\log n)$-bit registers, that has expected step complexity $O(\log^* n)$ against the oblivious adversary. Our algorithm provides the first sub-linear space upper bound for obstruction-free leader election. A lower bound of $\Omega(\log n)$ has been known since 1989 (Styer and Peterson, 1989). Our research is also motivated by the long-standing open problem whether there is an obstruction-free consensus algorithm which uses fewer than n registers.

This work was done in collaboration with Maryam Helmi, Lisa Higham, and Philipp Woelfel.

6.1.3. Broadcast in recurrent dynamic systems

Participants: Michel Raynal, Julien Stainer.

This work [50] proposes a simple broadcast algorithm suited to dynamic systems where links can repeatedly appear and disappear. The algorithm is proved correct and a simple improvement is introduced, that reduces the number and the size of control messages. As it extends in a simple way a classical network traversal algorithm (due to A. Segall, 1983) to the dynamic context, the proposed algorithm has also pedagogical flavor.

This work has been done in collaboration with Jiannong Cao and Weigang Wu.

6.1.4. Computing in the presence of concurrent solo executions

Participants: Michel Raynal, Julien Stainer.

In a wait-free model any number of processes may crash. A process runs solo when it computes its local output without receiving any information from other processes, either because they crashed or they are too slow. While in wait-free shared-memory models at most one process may run solo in an execution, any number of processes may have to run solo in an asynchronous wait-free message-passing model. This work [47] is on the computability power of models in which several processes may concurrently run solo. We introduced a family of round-based wait-free models, called the d-solo models, $1 \le d \le n$, where up to d processes may run solo. Then we gave a characterization of the colorless tasks that can be solved in each d-solo model. We also introduced the (d, ϵ)-solo approximate agreement task, which generalizes ϵ-approximate agreement, and proves that (d, ϵ)-solo approximate agreement can be solved in the d-solo model, but cannot be solved in the $(d + 1)$-solo model. We also studied the relation linking d-set agreement and (d, ϵ)-solo approximate agreement in asynchronous wait-free message-passing systems. These results establish for the first time a hierarchy of wait-free models that, while weaker than the basic read/write model, are nevertheless strong enough to solve non-trivial tasks.

This work was done in collaboration with Maurice Herlihy and Sergio Rajsbaum.

6.1.5. Relating message-adversaries and failure detectors

Participants: Michel Raynal, Julien Stainer.

A message adversary is a daemon that suppresses messages in round-based message-passing synchronous systems in which no process crashes. A property imposed on a message adversary defines a subset of messages that cannot be eliminated by the adversary. It has recently been shown that when a message adversary is constrained by a property denoted TOUR (for tournament), the corresponding synchronous system and the asynchronous crash-prone read/write system have the same computability power for task solvability. In this work [39] we introduced new message adversary properties (denoted SOURCE and QUORUM), and shown that the synchronous round-based systems whose adversaries are constrained by these properties are characterizations of classical asynchronous crash-prone systems (1) in which processes communicate through atomic read/write registers or point-to-point message-passing, and (2) enriched with failure detectors such as Ω and Σ. Hence these properties characterize maximal adversaries, in the sense that they define strongest message adversaries equating classical asynchronous crash-prone systems. They consequently provide strong relations linking round-based synchrony weakened by message adversaries with asynchrony restricted with failure detectors. This not only enriches our understanding of the synchrony/asynchrony duality, but also allows for the establishment of a meaningful hierarchy of property-constrained message adversaries.

6.1.6. A hierarchy of agreement problems from simultaneous consensus to set agreement

Participants: Michel Raynal, Julien Stainer.

In this work [38] we investigated the relation linking the s-simultaneous consensus problem and the k-set agreement problem in wait-free message-passing systems. To this end, we defined the (s, k)-SSA problem which captures jointly both problems: each process proposes a value, executes s simultaneous instances of a k-set agreement algorithm, and has to decide a value so that no more than sk different values are decided. We also introduced a new failure detector class denoted $Z_{s,k}$, which is made up of two components, one focused on the "shared memory object" that allows the processes to cooperate, and the other focused on the liveness of (s, k)-SSA algorithms. A novelty of this failure detector lies in the fact that the definition of its two components are intimately related. We designed a $Z_{s,k}$-based algorithm that solves the (s, k)-SSA problem, and shown that the "shared memory"-oriented part of $Z_{s,k}$ is necessary to solve the (s, k)-SSA problem (this generalizes and refines a previous result that showed that the generalized quorum failure detector Σ_k is necessary to solve k-set agreement). We finally, investigated the structure of the family of (s, k)-SSA problems and introduced generalized (asymmetric) simultaneous set agreement problems in which the parameter k can differ in each underlying k-set agreement instance. Among other points, it shows that, for $s, k > 1$, (a) the $(sk, 1)$-SSA problem is strictly stronger that the (s, k)-SSA problem which is itself strictly stronger than the $(1, ks)$-SSA problem, and (b) there are pairs (s_1, k_1) and (s_2, k_2) such that $s_1k_1 = s_2k_2$ and (s_1, k_1)-SSA and (s_2, k_2)-SSA are incomparable.

6.2. Large-scale and user-centric distributed systems

6.2.1. FreeRec: An anonymous and distributed personalization architecture

Participants: Antoine Boutet, Davide Frey, Arnaud Jégou, Anne-Marie Kermarrec, Heverson Borba Ribeiro.

FreeRec is an anonymous decentralized peer-to-peer architecture designed to bring personalization while protecting the privacy of its users [17], [30], [44]. FreeRec's decentralized approach makes it independent of any entity wishing to collect personal data about users. At the same time, its onion-routing-like gossip-based overlay protocols effectively hide the association between users and their interest profiles without affecting the quality of personalization. The core of FreeRec consists of three layers of overlay protocols: the bottom layer, rps, consists of a standard random peer sampling protocol ensuring connectivity; the middle layer, PRPS, introduces anonymity by hiding users behind anonymous proxy chains, providing mutual anonymity; finally, the top clustering layer identifies for each anonymous user, a set of anonymous nearest neighbors. We demonstrate the effectiveness of FreeRec by building a decentralized and anonymous content dissemination system. Our evaluation by simulation, our PlanetLab experiments, and our probabilistic analysis show that FreeRec effectively decouples users from their profiles without hampering the quality of personalized content delivery.

6.2.2. HyRec: A hybrid recommender system

Participants: Antoine Boutet, Davide Frey, Anne-Marie Kermarrec.

The ever-growing amount of data available on the Internet calls for personalization. Yet, the most effective personalization schemes, such as those based on collaborative filtering (CF), are notoriously resource greedy. HyRec is an online cost-effective scalable system for CF personalization. HyRec relies on a hybrid architecture, offloading CPU-intensive recommendation tasks to front-end client browsers, while retaining storage and orchestration tasks within back-end servers. HyRec has been fully implemented and extensively evaluated on several workloads from MovieLens and Digg. We convey the ability of HyRec to significantly reduce the operation costs of the content provider by up to 70% and drastically improve the scalability by up to 500%, with respect to a centralized (or cloud-based recommender approach), while preserving the quality of the personalization. We also show that HyRec is virtually transparent to the users and induces only 3% of the bandwidth consumption of a P2P solution.

6.2.3. Social market

Participants: Davide Frey, Arnaud Jégou, Anne-Marie Kermarrec, Michel Raynal, Julien Stainer.

The ability to identify people that share one's own interests is one of the most interesting promises of the Web 2.0 driving user-centric applications such as recommendation systems or collaborative marketplaces. To be truly useful, however, information about other users also needs to be associated with some notion of trust. Consider a user wishing to sell a concert ticket. Not only must she find someone who is interested in the concert, but she must also make sure she can trust this person to pay for it. Social Market (SM) solves this problem by allowing users to identify and build connections to other users that can provide interesting goods or information and that are also reachable through a trusted path on an explicit social network like Facebook. This year, we extended the contributions presented in 2011, by introducing two novel distributed protocols that combine interest-based connections between users with explicit links obtained from social networks ala Facebook. Both protocols build trusted multi-hop paths between users in an explicit social network supporting the creation of semantic overlays backed up by social trust. The first protocol, TAPS2, extends our previous work on *TAPS* (Trust-Aware Peer Sampling), by improving the ability to locate trusted nodes. Yet, it remains vulnerable to attackers wishing to learn about trust values between arbitrary pairs of users. The second protocol, PTAPS (*Private TAPS*), improves TAPS2 with provable privacy guarantees by preventing users from revealing their friendship links to users that are more than two hops away in the social network. In addition to proving this privacy property, we evaluate the performance of our protocols through event-based simulations, showing significant improvements over the state of the art. In addition to our previous publication on this topic, our recent work led to a paper that appeared in TCS [20].

6.2.4. Privacy-preserving P2P collaborative filtering

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

The huge amount of information available at any time in our connected society calls for a mechanism to filter it efficiently. Recommendation systems provide such a mechanism by personalizing the information displayed for each user. However, the collection of personal information by recommendation systems threatens the privacy of users. We address the two needs for recommendation and privacy through a peer-to-peer user-based collaborative filtering system. Recommendation is done ala GOSSPLE by building an overlay network which connects users with similar interests via clustering and random peer sampling. This overlay network is then used to make recommendations based on what similar users liked. Users' privacy is protected in two ways. Users are protected from a Big Brother adversary by the peer-to-peer design of the system in which profiles are stored only by their owners. Users are protected from other malicious users who would try to learn the content of their profiles by our landmark-based cosine similarity measure. It indirectly computes the similarity of two users by comparing their respective similarities with a set of randomly generated profiles, called landmarks. Thus, users can compute their similarity without revealing their profile, contrarily to the regular cosine similarity when used in a peer-to-peer system.

6.2.5. Gossip protocols for renaming and sorting

Participants: George Giakkoupis, Anne-Marie Kermarrec.

In [33] we devise efficient gossip-based protocols for some fundamental distributed tasks. The protocols assume an n-node network supporting point-to-point communication, and in every round, each node exchanges information of size $O(\log n)$ bits with (at most) one other node. We first consider the *renaming* problem, that is, to assign distinct IDs from a small ID space to all nodes of the network. We propose a renaming protocol that divides the ID space among nodes using a natural push or pull approach, achieving logarithmic round complexity with ID space $\{1, \cdots, (1+\epsilon)n\}$, for any fixed $\epsilon > 0$. A variant of this protocol solves the *tight* renaming problem, where each node obtains a unique ID in $\{1, \cdots, n\}$, in $O(\log^2 n)$ rounds. Next we study the following *sorting* problem. Nodes have consecutive IDs 1 up to n, and they receive numerical values as inputs. They then have to exchange those inputs so that in the end the input of rank k is located at the node with ID k. Jelasity and Kermarrec (2006) suggested a simple and natural protocol, where nodes exchange values with peers chosen uniformly at random, but it is not hard to see that this protocol requires $\Omega(n)$ rounds. We prove that the same protocol works in $O(\log^2 n)$ rounds if peers are chosen according to a non-uniform power law distribution.

This work has been done in collaboration with Philipp Woelfel.

6.2.6. Adaptive streaming

Participants: Ali Gouta, Anne-Marie Kermarrec.

HTTP Adaptive Streaming (HAS) is gradually being adopted by Over The Top (OTT) content providers. In HAS, a wide range of video bitrates of the same video content are made available over the internet so that clients' players pick the video bitrate that best fit their bandwidth. Yet, this affects the performance of some major components of the video delivery chain, namely CDNs or transparent caches since several versions of the same content compete to be cached. We investigated the benefits of a Cache Friendly HAS system (CF-DASH), which aims to improve the caching efficiency in mobile networks and to sustain the quality of experience of mobile clients. We presented a set of observations we made on large number of clients requesting HAS contents [34], [35]. Then, we evaluated CF-dash based on trace-driven simulations and testbed experiments. Our validation results are promising. Simulations on real HAS traffic show that we achieve a significant gain in hit-ratio that ranges from 15% up to 50%.

Work was done in collaboration with Yannick Le Louedec, Zied Aouini and Diallo Mamadou.

6.2.7. DynaSoRe: Efficient in-memory store for social applications

Participant: Arnaud Jégou.

Social network applications are inherently interactive, creating a requirement for processing user requests fast. To enable fast responses to user requests, social network applications typically rely on large banks of cache servers to hold and serve most of their content from the cache. The objective of this work is to build a memory cache system for social network applications that optimizes data locality while placing user views across the system. We call this system DynaSoRe (Dynamic Social stoRe). DynaSoRe storage servers monitor access traffic and bring data frequently accessed together closer in the system to reduce the processing load across cache servers and network devices. Our simulation results considering realistic data center topologies show that DynaSoRe is able to adapt to traffic changes, increase data locality, and balance the load across the system. The traffic handled by the top tier of the network connecting servers drops by 94% compared to a static assignment of views to cache servers while requiring only 30% additional memory capacity compared to the whole volume of cached data.

This work was conducted in collaboration with Xiao Bai, Flavio Junqueira, and Vincent Leroy. The product of this collaboration led to the publication of a paper at the Middleware 2013 conference [26].

6.2.8. Adaptive metrics on distributed recommendation systems

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

Current distributed recommendation systems are metric based. This means that recommendation quality depends on a single user comparison function. This is a simple solution that cannot cover the particularities of each system. Classically computing intensive data-mining methods have been used in the field of recommendation. However, they are not proper in distributed scenarios due to the lack of a global vision and the existing restrictions in terms of computing power. In this project, we study how to provide and model ad-hoc similarity metrics that can be automatically adapted to a different number of scenarios. We study our solution from two different points of view: recommendation and performance. In the first, we evaluate the capacity of data mining technics to give users relevant recommendations. Second, by exploring the performance of different approaches in order to obtain relevant recommendations we plan to study the trade-off between relevant recommendations and computational cost.

6.2.9. Cliff-Edge Consensus: Agreeing on the precipice

Participants: Michel Raynal, François Taïani.

In this project, we worked on a new form of consensus that allows nodes to agree locally on the extent of crashed regions in networks of arbitrary size. One key property of our algorithm is that it shows local complexity, i.e. its cost is independent of the size of the complete system, and only depends on the shape and extent of the crashed region to be agreed upon. In [40], we motivate the need for such an algorithm, formally define this new consensus problem, propose a fault-tolerant solution, and prove its correctness.

This work was done in collaboration with Geoff Coulson and Barry Porter.

6.2.10. Clustered network coding

Participants: Fabien André, Anne-Marie Kermarrec, Konstantinos Kloudas, Alexandre Van Kempen.

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

In this work, we propose a novel erasure code system, tailored for networked archival systems. The efficiency of our approach relies on a combination of the use of random codes coupled with a clever yet simple clustered placement strategy. Our repair protocol leverages network coding techniques to reduce by 50% the amount of data transferred during maintenance, as several cluster files are repaired simultaneously. We demonstrate both through an analysis and extensive experimental study conducted on a public testbed that our approach dramatically decreases both the bandwidth overhead during the maintenance process and the time to repair data lost upon failure.

This has been done in collaboration with Erwan le Merrer, Nicolas, Le Scouarnec and Gilles Straub.

ASCLEPIOS Project-Team

5. New Results

5.1. Medical Image Analysis

5.1.1. Segmentation of cardiac images from magnetic resonance

Participants: Jan Margeta [Correspondant], Kristin Mcleod, Antonio Criminisi [MSRC], Nicholas Ayache.

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

Cardiac imaging, Magnetic resonance, Image segmentation, Maching learning

- We contributed our previous method to build left ventricle myocardium segmentation consensus based on the STAPLE algorithm [26]
- We enhanced our segmentation method with extra features based on the distance transform and image vesselness measures in order to segment left atria (see Fig. 1) from 3d MRI images [49]. We participated with this method in the left atrium segmentation challenge at MICCAI 2013.

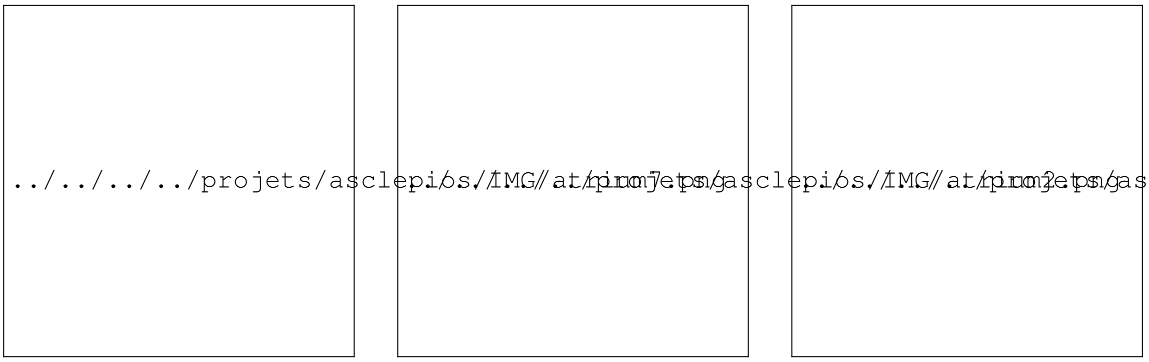


Figure 1. Segmented atria meshes from the validation dataset.

5.1.2. Brain tumor image processing and modeling

Participants: Bjoern Menze [Correspondant], Hervé Delingette, Nicholas Ayache, Nicolas Cordier, Erin Stretton, Jan Unkelbach.

We developed a new non-parametric lesion growth model for the analysis of longitudinal image sequences [59], evaluated the parametric tumor growth model of Konukoglu on longitudinal data, focusing on the relevance of DTI [40], [57], and addressed the question of how to detect tumor growth from longitudinal sequences of patients treated with angiogenesis inhibitors using registration techniques [47]. We also completed work for the 2012 MICCAI Challenge on Brain Tumor Image Segmentation (MICCAI-BRATS 2012) [79], where we also tested some of our own brain tumor image segmentation models based on random forests [42] and patch regression [38], we also participated in MICCAI-BRATS 2013 in Nagoya, Japan [67].

5.1.3. Further developing the random forest framework for medical computer vision tasks

Participants: Bjoern Menze [Correspondant], Matthias Schneider, Ezequiel Geremia, Rene Donner, Georg Langs, Gabor Szekely.

Methodological contributions include the further development of the random forest framework. We introduced the “spatially adaptive” random forest (SARF) classifier [42], and evaluated Hough regression forests for interest point detection in whole body CT image analysis, as well as for vessel detection and tracking [54]. We also evaluated alternative patch-based methods for whole body image registration [41]. As a related community effort, we organized the MICCAI-MCV workshop, also in conjunction with the MICCAI conference in Nagoya, Japan [65].

5.1.4. Statistical Analysis of Diffusion Tensor Images of the Brain

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

Diffusion Tensor Imaging of the Brain, Tractography, Super-resolution, Statistical analysis

Diffusion tensor imaging (DTI) is gaining interest as a clinical tool for studying a number of brain diseases pertaining to white matter tracts and also as an aid in neuro-surgical planning. Unfortunately, in a clinical environment, diffusion imaging is hampered by the long acquisition times, low signal to noise ratio and a prominent partial volume effect due to thick slices. We are developing a framework for increasing the resolution of the low-resolution clinical CTI images. The method uses a maximum likelihood strategy to account for the noise and an anisotropic regularization prior to promote smoothness in homogeneous areas while respecting edges. The technique is called Higher Resolution Tensor Estimation and it uses a single clinical acquisition to produce high resolution tensor images. We aim to replace resampling techniques used for tensor normalization in population based studies, with the present method. The method itself along with quantitative results on tractography 2 were presented in MICCAI 2013 [45].

5.1.5. 3D/2D coronary arteries registration

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

This work is done in collaboration with GE-Healthcare (Buc).

3D/2D registration; computed tomography angiography; CTA; X-ray fluoroscopy; coronary arteries

Endovascular treatment of coronary arteries involves catheter navigation through patient vasculature. Projective angiography guidance is limited in the case of chronic total occlusion where the occluded vessel can not be seen. Integrating standard preoperative CT angiography information with live fluoroscopic images addresses this limitation but requires alignment of both modalities.

We published the Iterative Closest Curve (ICC) algorithm [36] in the MICCAI 2013 conference :

- The ICC-algorithm mimics the ICP-algorithm [6], curves being considered instead of points.
- Contrary to closest point pairing, the resulting pairings assure a topological and geometrical coherence since a curve is paired to another one (cf Figure 3).
- The developed method can deal with differences in both datasets by considering outlier rejection at the level of curve and the level of point.

5.1.6. Automatic Registration of Endoscopic Images

Participants: Anant Vemuri [Correspondant], Stéphane Nicolau [IHU Strasbourg], Luc Soler [IHU Strasbourg], Nicholas Ayache.

This work is performed in collaboration with IHU Strasbourg.

Image registration; Endoscopic imaging; Biopsy Relocalization

[6] P.J. Besl and N.D. McKay. A method for registration of 3-D shapes

../../../../projets/asclepios/IMG/DTI.jpg

Figure 2. Middle column shows a comparatively dense fiber bundle in the fornix region for the Higher Resolution Tensor Estimation method (superior-inferior view) compared to tensor resampling (left column). Right column shows a quantitative comparison of fiber lengths.

Figure 3. Left : Pairing obtained with the ICP algorithm, Middle : Pairing obtained with the ICC algorithm, Right : ICC (green) and ICP (red) registration results.

The screening of cancer lesions in the oesophagus involves obtaining biopsies at different regions along the oesophagus. The localization and tracking of these biopsy sites inter-operatively poses a significant challenge for providing targeted treatments.

Our work [61] introduces a novel framework for accurate re-positioning of the endoscope at previously targeted sites:

- it includes an electromagnetic tracking system in the loop and provides a framework for utilizing it for re-localization inter-operatively.
- We have shown on three in-vivo porcine interventions that our system can provide accurate guidance information, which was qualitatively evaluated by five experts.

5.2. Biological Image Analysis

5.2.1. Pre-clinical molecular imaging: motionless 3D image reconstruction in micro-SPECT

Participants: Marine Breuilly [Correspondant, Inria], Grégoire Malandain [Inria], Nicholas Ayache [Inria], Jacques Darcourt [UNS-CAL], Philippe Franken [UNS-CAL], Thierry Pourcher [CEA].

This work is jointly conducted with the Transporter in Imagery and Oncologic Radiotherapy team (TIRO, CEA-CAL-UNSA) located in Nice.

SPECT/CT, small animal, respiratory motion, respiratory gating, 4D images, stomach, 99mTc-pertechnetate biodistribution, compartmental analysis

This work has been conducted on SPECT images acquired with a small animal device. Dynamic SPECT images provide functional information targeted by a specific radiotracer (99mTc-pertechnetate) that permit the tracking and quantifying of evolving phenomena.

- Respiratory motion induces an artificial enlargement of the moving structures (tumours, organs) in SPECT images, and biases the quantification.
- A full ad-hoc method was presented that allows the reconstruction of a single 3D SPECT image without motion artefacts [37], [6], [1].

5.2.2. Pre-clinical molecular dynamic imaging: ^{99}mTc-pertechnetate biodistribution model of murine stomach with micro-SPECT

Participants: Marine Breuilly [Correspondant, Inria], Grégoire Malandain [Inria], Nicholas Ayache [Inria], Jacques Darcourt [UNS-CAL], Philippe Franken [UNS-CAL], Thierry Pourcher [CEA].

This work is jointly conducted with the Transporter in Imagery and Oncologic Radiotherapy team (TIRO, CEA-CAL-UNSA) located in Nice.

SPECT/CT, small animal, 4D images, stomach, $^{99}\mathrm{m}$Tc-pertechnetate biodistribution, compartmental analysis

Using the coupled SPECT and CT device dedicated to small animals, functional information targeted by a specific radiotracer ($^{99}\mathrm{m}$Tc-pertechnetate) can be imaged dynamically.

- $^{99}\mathrm{m}$Tc-pertechnetate is an iodide analog related to the NIS gene. Thus iodide uptake kinetics can be studied through the study of $^{99}\mathrm{m}$Tc- pertechnetate biodistribution.
- Dynamic SPECT images exhibit a progressive accumulation of $^{99}\mathrm{m}$Tc-pertechnetate in the stomach wall and diffusion in the stomach cavity.
- A first simplified model for stomach $^{99}\mathrm{m}$Tc-pertechnetate biodistribution was proposed and studied with a compartmental analysis approach using a simplified two-compartment (stomach wall and cavity) model with one input (blood) (see Figure 4) [1].
- Time activity curves of each compartment were obtained from dynamic images thanks to an original layer-based decomposition of the stomach [1].
- The first estimation of the model transfer parameters K_{ij} was performed by numerically solving the inverse problem [1].

../../../../projets/asclepios/IMG/breuilly.png

Figure 4. Simplified two-compartment model of $^{99}\mathrm{m}$Tc-pertechnetate biodistribution in murine stomach with time-activity curves for each compartment.

5.3. Computational Anatomy

5.3.1. Longitudinal brain morphometry: statistical analysis and robust quantification of anatomical changes

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

Longitudinal analysis, Alzheimer's Disease, non-linear registration, brain morphometry

This project is based on the PhD thesis defended in 2012 by Marco Lorenzi, and aims at developing robust and effective instruments for the analysis of longitudinal brain changes, with special focus on the study of brain atrophy in Alzheimer's disease. The project relies on the analysis of follow-up magnetic resonance images of the brain by means of non-linear registration. During 2013 the main scientific achievements were the following:

- We developed and distributed the LCC-logDemons, an accurate and robust diffeomorphic non-linear registration algorithm [14], [16]. The algorithm implements the symmetric Local Correlation Coefficient (LCC) and is suited for both inter and intra-subject registration. The software is freely available for research purposes here.
- We investigated the problem of comparing the trajectories of longitudinal morphological changes estimated in different patients. Based on our previous work on parallel transport in diffeomorphic registration, we proposed the "pole ladder" for the efficient normalization of longitudinal trajectories in a common reference space [15], [48].
- We defined an effective framework for the statistical analysis of longitudinal brain changes in clinical groups. The proposed framework enabled the characterization of abnormal morphological changes in healthy subjects at risk for Alzheimer's disease [46].
- We addressed the multi-scale analysis of longitudinal volume changes encoded by deformation fields. We provided a probabilistic framework for the consistent definition of anatomical regions of longitudinal brain atrophy across spatial scales, in order to robustly quantify regional volume changes in populations or in single patients. The framework was applied to the longitudinal analysis of group-wise atrophy in Alzheimer's disease (Figure 5), and to the tracking and quantification of treatment efficacy on brain tumors [47].

5.3.2. Longitudinal Analysis and Modeling of Brain Development during Adolescence

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

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

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

Due to the lack of tools to capture the subtle changes in the brain, little is known about its development during adolescence. The aim of this project is then to provide quantification and models of brain development during adolescence based on non-rigid registration of longitudinal MRIs (enabling us to capture these changes). The analysis pipeline is the following (Figure 6) :

- Register each patient's pair of images in order to get access to the longitudinal changes defined by a transformation field (parameterized by a Stationary Velocity Field).
- Transport every deformation field in a common space (template) to obtain the mean scenario and quantify the changes.
- Propose simplified models of the anatomical changes occurring during adolescence abstracting the results of the analysis.

5.3.3. Reduced-Order Statistical Models of Cardiac Growth, Motion and Blood Flow

Participants: Kristin Mcleod [Correspondant], Maxime Sermesant, Xavier Pennec.

This work was partially funded by the EU projects Care4me (ITEA2) and MD-Paedigree (FP7).

../../../../projets/asclepios/IMG/FigureMarcoLorenzi.png

Figure 5. Group-wise scale-space analysis for the 1-year brain atrophy in 30 AD patients.

../../../../projets/asclepios/IMG/Pipeline_all_wt.png

Figure 6. Pipeline for the longitudinal analysis of brain development during adolescence.

Statistical analysis, image registration, Demons algorithm, reduced models, CFD, Polyaffine, cardiac motion tracking

This work involves developing reduced models of cardiac growth, motion and blood flow, with application to the Tetralogy of Fallot heart [28].

- Extending the 2012 reduced order model of cardiac motion based on a polyaffine log-demons registration proposed at the 2012 STACOM MICCAI workshop, an additional cardiac-specific prior was added to the model to give more physiologically meaningful weight functions. Using this method, the trace of the affine matrix per region was plotted over time to establish differences between healthy subjects and asynchronous heart failure patients. The method and results were presented at the 2013 FIMH conference [52].
- Going further in analysing the affine parameters per region, statistical methods were applied to the registration parameters of the method proposed at the 2012 STACOM MICCAI workshop [50]. By applying principal component analysis to the transformation parameters stacked either column-wise or row-wise, population-based descriptors of motion in terms of the temporal or spatial components were obtained. The method was applied to 15 healthy subjects and 2 heart failure patients and presented at the 2013 MICCAI conference [51].
- The analysis of a statistical model for reduced blood flow simulations in the pulmonary artery proposed in the 2010 STACOM workshop was extended to a journal version [10], [64]. The previous work was extended to re-solve the obtained pressure and velocity bases for the subject-specific geometry by solving the Navier Stokes equations on the reduced bases. The method was applied to a data-set of 17 Tetralogy of Fallot patients.

5.3.4. Geometric Statistics

Participants: Xavier Pennec [Correspondant], Nina Miolane, Christof Seiler [Stanford], Susan Holmes [Stanford].

This work is partly funded through a France Stanford collaborative project grant (2013-2014).

Statistics, manifolds, Lie groups

The study of bi-invariant means on Lie groups [53] was further pushed by looking for the conditions of existence of bi-invariant semi-Riemannian metrics, thus relaxing the positivity constraint of Riemannian metrics [4]. This idea was based on the fact that such a bi-invariant semi-Riemannian metric exists of SE(3). Unfortunately, this does not generalize to higher dimensions. Other results on geometric statistics on regions for in the context of group-valued trees for deformation analysis were presented in [55].

5.4. Computational Physiology

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

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

Alzheimer's Disease (AD), modeling atrophy, bio-physical model, simulation

We have implemented a 3D bio-physical model for the deformation of the brain with Alzheimer's Disease (AD).The model produces a deformation field of the brain when a known distribution of local volume change (atrophy) is given as the input. The obtained deformation is then used to warp the original 3D MR image. The major contribution of this work corresponds to the block "Brain Deformation" in Figure 7 .

5.4.2. Registration of time series of cardiac images

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

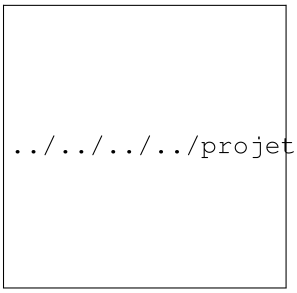
This work has been partly supported by Microsoft Research through its PhD Scholarship Programme and the European Research Council through the ERC Advanced Grant MedYMA (on Biophysical Modeling and Analysis of Dynamic Medical Images).

../../../../projets/asclepios/IMG/adModel.png

Figure 7. Modeling and simulation of longitudinal brain MRIs with atrophy in Alzheimer's Disease

Registration, Automatic Relevance Determination, Magnetic resonance, 3D-US

- We developed a generic approach to registration building on the framework of Automatic Relevance Determination. We applied this framework to the tracking of heart motion throughout time series of images from cine-MR, tagged-MR and 3D-US modalities.
- Our approach allows for the joint determination of model parameters, such as noise and regularization parameters, decreasing the need for manual tuning and preprocessing. Moreover, it is suitable for further analysis of uncertainty in the output of the registration.



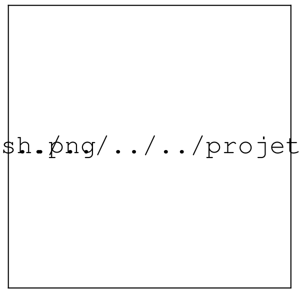


Figure 8. An instance of motion tracking on a cine MR frame. (Left) Mesh contour propagated to end systole via the registration output (Right) Computed displacement field.

5.4.3. Real-Time Cardiac Electrophysiology Computing for Training Simulator

Participants: Hugo Talbot [Correspondant], Hervé Delingette, Stephane Cotin, Maxime Sermesant, Christian Duriez.

This work was performed in collaboration with the SHACRA team in Lille.

Cardiac electrophysiology simulation, Cryoablation simulation, SOFA framework, GPU computing, patient-specific study

- Cardiac arrhythmia is a very frequent pathology that comes from abnormal electrical activity in the myocardium. This work aims at developing a training simulator for cardiologists in the context of catheterization and thermo-ablation of these arrhythmias. After tackling the issue of fast electrophysiology computation [27], a first version of our training simulator was proposed which combines virtual catheterization and interactive GPU electrophysiology modeling [70]. This year, the simulator has been improved by tackling the issue of interactive catheter navigation inside a moving venous system and a beating heart [70]. The simulator was demonstrated during the VRIPHYS 2013 workshop in Lille and the Inria-industry meeting in Paris. Personalization of the electrophysiological model using the data assimilation library Verdandi has been initiated.
- Cryotherapy simulation in collaboration with the IHU Strasbourg has been performed. This technique consists in inserting needles that freeze the surrounding tissues, thus immediately leading to cellular death of the tissues. We built a simulator able to place the cryoprobes and run a simulation representing the evolution of iceballs in living tissues [58].

5.4.4. Personalized model of the heart for cardiac therapy planning

Participants: Stéphanie Marchesseau [Correspondant], Maxime Sermesant, Hervé Delingette, Nicholas Ayache.

Figure 9. Setup of our simulator dedicated to thermo-ablation for cardiac electrophysiology.

This work is performed in the context of the PhD of Stéphanie Marchesseau in collaboration with St Thomas Hospital in London and was partially funded by ERC MedYMA.

- Personalization of the mechanical function of the heart from time series of cardiac images has been achieved by combining global calibration of a few global parameters [18] with estimation of regional contractility parameters [17] using data assimilation techniques.
- Personalized cardiac models were used to create synthetic images [22] of cardiac motion thus allowing the benchmarking of motion tracking algorithms [8], [39].

5.4.5. Cardiac Arrhythmia Radio-frequency Ablation Planning

Participants: Rocio Cabrera Lozoya [Correspondant], Maxime Sermesant, Hervé Delingette, Nicholas Ayache.

This work is performed in the context of the PhD of Rocío Cabrera Lozoya in collaboration with the CHU LIRYC Bordeaux and is funded by ERC MedYMA.

- Biophysical model development for the prediction of radio frequency ablation sites for ventricular tachycardias.
- Target site map generation for ablation therapy guidance
- Structural and functional characterization of target sites using 3D imaging and EP measurements through machine learning algorithms.
- Prediction validation with acquired clinical data

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

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

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

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

- In order to obtain a computational framework for patient-specific planning of radiofrequency ablation, a patient-specific detailed anatomical model of the liver has been extracted from a standard CT image and then meshed with tetrahedra. The structures of interest include : parenchyma, lesion, hepatic vein and vena cava.
- A computational fluid dynamic model is used to estimate the patient-specific blood flow in the hepatic circulatory system. It was combined with a porous media model to compute the patient-specific blood flow distribution inside the parenchyma using the results of the CFD solver (pressures).
- Bio-heat equations and a cell death model to account for cellular necrosis have been implemented with FEM using SOFA and a Lattice Boltmann Model to model heat propagation in biological tissues [35] leading to improved accuracy and computational efficiency.

5.4.7. Tumor growth assessment based on biophysical modeling

Participants: Erin Stretton [Correspondant], Bjoern Menze, Nicholas Ayache, Hervé Delingette.

This work was carried out during the Phd of Erin Stretton and was funded by the Care4Me project. It was performed in collaboration with Pr Mandonnet, Lariboisière hospital in Paris, and the German Cancer Research Center (DKFZ)

Glioma simulation, tumor growth.

We aim at developing image analysis methods [23] using tumor growth models in order to guide the planning of therapies (surgical removal and chemotherapy) for brain cancer (glioma) patients. Our work is focused on these objectives :

- Predicting the location of glioma recurrence after a resection surgery;
- Determining the description the of tumor cell diffusion tensor in white matter (patient-based, atlas based or isotropic) which leads to the most accurate results for predicting future tumor growth [57];
- Comparing tumor growth speeds between 8 patient cases based on biophysical modeling and various manual methods.

5.4.8. Brain tumor growth modeling : Application to radiation therapy

Participants: Matthieu Lê [Correspondant], Jan Unkelbach, Nicholas Ayache, Hervé Delingette.

This work is carried out between Asclepios research group, and the Department of Radiation Oncology of the Massachusetts General Hospital, Boston, USA. Part of this work was funded by the European Research Council through the ERC Advanced Grant MedYMA.

Glioma simulations, radiation therapy, target delineation, vasogenic edema

- We developed a tumor growth model for high grade gliomas, based on different types of cell and the vascularization of the brain.
- We studied multimodal brain tumor images to evaluate tumor infiltration.
- We used a Fisher-Kolmogorov model to improve target volume delineation for radiation therapy (see Figure 11)

5.4.9. Multimodal patch-based glioma segmentation

Participants: Nicolas Cordier [Correspondant], Bjoern Menze, Hervé Delingette, Nicholas Ayache.

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

../../../../projets/asclepios/IMG/pipelinebis.png

Figure 10. Steps of the proposed method (blue: input, green: processes, purple: output).

../../../../projets/asclepios/IMG/dose_distribution.png

Figure 11. Comparison of the dose distribution (in Grey) clinically delivered and based on the Fisher-Kolmogorov model.

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

- A patch-based approach was developed for glioma segmentation based on multi-channel 3D MRI. The method is fully automatic and does not require any learning step.
- Features: multi-channel MR intensities in local neighborhoods.
- A heuristic label fusion strategy was introduced and showed promising results, as shown in Figure 12 .
- The algorithm was ranked 5th in the Brain Tumor Segmentation Challenge (BraTS) at MICCAI 2013 [67].
- Large unlabeled glioma MRI databases are being incorporated in the framework.

../../../../projets/asclepios/IMG/CordierFig3.png

Figure 12. Real high-grade case. From left to right: Vote maps for background, necrosis and non-enhancing tumor (merged), edema, enhancing tumor; Segmentation map. From top to bottom: saggital, axial, and coronal views.

ASCOLA Project-Team

6. New Results

6.1. Software composition

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

6.1.1. Service-oriented computing

Services are frequently implemented using object-oriented frameworks. In this context, two properties are particularly important: (i) a loose coupling between the service layer and the object layer, allowing evolution of the service layer with a minimal impact on the object layer, (ii) interoperability induced by the substitution principle associated to subtyping in the object layer, thus allowing to freely convert a value of a subtype into a supertype. However, through experimentation with Apache's popular service framework CXF, we observed some undesirable coupling and interoperability issues due to the failure of the substitution principle [23]. Therefore we have proposed a new specification method for the data binding used to translate data between the object and service layers [24]. We have shown that if the CXF framework follows the specification, the substitution principle is satisfied, with all its advantages.

6.1.2. Modularity and program transformations

Refactoring tools are commonly used for remodularization tasks. Basic refactoring operations are combined to perform complex program transformations, but the resulting composed operations are rarely reused, even partially, because popular tools have few support for composition. In [31], we have recast two calculus for static composition of refactorings in a type system framework and we have discussed their use for inferring useful properties. We have illustrated the value of support for static composition in refactoring tools with a complex remodularization use case: a round-trip transformation between programs conforming to the Composite and Visitor patterns. Composite and Visitor design patterns have dual properties with respect to modularity, thus they are good candidates to explore their transformations. In [22] we have extended our initial refactoring-based round-trip transformation between these two structures and we have studied how that transformation is impacted by four variations in the implementation of these patterns. We have validated that study by computing the smallest preconditions for the resulting transformations. We have also automated the transformation and applied it to JHotDraw, where the studied variations occur. Finally, [11] presents more exhaustively modular transformations and design patterns. We have also proposed a reversible transformation in the Singleton pattern to benefit from optimization by introducing this pattern and flexibility by its suppression according to the requirements of the software user.

6.1.3. Domain specific languages

In the context of Charles Prud'hommes' PhD Thesis, we have developed a domain specific language in order to specify strategies of filtering propagation in constraint solvers. Indeed, constraint programming replaces brute force generate-and-test by the exploration of the solution space based on incremental instantiation and constraint propagation. Strategies of incremental instantiation (also known as heuristics) have been heavily studied. However, most solvers propagate constraints with a simple fix point computation based on a queue of constraints to propagate (or several queues in order to deal with the grain/cost of filtering algorithms). This technique has a good behavior in general but for a given problem a dedicated strategy can be more efficient. Our declarative DSL and its support in the new version of the constraint solver Choco [19], [52] enables us to easily experiment with different propagation strategies. Moreover, our DSL supports properties such as completeness, intended incompleteness or non ambiguity.

6.1.4. Constructive security

In the field of techniques for the development of secure software systems we have presented results on the enforcement of security properties in service-oriented systems and Javascript programs.

Concerning the security of service-based systems, we have first presented a software framework that harnesses a type based policy language and aspect-based support for protocol adaptation in service-oriented systems by means of flexible reference monitors [29], [28]. We have shown how this framework improves the security, interoperability and evolution issues of service systems using the OAuth 2.0 standard for the authorization of resource accesses. The OAuth 2 protocol is a recent IETF standard devoted to providing authorization to clients requiring access to specific resources over HTTP. It was recently adopted by major internet companies and software editors, such as Google, Facebook, Microsoft, and SAP. We have shown how to improve the security of software systems that use OAuth 2 in the presence of different kinds of clients.

Furthermore, we have developed a new notion of transformation operators, so-called workflow adaptation schemas (WASs) for service compositions that facilitates the integration and modification of security functionalities of service-oriented systems [30]. These schemas may be generic and specialized through parameter instantiation. A set of schemas therefore effectively provides a domain-specific language for the transformation of service-oriented applications. We have developed a set of specific schemas and applied them to the OAuth 2 standard in order to implement state-based security hardening strategies. We have also implemented tool support for WASs and implemented some of the security scenarios involving OAuth 2 (see Sec. 5.4).

Finally, we have shown that a wide range of strategies to make secure JavaScript-based applications can be described pertinently using aspects [42]. To this end, we have reviewed major categories of approaches to make client-side applications secure and have discussed uses of aspects that exist for some of them. We also propose aspect-based techniques for the categories that have not been studied previously. We have given examples of applications where aspects are useful as a general means to flexibly express and implement security policies for JavaScript.

6.2. Aspect-Oriented Programming

Participants: Rémi Douence, Ismael Figueroa, Jacques Noyé, Mario Südholt, Nicolas Tabareau, Jurgen Van Ham.

6.2.1. Aspects in a concurrent and distributed setting

Aspect oriented programming modularizes crosscutting concerns by gathering several join points. In the context of distributed applications these point cuts can be on different machines. In this case, a sequence of join points must be defined as a sequence of logical joint points (à la Lamport). We propose an aspect oriented languages to define distributed aspects in JavaScript in a distributed context. Our proposal [18] is based on vector clocks in order to logically relate join points and can ignore "illogical" (that is late or early) join points. In can also enforce causal communications when no join point must be discarded. We have exemplified the advantages of our technique with different applications such as a discussion forum, a retweet scenario and a web browser.

Multiparty session types allow the definition of distributed processes with strong communication safety properties. A global type is a choreographic specification of the interactions between peers, which is then projected locally in each peer. Well-typed processes behave accordingly to the global protocol specification. Multiparty session types are however monolithic entities that are not amenable to modular extensions. Also, session types impose conservative requirements to prevent any race condition, which prohibit the uniform application of extensions at different points in a protocol. We have proposed a means to support modular extensions with *aspectual session types* [47], a static pointcut/advice mechanism at the session type level. To support the modular definition of crosscutting concerns, we augment the expressivity of session types to allow harmless race conditions. We formally prove that well-formed aspectual session types entail communication safety. As a result, aspectual session types make multiparty session types more flexible, modular, and extensible.

We have added dedicated concurrency support to EScala, our extension of Scala that introduces composable *declarative events* as a way to integrate Aspect-Oriented Programming and Event-Based Programming in the context of Object-Oriented Programming. In JEScala, Events, which were synchronous in EScala, can be declared as *asynchronous* so that they are handled concurrently to their emitter. Moreover, two new operators, a join and a choice operator, inherited from the join calculus - hence the name of the new prototype, can now be used to compose events and control concurrency. In [48], we present JEScala, show that it captures coordination schemas in a more expressive and modular way than plain join languages and provide a first performance assessment.

6.2.2. Effective aspects

We have proposed a novel approach to embed pointcut/advice aspects in a typed functional programming language like Haskell. Aspects are first-class, can be deployed dynamically, and the pointcut language is extensible. Type soundness is guaranteed by exploiting the underlying type system, in particular phantom types and a new anti-unification type class. The use of monads brings type-based reasoning about effects for the first time in the pointcut/advice setting and enables modular extensions of the aspect language [46], [16].

To allow a type-safe embedding of aspects in Haskell, we had to develop a notion of anti-unification in Haskell type system. The anti-unification problem is that of finding the most specific pattern of two terms. While dual to the unification problem, anti-unification has rarely been considered at the level of types. We have developed an algorithm to compute the least general type of two types in Haskell, using the logic programming power of type classes [53]. That is, we have defined a type class for which the type class instances resolution performs anti-unification.

6.2.3. Reasoning about aspect interference

When a software system is developed using several aspects, special care must be taken to ensure that the resulting behavior is correct. This is known as the *aspect interference problem*, and existing approaches essentially aim to detect whether a system exhibits problematic interferences of aspects. We have described how to control aspect interference by construction by relying on the type system. More precisely, we combine a monadic embedding of the pointcut/advice model in Haskell with the notion of membranes for aspect-oriented programming [34]. Aspects must explicitly declare the side effects and the context they can act upon. Allowed patterns of control flow interference are declared at the membrane level and statically enforced. Finally, computational interference between aspects is controlled by the membrane topology. To combine independent and reusable aspects and monadic components into a program specification we use *monad views*, a recent technique for conveniently handling the monadic stack.

Oliveira and colleagues recently developed a powerful model to reason about mixin-based composition of effectful components and their interference, exploiting a wide variety of techniques such as equational reasoning, parametricity, and algebraic laws about monadic effects. Our work addresses the issue of reasoning about interference with effectful aspects in the presence of unrestricted quantification through pointcuts. While global reasoning is required, we have shown that it is possible to reason in a compositional manner, which is key for the scalability of the approach in the face of large and evolving systems. We have established a general equivalence theorem that is based on a few conditions that can be established, reused, and adapted separately as the system evolves. Interestingly, one of these conditions, local harmlessness, can be proven by a translation to the mixin setting, making it possible to directly exploit previously established results about certain kinds of harmless extensions [33].

In aspect-oriented programming (AOP) languages, advice evaluation is usually considered as part of the base program evaluation. While viewing aspects as part of base level computation clearly distinguishes AOP from reflection, it also comes at a price: because aspects observe base level computation, evaluating pointcuts and advice at the base level can trigger infinite regression. To avoid these pitfalls, we have introduced levels of execution in the programming language, thereby allowing aspects to observe and run at specific, possibly different, levels. We adopt a defensive default that avoids infinite regression, and gives advanced programmers the means to override this default using level-shifting operators [21].

6.3. Resource management in Cloud computing

Participants: Frederico Alvares, Gustavo Bervian Brand, Yousri Kouki, Adrien Lèbre, Thomas Ledoux, Guillaume Le Louët, Jean-Marc Menaud, Jonathan Pastor, Flavien Quesnel, Mario Südholt.

We have contributed on several topics: multiple autonomic managers for Cloud infrastructure, SLA management for Cloud elasticity, fully distributed and autonomous virtual machine scheduling, and simulator toolkits for IaaS platforms.

6.3.1. Cloud infrastructure based on multiple autonomic managers

One of the main reasons for the wide adoption of Cloud Computing is the concept of elasticity. Implementing elasticity to tackle varying workloads while optimizing infrastructures (e.g. utilization rate) and fulfilling the application requirements on Quality of Service should be addressed by self-adaptation techniques able to manage complexity and dynamism. However, since Cloud systems are organized in different but dependent Cloud layers, self-management decisions taken in isolation in a certain layer may indirectly interfere with the decision taken by an other layer. Indeed, non-coordinated managers may lead to conflicting decisions and consequently to non-desired states.

We have proposed a framework for the coordination of multiple autonomic managers in cloud environments [25]. The PhD thesis of Frederico Alvares [12], defended in April 2013, is based on this framework. This thesis proposes a self-adaptation approach that considers both application internals (architectural elasticity) and infrastructure (resource elasticity), managed by multiple autonomic managers, to reduce the energy footprint in Cloud infrastructures.

6.3.2. SLA Management for Cloud elasticity

Elasticity is the intrinsic element that differentiates Cloud Computing from traditional computing paradigms, since it allows service providers to rapidly adjust their needs for resources to absorb the demand and hence guarantee a minimum level of Quality of Service (QoS) that respects the Service Level Agreements (SLAs) previously defined with their clients. However, due to non-negligible resource initiation time, network fluctuations or unpredictable workload, it becomes hard to guarantee QoS levels and SLA violations may occur. The main challenge of service providers is to maintain its consumer's satisfaction while minimizing the service costs due to resources fees. The PhD thesis of Yousri Kouki [13], defended in December, proposes different contributions to address this issue: CSLA, a specific language to describe SLA for Cloud services ; HybridScale, an auto-scaling framework driven by SLA [39], [17].

6.3.3. Fully Distributed and Autonomous Virtualized Environments

We have consolidated the DVMS system to obtain a fully distributed virtual machine scheduler [44]. This system makes it possible to schedule VMs cooperatively and dynamically in large scale distributed systems. Simulations (up to 64K VMs) and real experiments both conducted on the Grid'5000 large-scale distributed system [44] showed that DVMS is scalable. This building block is a first element of a more complete cloud OS, entitled DISCOVERY (DIStributed and COoperative mechanisms to manage Virtual EnviRonments autonomicallY) [56]. The ultimate goal of this system is to overcome the main limitations of the traditional server-centric solutions. The system, currently under investigation in the context of the Jonathan Pastor's PhD, relies on a peer-to-peer model where each agent can efficiently deploy, dynamically schedule and periodically checkpoint the virtual environments it manages.

6.3.4. Testing the cloud

Computer science, as other sciences, needs instruments to validate theoretical research results, as well as software developments. Although simulation and emulation are generally used to get a glance of the behavior of new algorithms, they use over-simplified models in order to reduce their execution time and thus cannot be accurate enough. Leveraging a scientific instrument to perform actual experiments is an undeniable advantage. However, conducting experiments on real environments is still too often a challenge for researchers, students, and practitioners: first, because of the unavailability of dedicated resources, and second, because of the inability to create controlled experimental conditions, and to deal with the wide variability of software requirements. During 2013, we have contributed to a new topic addressing the "testing the cloud" challenge. First, we have presented the latest mechanisms we have designed to enable the automated deployment of the major open-source IaaS cloudkits (i.e., Nimbus, OpenNebula, CloudStack, and OpenStack) on Grid'5000 [26]. Providing automatic, isolated and reproducible deployments of cloud environments lets end-users study and compare each solution or simply leverage one of them to perform higher-level cloud experiments (such as investigating Map/Reduce frameworks or applications). Moreover, we have presented EXECO, a library that provides easy and efficient control of large scale experiments through a set of tools well as tools designed for scripting distributed computing experiments on any computing platform. We have illustrated its interest by presenting two experiments dealing with virtualization technologies on the Grid'5000 testbed [37].

6.3.5. Adding virtualization abstractions into the Simgrid toolkit

In the context of the ANR SONGS project and in collaboration with Takahiro Hirofuchi, researcher at AIST (Japan), we have extended the Simgrid framework to be able to simulate virtualized distributed infrastructures [35]. In addition, we have proposed the first class support of live migration operations within such a simulator toolkit for large scale distributed infrastructures. We have developed a resource share calculation mechanism for VMs and a live migration model implementing the precopy migration algorithm of Qemu/KVM. We have confirmed that our simulation framework correctly reproduced live migration behaviors of the real world under various conditions [36].

6.3.6. Power and energy management in the cloud

Power management has become one of the main challenges for data center infrastructures. Currently, the cost of powering a server is approaching the cost of the server hardware itself, and, in a near future, the former will continue to increase, while the latter will go down. In this context, virtualization is used to decrease the number of servers, and increase the efficiency of the remaining ones.

First, in [43] we have proposed an approach and a model to estimate the total power consumption of a virtual machine, by taking into account its static (e.g. memory) and dynamic (e.g. CPU) consumption of resources. Second, we have rewritten the Entropy framework (in OptiPlace) to give it the support of external models, named views. Entropy, based on the Constraint Programming solver Choco written in Java, does not really scale well. We have studied Entropy's scalability properties [32] and have then integrated heuristics and constraints in OptiPlace [40].

The evaluation of these policies on real infrastructures has become an important and difficult issue. The corresponding techniques have become so complex that there is a need for load injection frameworks able to inject resource load in a tested datacenter instead of model-driven simulation. For this reason we have developed StressCloud [41], [51], a framework to manipulate the activities of a group of Virtual Machines and observe the resulting performance.

ASPI Project-Team

5. New Results

5.1. Iterative isotone regression

Participant: Arnaud Guyader.

This is a collaboration with Nicolas Hengartner (Los Alamos), Nicolas Jégou (université de Rennes 2) and Eric Matzner–Løber (université de Rennes 2), and with Alexander B. Németh (Babeş Bolyai University) and Sándor Z. Németh (University of Birmingham).

We explore some theoretical aspects of a recent nonparametric method for estimating a univariate regression function of bounded variation. The method exploits the Jordan decomposition which states that a function of bounded variation can be decomposed as the sum of a non-decreasing function and a non-increasing function. This suggests combining the backfitting algorithm for estimating additive functions with isotonic regression for estimating monotone functions. The resulting iterative algorithm is called IIR (iterative isotonic regression). The main result in this work [22] states that the estimator is consistent if the number of iterations k_n grows appropriately with the sample size n. The proof requires two auxiliary results that are of interest in and by themselves: firstly, we generalize the well-known consistency property of isotonic regression to the framework of a non-monotone regression function, and secondly, we relate the backfitting algorithm to the von Neumann algorithm in convex analysis. We also analyse how the algorithm can be stopped in practice using a data-splitting procedure.

With the geometrical interpretation linking this iterative method with the von Neumann algorithm, and making a connection with the general property of isotonicity of projection onto convex cones, we derive in [14] another equivalent algorithm and go further in the analysis.

5.2. Mutual nearest neighbors

Participant: Arnaud Guyader.

This is a collaboration with Nicolas Hengartner (Los Alamos).

Motivated by promising experimental results, this work [13] investigates the theoretical properties of a recently proposed nonparametric estimator, called the MNR (mutual nearest neighbors) rule, which estimates the regression function $m(x) = E[Y|X = x]$ as follows: first identify the k nearest neighbors of x in the sample, then keep only those for which x is itself one of the k nearest neighbors, and finally take the average over the corresponding response variables. We prove that this estimator is consistent and that its rate of convergence is optimal. Since the estimate with the optimal rate of convergence depends on the unknown distribution of the observations, we also have adaptation results by data-splitting.

5.3. Adaptive multilevel splitting

Participants: Frédéric Cérou, Arnaud Guyader, Florent Malrieu.

This is a collaboration with Pierre Del Moral (EPI ALEA, Inria Bordeaux—Sud Ouest).

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

5.4. Total variation estimates for the TCP process

Participant: Florent Malrieu.

This is a collaboration with Jean-Baptiste Bardet (université de Rouen), Alejandra Christen (University of Chile), Arnaud Guillin (université de Clermont–Ferrand), and Pierre–André Zitt (université de Paris–Est Marne–la–Vallée).

The TCP window size process appears in the modeling of the famous Transmission Control Protocol used for data transmission over the Internet. This continuous time Markov process takes its values in $[0, \infty)$, is ergodic and irreversible. The sample paths are piecewise linear deterministic and the whole randomness of the dynamics comes from the jump mechanism. The aim of [27] is to provide quantitative estimates for the exponential convergence to equilibrium, in terms of the total variation and Wasserstein distances, using coupling methods. The technique could be applied to a large class of Markov processes as well.

5.5. On the stability of planar randomly switched systems

Participant: Florent Malrieu.

This is a collaboration with Michel Benaïm (université de Neuchâtel), Stéphane Le Borgne (IRMAR) and Pierre–André Zitt (université de Paris–Est Marne–la–Vallée).

The paper [28] illustrates some surprising instability properties that may occur when stable ODE's are switched using Markov dependent coefficients. Consider the random process (X_t) solution of $dX_t/dt = A(I_t)X_t$ where (I_t) is a Markov process on $\{0, 1\}$ and A_0 and A_1 are real Hurwitz matrices on $\mathbb{R}^2$. Assuming that there exists $\lambda \in (0, 1)$ such that $(1 - \lambda)A_0 + \lambda A_1$ has a positive eigenvalue, we establish that the norm of X_t may converge to 0 or infinity, depending on the the jump rate of the process I. An application to product of random matrices is studied. This work can be viewed as a probabilistic counterpart of the paper [26] by Baldé, Boscain and Mason.

5.6. Marginalization in rare event simulation for switching diffusions

Participant: François Le Gland.

This is a collaboration with Anindya Goswami (IISER, Poone).

Switching diffusions are continuous–time Markov processes with a hybrid continuous / finite state space. A rare but critical event (such as a scalar function of the continuous component of the state exceeding a given threshold) can occur for several reasons:

- the process can remain in *nominal* mode, where the critical event is very unlikely to occur,
- or the process can switch in some *degraded* mode, where the critical event is much more likely to occur, but the switching itself is very unlikely to occur.

Not only is it important to accurately estimate the (very small) probability that the critical event occurs before some fixed final time, but it is also important to have an accurate account on the reason why it occured, or in other words to estimate the probability of the different modes. A classical implementation of the multilevel splitting would not be efficient. Indeed, as soon as (even a few) samples paths switch to a *degraded* mode, these sample paths will dominate and it will not be possible to estimate the contribution of samples paths in the *nominal* mode. Moreover, sampling the finite component of the state is not efficient to accurately estimate the (very small) probability of rare but critical modes. A more efficient implementation is based on marginalization, i.e. in sampling jointly the continuous component and the probability distribution of the finite component given the past continuous component [18]. The latter is a probability vector, known as the Wonham filter, that satisfies a deterministic equation.

5.7. Combining importance sampling and multilevel splitting for rare event simulation

Participants: François Le Gland, Damien–Barthélémy Jacquemart.

This is a collaboration with Jérôme Morio (ONERA, Palaiseau).

The problem is to accurately estimate the (very small) probability that a rare but critical event (such as a scalar function of the state exceeding a given threshold) occurs before some fixed final time. Multilevel splitting is a very efficient solution, in which sample paths are propagated and are replicated when some intermediate events occur. Events that are defined in terms of the state variable only (such as a scalar function of the state exceeding an intermediate threshold) are not a good design. A more efficient but more complicated design would be to let the intermediate events depend also on time. An alternative design is to keep intermediate events simple, defined in terms of the state variable only, and to make sure that samples that exceed the threshold early are replicated more than samples that exceed the same threshold later [19].

5.8. Sequential data assimilation: ensemble Kalman filter vs. particle filter

Participants: François Le Gland, Valérie Monbet.

The contribution has been to prove (by induction) the asymptotic normality of the estimation error, i.e. to prove a central limit theorem for the ensemble Kalman filter. Explicit expression of the asymptotic variance has been obtained for linear Gaussian systems (where the exact solution is known, and where EnKF is unbiased). This expression has been compared with explicit expressions of the asymptotic variance for two popular particle filters: the bootstrap particle filter and the so–called optimal particle filter, that uses the next observation in the importance distribution.

5.9. Non–homogeneous Markov–switching models

Participant: Valérie Monbet.

This is a collaboration with Pierre Ailliot (université de Bretagne occidentale, Brest).

We have developped various hidden non–homogeneous Markov–switching models for description and simulation of univariate and multivariate time series. Considered application are in weather variables modelling but also in economy. The main originality of the proposed models is that the hidden Markov chain is not homogeneous, its evolution depending on the past wind conditions or other covariates. It is shown that it permits to reproduce complex non–linearities.

5.10. Dynamical partitioning of directional ocean wave spectra

Participant: Valérie Monbet.

This is a collaboration with Pierre Ailliot (université de Bretagne occidentale, Brest) and Christophe Maisondieu (IFREMER, Brest).

Directional wave spectra generally exhibit several peaks due to the coexistence of wind sea generated by local wind conditions and swells originating from distant weather systems. The paper [24] proposes a new algorithm for partitioning such spectra and retrieving the various systems which compose a complex sea-state. It is based on a sequential Monte Carlo algorithm which allows to follow the time evolution of the various systems. The proposed methodology is validated on both synthetic and real spectra and the results are compared with a method commonly used in the literature.

5.11. Track–before–detect

Participants: François Le Gland, Alexandre Lepoutre.

This is a collaboration with Olivier Rabaste (ONERA, Palaiseau).

The problem considered in [20] is tracking one or several targets in a track–before–detect (TBD) context using particle filters. These filters require the computation of the likelihood of the complex measurement given the target states. This likelihood depends on the complex amplitudes of the targets. When the complex amplitude fluctuates over time, time coherence of the target cannot be taken into account. However, for the single target case, spatial coherence of this amplitude can be taken into account to improve the filter performance, by marginalizing the likelihood of the complex measurement over the amplitude parameter. The marginalization depends on the fluctuation law considered. We show that for the Swerling 1 model the likelihood of the complex measurement can be obtained analytically in the multi-target case. For the Swerling 0 model no closed form can be obtained in the general multi–target setting. Therefore we resort to some approximations to solve the problem. Finally, we demonstrate with Monte Carlo simulations the gain of this method both in detection and in estimation compared to the classic method that works with the square modulus of the complex signal.

The problem considered in [21] is detecting and tracking a single radar target with amplitude fluctuation Swerling 1 and 3 in a track–before–detect context with particle filter. Those fluctuations are difficult to take into account as they are uncoherent from measurement to measurement. Thus, conventionnal filters work on square modulus of the complexe signal to remove the unknown phase of complex amplitude and the marginalized over the law of the modulus but they lose the spatial coherence of the amplitude in the measurement. We show in this paper that complex measurements can be marginalized directly while taking into account the spatial coherence of the complex amplitude. Finally, we show the benefit of this method both in detection and in estimation via Monte Carlo simulations.

ATEAMS Project-Team

5. New Results

5.1. Empirical analyses of source code

Rascal was used to perform empirical investigations of existing source code bases. First of all, Davy Landman performed an analysis of project management source code to investigate if domain knowledge is present in source code and, if so, how easy it is to extract that knowledge [26]. An earlier experiment in static analysis of PHP code was finalized by Mark Hills. The result is a deep study of feature usage in a large number of well-known PHP projects [25]. Vadim Zaytsev conducted an experiment to recognize micro-patterns in grammars and meta-models [32]. Finally, Jeroen van den Bos performed a deep empirical study to find out as to how far a domain-specific language facilitates evolution [34]. The results showed that the Derric DSL did indeed cover most evolution scenarios, but there is still room for improving the language. In all cases Rascal proved to be instrumental in performing the experiments.

5.2. Better parsing and disambiguation

Ali Afroozeh worked on a new implementation of GLL parsing, called Iguana. Unlike traditional parser generators, Iguana adopted the interpretive approach that is also used in the Ensō parser. This experiment is still ongoing, but the new parser is expected to be integrated into Rascal beginning of 2014. Additionally, a longstanding problem of disambiguation using operator precedence was solved [23]. Traditional approaches are either not safe (i.e. they make the language smaller), or they do not support complex precedence rules as found in, for instance, OCaml.

5.3. Extensible Programming

Modular and extensible implementation of languages could have major impact on how DSLs will be implemented. Anastasia Izmaylova continued here work on improving the extensibility of Rascal's module system, by providing open recursive function combinators.

Extensible programming is traditionally plagued by what has become known as "the expression problem", which captures the fact that most programming languages either support extension of data variants, or extension of operations, but not both. Object Algebras are simple solution to this problem. In [30] we have extended this model to support feature-oriented programming. These results are currently being integrated into the Ensō system.

5.4. DSLs for Games

In collaboration with the Hogeschool van Amsterdam, Riemer van Rozen developed a workbench for MicroMachinations, a DSL for game economies [28]. Completely built using Rascal, this DSL environment features syntax highlighting, static analysis, interactive simulation, and SPIN-based model-checking of process models describing the economy of a game. The project shows the versatility of Rascal as a language workbench for the development of DSLs.

5.5. DSLs for Questionnaires

In the context of computational auditing we have intensified our research on DSLs for questionnaires. It was proposed by Tijs van der Storm as the benchmark task for the Language Workbench Challenge 2013 (LWC'13), which has resulted in a thorough overview and qualitative comparison of language workbenches [24]. As a side-effect, there are now two publicly available Rascal implementations of the questionnaire DSL (QL-R-Kemi and Demoqles). A first step has been made to collect all implementations to create a "chrestomathy" for further study and dissemination of language workbench concepts and DSL implementation patterns. Other results include a formal semantics of the dynamics of questionnaires [21], and an initial prototype of a questionnaire model for modeling the Dutch Tax Income filing application by Pablo Inostroza Valdera.

5.6. Live Programming

Live programming aims to bring the dynamic execution of programs closer to the programmer, ideally almost obliterating the gap between editing and executing the program. We are working on applying such principles in the context of DSLs. This has lead to two results: a live programming environment for a DSL for questionnaires [36], and Trinity, a data-driven IDE for Derric [35]. Riemer van Rozen has worked on applying similar techniques to MicroMachinations, so that game economies can be adapted at runtime.

5.7. Visualization and interaction

Atze van der Ploeg worked on designing new algorithms and abstractions in the domain of visualization and abstraction. His first result is a fast algorithm for drawing non-layered, tidy trees [20]. DeForm is a library for the declarative specification of resolution-independent 2D graphics [27]. In [31] he proposed a reformulation of the traditional functional reactive programming (FRP) framework, which is both simple and efficient to implement.

5.8. Guarded Coroutines

Anastasia Izmaylova and Paul Klint have built an initial version of a compiler for Rascal. The performance improvements with respect to the interpreter are impressive. Moreover, the design of compiler is based on a new construct for implementing languages with complex backtracking and pattern matching semantics: guarded coroutines. This construct will be instrumental in extending the Rascal language with new kinds of control-flow and concurrency.

5.9. Data structures for meta programming

The efficiency of many meta programs is dependent on the internal data structures used to represent collections, trees, relations etc. Michael Steindorfer has worked on comparing the performance of various persistent collection libraries (e.g., those used in Rascal, Clojure, and Scala). This has lead to a redesign of the PDB collection library that underlies the data structures of Rascal. Furthermore, he developed the Orpheus tool, an object redunancy profiler to assess the effects of maximal sharing.

ATHENA Project-Team

6. New Results

6.1. Improving Diffusion MRI Signal and Acquisition

6.1.1. Design of multishell sampling schemes with uniform coverage in diffusion MRI

Participants: Emmanuel Caruyer [SBIA, University of Pennsylvania Medical School,USA], Christophe Lenglet [CMRR, Department of Radiology, University of Minnesota,USA], Guillermo Sapiro [Electrical & Computer Engineering Dept, Duke University,USA], Rachid Deriche.

In diffusion MRI, a technique known as diffusion spectrum imaging reconstructs the propagator with a discrete Fourier transform, from a Cartesian sampling of the diffusion signal. Alternatively, it is possible to directly reconstruct the orientation distribution function in q-ball imaging, providing so-called high angular resolution diffusion imaging. In between these two techniques, acquisitions on several spheres in q-space offer an interesting trade-off between the angular resolution and the radial information gathered in diffusion MRI. A careful design is central in the success of multishell acquisition and reconstruction techniques.

The design of acquisition in multishell is still an open and active field of research, however. In this work, we provide a general method to design multishell acquisition with uniform angular coverage. This method is based on a generalization of electrostatic repulsion to multishell.

The impact of our method on the angular resolution in one and two bundles of fiber configurations is evaluated using simulations. Compared to more commonly used radial sampling, we show that our method improves the angular resolution, as well as fiber crossing discrimination.

This work has been published in [14].

6.1.2. Motion detection in diffusion MRI via online ODF estimation

Participants: Emmanuel Caruyer [SBIA, University of Pennsylvania Medical School,USA], Iman Aganj [Martinos Center for Biomedical Imaging, MGH, Harvard Medical School,USA], Christophe Lenglet [CMRR, Department of Radiology, University of Minnesota,USA], Guillermo Sapiro [Electrical & Computer Engineering Dept, Duke University,USA], Rachid Deriche.

The acquisition of high angular resolution diffusion MRI is particularly long and subject motion can become an issue. The orientation distribution function (ODF) can be reconstructed online incrementally from diffusion-weighted MRI with a Kalman filtering framework. This online reconstruction provides real-time feedback throughout the acquisition process. In this work, the Kalman filter is first adapted to the reconstruction of the ODF in constant solid angle. Then, a method called STAR (STatistical Analysis of Residuals) is presented and applied to the online detection of motion in high angular resolution diffusion images. Compared to existing techniques, this method is image based and is built on top of a Kalman filter. Therefore, it introduces no additional scan time and does not require additional hardware. The performance of STAR is tested on simulated and real data and compared to the classical generalized likelihood ratio test. Successful detection of small motion is reported (rotation under 2 degrees) with no delay and robustness to noise.

This work has been published in [13].

6.1.3. A Robust variational approach for simultaneous smoothing and estimation of DTI

Participants: Rachid Deriche, Meizhu Liu [Department of CISE, University of Florida, Gainesville, USA], Baba C. Vemuri [Department of CISE, University of Florida, Gainesville, USA].

Estimating diffusion tensors is an essential step in many applications — such as diffusion tensor image (DTI) registration, segmentation and fiber tractography. Most of the methods proposed in the literature for this task are not simultaneously statistically robust and feature preserving techniques. In this work, we propose a novel and robust variational framework for simultaneous smoothing and estimation of diffusion tensors from diffusion MRI. Our variational principle makes use of a recently introduced total Kullback–Leibler (tKL) divergence for DTI regularization. tKL is a statistically robust dissimilarity measure for diffusion tensors, and regularization by using tKL ensures the symmetric positive definiteness of tensors automatically. Further, the regularization is weighted by a non-local factor adapted from the conventional non-local means filters. Finally, for the data fidelity, we use the nonlinear least-squares term derived from the Stejskal–Tanner model. We present experimental results depicting the positive performance of our method in comparison to competing methods on synthetic and real data examples.

This work has been published in [20].

6.1.4. *Tensor estimation and visualization using dMRI*

Participants: Dalila Cherifi [University of Boumerdes, Algeria], Ali Chellouche [University of Boumerdes, Algeria], Amazigh Ait-Ouakli [University of Boumerdes, Algeria], Youcef Benamara [University of Boumerdes, Algeria], Rachid Deriche.

Diffusion tensor imaging in a non-invasive in vivo image modality that allows us to measure molecular diffusion of water in tissues. We characterize diffusion transport of water by an effective diffusion tensor D. The practical importance of the effective diffusion tensor is that it contains new and useful structural and physiological informations about tissues that were previously unobtainable. In this work, we present a software implementation of the estimation of these tensors and their visualization in order to extract these informations.

This work has been published in [28]

6.2. Modeling in Diffusion MRI

6.2.1. *A computational diffusion MRI and parametric dictionary learning framework for modeling the diffusion signal and its features*

Participants: Sylvain Merlet, Emmanuel Caruyer [SBIA, University of Pennsylvania Medical School,USA], Aurobrata Ghosh, Rachid Deriche.

In this work, we first propose an original and efficient computational framework to model continuous diffusion MRI (dMRI) signals and analytically recover important diffusion features such as the Ensemble Average Propagator (EAP) and the Orientation Distribution Function (ODF). Then, we develop an efficient parametric dictionary learning algorithm and exploit the sparse property of a well-designed dictionary to recover the diffusion signal and its features with a reduced number of measurements. The properties and potentials of the technique are demonstrated using various simulations on synthetic data and on human brain data acquired from 7T and 3T scanners. It is shown that the technique can clearly recover the dMRI signal and its features with a much better accuracy compared to state-of-the-art approaches, even with a small and reduced number of measurements. In particular, we can accurately recover the ODF in regions of multiple fiber crossing, which could open new perspectives for some dMRI applications such as fiber tractography.

This work has been published in Medical Image Analysis [21]. It is part of Merlet's PhD thesis defended on Sept. 11th, 2013 [11].

6.2.2. *Continuous diffusion signal, EAP and ODF estimation via compressive sensing in diffusion MRI*

Participants: Sylvain Merlet, Rachid Deriche.

In this work, we exploit the ability of Compressed Sensing (CS) to recover the whole 3D Diffusion MRI (dMRI) signal from a limited number of samples while efficiently recovering important diffusion features such as the Ensemble Average Propagator (EAP) and the Orientation Distribution Function (ODF). Some attempts to use CS in estimating diffusion signals have been done recently. However, this was mainly an experimental insight of CS capabilities in dMRI and the CS theory has not been fully exploited. In this work, we also propose to study the impact of the sparsity, the incoherence and the RIP property on the reconstruction of diffusion signals. We show that an efficient use of the CS theory enables to drastically reduce the number of measurements commonly used in dMRI acquisitions. Only 20–30 measurements, optimally spread on several b-value shells, are shown to be necessary, which is less than previous attempts to recover the diffusion signal using CS. This opens an attractive perspective to measure the diffusion signals in white matter within a reduced acquisition time and shows that CS holds great promise and opens new and exciting perspectives in diffusion MRI (dMRI).

This work has been published in Medical Image Analysis [22]. It is part of Merlet's PhD thesis defended on Sept. 11th, 2013 [11].

6.2.3. Constrained diffusion kurtosis imaging using ternary quartics & MLE

Participants: Aurobrata Ghosh, Tristan Milne, Rachid Deriche.

Diffusion kurtosis imaging (DKI) is a recent improvement over diffusion tensor imaging that characterizes tissue by quantifying non-gaussian diffusion using a 3D fourth-order kurtosis tensor. DKI needs to consider three constraints to be physically relevant. Further, it can be improved by considering the Rician signal noise model. A DKI estimation method is proposed that considers all three constraints correctly, accounts for the signal noise and incorporates efficient gradient-based optimization to improve over existing methods.

In this work, the ternary quartic parameterization is utilized to elegantly impose the positivity of the kurtosis tensor implicitly. Sequential quadratic programming with analytical gradients is employed to solve nonlinear constrained optimization efficiently. Finally, a maximum likelihood estimator based on Rician distribution is considered to account for signal noise.

Extensive experiments conducted on synthetic data verify a MATLAB implementation by showing dramatically improved performance in terms of estimation time and quality. Experiments on in vivo cerebral data confirm that in practice the proposed method can obtain improved results.

This work has been published in [18].

6.2.4. Compressive Sensing DSI

Participants: Sylvain Merlet, Michael Paquette [Sherbrooke Connectivity Imaging Laboratory, Computer Science Departement, Université de Sherbrooke, Québec, Canada], Maxime Descoteaux [Sherbrooke Connectivity Imaging Laboratory, Computer Science Departement, Université de Sherbrooke, Québec, Canada], Rachid Deriche.

Compressive Sensing (CS) offers an efficient way to decrease the number of measurements required in Diffusion Spectrum Imaging (DSI). This method aims to reconstruct the Ensemble Average Propagator (EAP) and, for the purpose of this contest, we compute the numerical Orientation Distribution Function (ODF) by integrating the EAP over a solid angle. In this abstract, we briefly describe three important points underlying the CS technique in order to accelerate DSI, namely the sparsity, the Restricted Isometry Property (RIP) and the L1 reconstruction scheme. Due to the high b-values required in the sampling protocol, our approach enters the heavyweight sampling category. Nevertheless, only 64 measurements are used for the reconstruction.

This work has been published in [31]. It is part of Merlet's PhD thesis defended on Sept. 11th, 2013 [11].

6.2.5. 4th Order symmetric tensors and positive ADC modelling

Participants: Aurobrata Ghosh, Rachid Deriche.

High Order Cartesian Tensors (HOTs) were introduced in Generalized DTI (GDTI) to overcome the limitations of DTI. HOTs can model the apparent diffusion coefficient (ADC) with greater accuracy than DTI in regions with fiber heterogeneity. Although GDTI HOTs were designed to model positive diffusion, the straightforward least square (LS) estimation of HOTs doesn't guarantee positivity. In this work, we address the problem of estimating 4th order tensors with positive diffusion profiles. Two known methods exist that broach this problem, namely a Riemannian approach based on the algebra of 4th order tensors, and a polynomial approach based on Hilbert's theorem on non-negative ternary quartics. In this work, we review the technicalities of these two approaches, compare them theoretically to show their pros and cons, and compare them against the Euclidean LS estimation on synthetic, phantom and real data to motivate the relevance of the positive diffusion profile constraint.

This work has been published in [37].

6.2.6. Higher-Order tensors in diffusion imaging: A survey

Participants: Thomas Schultz [MPI for Intelligent Systems, Tubingen, Germany], Andrea Fuster [Eindhoven University of Technology, The Netherlands], Aurobrata Ghosh, Luc Florack [Eindhoven University of Technology, The Netherlands], Rachid Deriche, Lek-Heng Lim [University of Chicago, USA].

Diffusion imaging is a noninvasive tool for probing the microstructure of fibrous nerve and muscle tissue. Higher-order tensors provide a powerful mathematical language to model and analyze the large and complex data that is generated by its modern variants such as High Angular Resolution Diffusion Imaging (HARDI) or Diffusional Kurtosis Imaging. This survey gives a careful introduction to the foundations of higher-order tensor algebra, and explains how some concepts from linear algebra generalize to the higher-order case. From the application side, it reviews a variety of distinct higher-order tensor models that arise in the context of diffusion imaging, such as higher-order diffusion tensors, q-ball or fiber Orientation Distribution Functions (ODFs), and fourth-order covariance and kurtosis tensors. By bridging the gap between mathematical foundations and application, it provides an introduction that is suitable for practitioners and applied mathematicians alike, and propels the field by stimulating further exchange between the two.

This work has been published in [39].

6.2.7. Regularized spherical polar fourier diffusion MRI with optimal dictionary learning

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

One important problem in diffusion MRI (dMRI) is to recover the diffusion weighted signal from only a limited number of samples in q-space. An ideal framework for solving this problem is Compressed Sensing (CS), which takes advantage of the signal's sparseness or compressibility, allowing the entire signal to be reconstructed from relatively few measurements. CS theory requires a suitable dictionary that sparsely represents the signal. To date in dMRI there are two kinds of Dictionary Learning (DL) methods: 1) discrete representation based DL (DR-DL), and 2) continuous representation based DL (CR-DL). Due to the discretization in q-space, DR-DL suffers from the numerical errors in interpolation and regridding. By considering a continuous representation using Spherical Polar Fourier (SPF) basis, this work proposes a novel CR-DL based Spherical Polar Fourier Imaging, called DL-SPFI, to recover the diffusion signal as well as the Ensemble Average Propagator (EAP) in continuous 3D space with closed form. DL-SPFI learns an optimal dictionary from the space of Gaussian diffusion signals. Then the learned dictionary is adaptively applied for different voxels in a weighted LASSO framework to robustly recover the di ffusion signal and the EAP. Compared with the start-of-the-art CR-DL method by Merlet et al. and DRDL by Bilgic et al., DL-SPFI has several advantages. First, the learned dictionary, which is proved to be optimal in the space of Gaussian diffusion signal, can be applied adaptively for different voxels. To our knowledge, this is the first work to learn a voxel-adaptive dictionary. The importance of this will be shown theoretically and empirically in the context of EAP estimation. Second, based on the theoretical analysis of SPF basis, we devise an efficient learning process in a small subspace of SPF coefficients, not directly in q-space as done by Merlet et al.. Third, DL-SPFI also devises different regularization for different atoms in the learned dictionary for robust estimation, by considering the structural prior in the space of signal exemplars. We evaluate DL-SPFI in comparison to L1-norm regularized SPFI (L1-SPFI) with fixed SPF basis, and the DR-DL by Bilgic et al. The experiments on synthetic data and real data demonstrate that the learned dictionary is sparser than SPF basis and yields lower reconstruction error than Bilgic's method, even though only simple synthetic Gaussian signals were used for training in DL-SPFI in contrast to real data used by Bilgic et al.

This work has been published in [27].

6.2.8. Fiber orientation distribution from non-negative sparse recovery

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

We revisit the theory of spherical deconvolution and propose a new fiber orientation distribution (FOD) model that can efficiently reconstruct extremely narrow fiber-crossings from limited number of acquisitions. First, we show how to physically model fiber-orientations as rank-1 tensors. Then, we parameterize the FODs with tensors that are decomposable into non-negative sums of rank-1 tensors and finally, we propose a non-negative sparse recovery scheme to estimate FODs of any tensor order from limited acquisitions. Our method features three important advantages: (1) it estimates non-negative FODs, (2) it estimates the number of fiber-compartments, which need not be predefined and (3) it computes the fiber-directions directly, rendering maxima detection superfluous. We test for various SNRs on synthetic, phantom and real data and find our method accurate and robust to signal-noise: fibers crossing up to 23° are recovered from just 21 acquisitions. This opens new and exciting perspectives in diffusion MRI (dMRI), where our improved characterization of the FOD can be of great help for applications such as tractography.

This work has been published in [29].

6.2.9. A polynomial approach for extracting the extrema of a spherical function and its application in diffusion MRI

Participants: Aurobrata Ghosh, Elias Tsigaridas [PolSys Project-Team, Inria, Paris Rocquencourt, France], Bernard Mourrain [Galaad Project-Team, Inria, Sophia Antipolis, Méditerranée, France], Rachid Deriche.

This work has been partially supported by the ANR project NucleiPark and the France-Parkinson Association.

Antipodally symmetric spherical functions play a pivotal role in diffusion MRI in representing sub-voxel-resolution microstructural information of the underlying tissue. This information is described by the geometry of the spherical function. In this work we propose a method to automatically compute all the extrema of a spherical function. We then classify the extrema as maxima, minima and saddle-points to identify the maxima. We take advantage of the fact that a spherical function can be described equivalently in the spherical harmonic (SH) basis, in the symmetric tensor (ST) basis constrained to the sphere, and in the homogeneous polynomial (HP) basis constrained to the sphere. We extract the extrema of the spherical function by computing the stationary points of its constrained HP representation. Instead of using traditional optimization approaches, which are inherently local and require exhaustive search or re-initializations to locate multiple extrema, we use a novel polynomial system solver which analytically brackets all the extrema and refines them numerically, thus missing none and achieving high precision.

To illustrate our approach we consider the Orientation Distribution Function (ODF). In diffusion MRI the ODF is a spherical function which represents a state-of-the-art reconstruction algorithm whose maxima are aligned with the dominant fiber bundles. It is, therefore, vital to correctly compute these maxima to detect the fiber bundle directions. To demonstrate the potential of the proposed polynomial approach we compute the extrema of the ODF to extract all its maxima. This polynomial approach is, however, not dependent on the ODF and the framework presented in this work can be applied to any spherical function described in either the SH basis, ST basis or the HP basis.

This work has been published in [19].

6.2.10. ODF maxima computation using hill climbing algorithm

Participants: Thinhinane Megherbi [USTHB, Algeria], Makhlouf Laouchedi [EMP, Algeria], Houssem Khabatti [EMP, Algeria], Linda Oulebsir-Boumghar [USTHB, Algeria], Ishak Serrat [EMP, Algeria], Vincent Perlbarg [LIF, UMRS 678, INSERM, UPMC - Paris 6], Rachid Deriche.

Diffusion MRI (dMRI) is the only technique to probe in-vivo and non-invasively fiber structure of white matter. Diffusion was first modeled using the classical Second Order Diffusion Tensor model. However, this model is limited in regions of multiple fiber crossings and this has motivated the development of many approaches to extract crossing fibers. Methods like Diffusion Spectrum Imaging (DSI), High Angular Resolution Diffusion Imaging (HARDI) and the High Order Tensor techniques have been proposed to reconstruct specific functions like the Orientation Distribution Function (ODF) whose maxima do correspond to the directions of the multiple fibers.

In this work, we are interested to extract all the crossing fibers characterized as the maxima of the Orientation Distribution Function (ODF). A Hill Climbing algorithm based approach has been developed and implemented to efficiently and accurately extract all the fibers. Promising experimental results obtained with synthetic and real data illustrate the potential of the technique.

This work has been submitted to ISBI'2014 and accepted for presentation and publication.

6.2.11. On SHORE and SPF bases

Participants: Elodie Pozzi, Gonzalo Sanguinetti, Rachid Deriche.

The 3D Simple Harmonic Oscillation Reconstruction and Estimation (SHORE) basis and the Spherical Polar Fourier (SPF) basis were introduced recently to represent the dMRI signal in the full 3D Q-space. SPF presents some continuity problems at the origin which led to our development of the modified SPF basis we introduced to overcome this issue. These bases can be written with radial and angular functions. The radial part of the decomposition is a family of orthogonal functions (the Gauss-Laguerre functions) and the angular component are the spherical harmonic functions. Even though they look similar, they have different properties. The first objective of this work has been to analyse and clarify the differences between those bases. This has been accomplished by describing the spanned spaces. The second goal has been to classify the bases according to their continuity and differentiability and thus draw a more focused comparison between.

This on-going work will be submitted to a journal.

6.3. From DW-MRI to Fiber Pathways and Microstructures Recovery

6.3.1. Mapping Average axon diameters under long diffusion time

Participants: Gonzalo Sanguinetti, Rachid Deriche.

This work proposes an original method to recover axon diameter distribution (ADD) parameters using nuclear magnetic resonance. White matter (WM) is modelled as a bi-compartmental medium composed of an intra axonal space where the diffusion is restricted and an extra axonal space where diffusion is hindered. Under the assumption of long diffusion time, we provide a novel and efficient model for the component of the signal due to the restricted part. This technique might be interpreted as an interesting simplification of the AxCaliber framework, which leads to a simpler model and an extremely faster acquisition protocol. To test and validate our method, we use the open-source toolkit Camino for computing Monte-Carlo simulations of NMR data and model the WM as 3D cubic environments, formed by parallel cylinders with gamma distributed radii. Promising experimental results illustrate the potential of the proposed method.

This work has been submitted to ISBI'2014 and accepted for presentation and publication.

6.3.2. NMR characterization of cylinder radii distribution using a SHORE-based regularization method.

Participants: Gonzalo Sanguinetti, Daniel Alexander [Centre for Medical Image Computing, Dept. Computer Science, UCL], Matt Hall [Centre for Medical Image Computing, Dept. Computer Science, UCL], Rachid Deriche.

In this work, we extend the framework presented by Ozarslan et al [79] by adding a regularization term for better measuring the moments of a cylinder radii distribution by means of NMR acquisitions. The added value of the regularization term is tested and validated using Monte Carlo simulations of NMR signals from complex white matter-like environment. The open source toolkit CAMINO [50] is used for computing the simulations and an excellent agreement is obtained between the ground truth and the estimated moments.

This work has been submitted to ISMRM'2014.

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

Participants: Emmanuel Caruyer [SBIA, University of Pennsylvania Medical School, USA], Sylvain Merlet, Rachid Deriche.

In diffusion MRI, a technique known as diffusion spectrum imaging reconstructs the propagator with a discrete Fourier transform, from a Cartesian sampling of the diffusion signal. Alternatively, it is possible to directly reconstruct the orientation distribution function in q-ball imaging, providing so-called high angular resolution diffusion imaging. In between these two techniques, acquisitions on several spheres in q-space offer an interesting trade-off between the angular resolution and the radial information gathered in diffusion MRI. A careful design is central in the success of multishell acquisition and reconstruction techniques and the design of acquisition in multishell is still an open and active field of research.

In this work, we propose a novel method to design sampling schemes with optimal angular coverage and show the positive impact on angular resolution in diffusion MRI. Our method is based on a generalization of electrostatic repulsion to multishell and allows to design multishell acquisition with uniform angular coverage.

We evaluated the impact of our method using simulations, on the angular resolution in one and two bundles of fiber configurations. Compared to more commonly used radial sampling, we show that our method improves the angular resolution, as well as fiber crossing discrimination.

This work has been published in [16].

6.3.4. Choosing tractography parameters to improve connectivity mapping

Participants: Gabriel Girard [SCIL Lab., Sherbrooke University], Kevin Whittingstall [SCIL Lab., Sherbrooke University], Kevin Whittingstall [SCIL Lab., Sherbrooke University], Rachid Deriche.

Diffusion-weighted imaging (DWI) is often used as a starting point for in vivo white matter (WM) connectivity to reconstruct potential WM pathways between brain areas. Tractography algorithms have many parameters which can influence reconstruction and connectivity. Various choices of parameters have been proposed. But how does one choose the best set of parameters ? In this study, we varied three critical parameters while monitoring connectivity score using the Tractometer evaluation system on the International Symposium on Biomedical Imaging (ISBI) Challenge synthetic dataset. The three parameters were: The maximum deviation angle between two consecutive tractography steps (this addresses the hypothesis of smoothness of the WM pathways), the spherical function (SF) threshold (this aims at removing noisy propagation directions during the tractography process) and the initial SF threshold (this aims at removing initial noise at the seeds and to start tractography in a good tangent direction to the WM bundle).

This work has been submitted to ISMRM'2014.

6.3.5. Improved tractography using structural priors

Participants: Gabriel Girard [SCIL Lab., Sherbrooke University], Maxime Descoteaux [SCIL Lab., Sherbrooke University], Kevin Whittingstall [SCIL Lab., Sherbrooke University], Rachid Deriche.

In this work, we propose better tractography parameters in term of global connectivity and a novel tractography stopping criterion based on partial volume estimation maps, calculated from a T1-weighted image. We also propose a particle filtering method using anatomical information as prior for tractography to enforce streamlines connecting gray matter regions and reducing the proportion of erroneous streamlines. Results show streamlines more uniformly distributed among long and short, and small and large white matter bundles. This provides connectivity estimation not underestimated for bundles having higher complexity. Quantitative analysis is done on synthetic datasets and qualitative results are shown on real data. The proposed method takes advantage of prior information on the brain to change the dMRI-based tracking direction and help providing streamlines that can quantify the brain structure.

This on-going work will be submitted to NeuroImage.

6.3.6. From diffusion MRI to brain connectomics

Participants: Aurobrata Ghosh, Rachid Deriche.

Diffusion MRI (dMRI) is a unique modality of MRI which allows one to indirectly examine the microstructure and integrity of the cerebral white matter in vivo and non-invasively. Its success lies in its capacity to reconstruct the axonal connectivity of the neurons, albeit at a coarser resolution, without having to operate on the patient, which can cause radical alterations to the patient's cognition. Thus dMRI is beginning to assume a central role in studying and diagnosing important pathologies of the cerebral white matter, such as Alzheimer's and Parkinson's diseases, as well as in studying its physical structure in vivo. In this work, we present an overview of the mathematical tools that form the framework of dMRI – from modelling the MRI signal and measuring diffusion properties, to reconstructing the axonal connectivity of the cerebral white matter, i.e., from Diffusion Weighted Images (DWIs) to the human connectome.

This work has been published in [38].

6.4. Forward and Inverse Problems in MEEG

6.4.1. Source localization using rational approximation on plane sections

Participants: Maureen Clerc [Athena Project-Team, Inria, Sophia Antipolis, Méditerranée, France], Théodore Papadopoulo [Athena Project-Team, Inria, Sophia Antipolis, Méditerranée, France], Juliette Leblond [Apics Project-Team, Inria, Sophia Antipolis, Méditerranée, France], Jean-Paul Marmorat [CMA, Ecole des Mines Paristech, Sophia Antipolis, France].

In functional neuroimaging, a crucial problem is to localize active sources within the brain non-invasively, from knowledge of electromagnetic measurements outside the head. Identification of point sources from boundary measurements is an ill-posed inverse problem. In the case of electroencephalography (EEG), measurements are only available at electrode positions, the number of sources is not known in advance and the medium within the head is inhomogeneous. We have presented [49] a new method for EEG source localization, based on rational approximation techniques in the complex plane. The method is used in the context of a nested sphere head model, in combination with a cortical mapping procedure. Results on simulated data prove the applicability of the method in the context of realistic measurement configurations. In the continuation of this work, we are in discussion with an industrial partner (BESA, Munich) for a scientific partnership.

6.4.2. Dictionary learning for multitrial datasets

Participants: Maureen Clerc, Sebastian Hitziger, Théodore Papadopoulo.

Following the path opened with the Consensus matching Pursuit method (CMP) [46], we continue our endeavour to avoid signal averaging using directly the raw signal with the assumption that events of interest are those that repeat in each trial [36]. Towards such a goal, and to improve the simple dictionary used in CMP, we have adapted dictionary learning methods to multitrial bio-electric signals, by explicitly implementing jitter invariance [30]. This allows for a much more detailed data-driven description of events. For example, using local field potential signals of chemically induced spikes (in a rat model), we have been able to distinguish several spike shapes which show some coherence in time. The method has been recently extended to detect spike events in continuous signals (i.e. not organized in epochs). While it requires a good signal to noise ratio, the method is very general and has also been used for various other signal types (see section 6.5).

6.5. Coupling functional and structural models

6.5.1. Cortex parcellation via diffusion data as prior knowledge for the MEG inverse problem

Participants: Anne-Charlotte Philippe, Maureen Clerc, Théodore Papadopoulo, Rachid Deriche.

In this work, a new approach is presented for the recovery of dipole magnitudes in a distributed source model for magnetoencephalographic (MEG) imaging. This method consists in introducing prior knowledge regarding the anatomical connectivity in the brain to this ill-posed inverse problem. Towards this goal, a cortex parcellation is performed using structural information coming from diffusion MRI (dMRI), the only non-invasive modality allowing to have access to the structure of the WM tissues. Then, sources in the same diffusion parcel are constrained, in the MEG inverse problem, to have close magnitude values. Results of our method on MEG simulations are presented and favorably compared with classical source reconstruction methods.

This work has been published in [32], and is part of A.C. Philippe's Ph.D thesis [12].

6.5.2. Diffusion-Weighted Imaging tractography-based parcellation of the human cortex as regularization term for the MEG inverse problem

Participants: Anne-Charlotte Philippe, Maureen Clerc, Théodore Papadopoulo, Rachid Deriche.

The purpose of this work is to advocate the use of structural connectivity information to regularize the ill-posed MEG inverse problem. Diffusion MRI being the only non invasive modality allowing to have access to the connectivity profile of cortical sources, the proposed method called Diff-MNE consists in the introduction of a cortex parcellation based on diffusion data regularization term to the MEG inverse problem. Our method is tested on synthetic and real human brain data and compared to the classical minimum-norm method. Results show that a diffusion-based cortex parcellation as a regularization term for the MEG inversion process improves the source reconstruction. This proves the interest of merging diffusion MRI and MEG data.

This work is under submission to a Neuroimage and is part of A.C. Philippe's Ph.D thesis [12]

6.5.3. Propagation of epileptic spikes revealed by diffusion-based constrained MEG source reconstruction

Participants: Anne-Charlotte Philippe, Maureen Clerc, Théodore Papadopoulo, Rachid Deriche.

In this work, we study the propagation of an epileptic spike (from single event data). As in the two previous sections, a cortex parcellation is performed using structural information coming from diffusion MRI Then, a MEG inverse problem is defined on a parcellated source space which imposes constant activity on each parcel. This inverse problem is applied separately for measurements obtained in a given time range. The most active parcels over the time range are located and their time course are displayed. This allowed the study of the propagation of an epileptic spike via those active parcels. Results on real data shows varying spatial propagations of an epileptic spike for the same subject.

This work has been published in [40], and is part of A.C. Philippe's Ph.D thesis [12].

6.6. Brain Computer Interfaces

6.6.1. Combining ERD and ERS features to create a system-paced BCI

Participants: Maureen Clerc, Joan Fruitet, Théodore Papadopoulo, Eoin Thomas.

An important factor in the usability of a brain computer interface (BCI) is the setup and calibration time required for the interface to function accurately. Recently, brain-switches based on the rebound following motor imagery of a single limb effector have been investigated as basic BCIs due to their good performance with limited electrodes, and brief training session requirements. In this work, a BCI is proposed which expands the methodology of brain-switches to design an interface composed of multiple brain-buttons. The algorithm is designed as a system paced interface which can recognise 2 intentional-control tasks and a no-control state based on the activity during and following motor imagery in only 3 electroencephalogram channels. An online experiment was performed over 6 subjects to validate the algorithm, and the results show that a working BCI can be trained from a single calibration session and that the post motor imagery features are both informative and robust over multiple sessions.

This work has been published in [24].

6.6.2. An analysis of performance evaluation for motor-imagery based BCI

Participants: Maureen Clerc, Matthew Dyson [Laboratoire de Neurosciences Cognitives, Marseille], Eoin Thomas.

In recent years, numerous brain–computer interfaces (BCIs) based on motor-imagery have been proposed which incorporate features such as adaptive classification, error detection and correction, fusion with auxiliary signals and shared control capabilities. Due to the added complexity of such algorithms, the evaluation strategy and metrics used for analysis must be carefully chosen to accurately represent the performance of the BCI. In this work, metrics are reviewed and contrasted using both simulated examples and experimental data. Furthermore, a review of the recent literature is presented to determine how BCIs are evaluated, in particular, focusing on the relationship between how the data are used relative to the BCI subcomponent under investigation. From the analysis performed in this study, valuable guidelines are presented regarding the choice of metrics and evaluation strategy dependent upon any chosen BCI paradigm.

This work has been published in [23].

6.6.3. Bandit algorithms for faster task selection in BCI

Participants: Maureen Clerc, Aurélien Emmanuel, Joan Fruitet [former Athena PhD student], Alexandra Carpentier [Sequel Project-Team, Inria Lille], Rémi Munos [Sequel Project-Team, Inria Lille].

Brain–computer interfaces (BCIs) based on sensorimotor rhythms use a variety of motor tasks, such as imagining moving the right or left hand, the feet or the tongue. Finding the tasks that yield best performance, specifically to each user, is a time-consuming preliminary phase to a BCI experiment. This study presents a new adaptive procedure to automatically select (online) the most promising motor task for an asynchronous brain-controlled button. We have developed for this purpose an adaptive algorithm UCB-classif based on the stochastic bandit theory and design an EEG experiment to test our method. We compare (offline) the adaptive algorithm to a naïve selection strategy which uses uniformly distributed samples from each task. We also run the adaptive algorithm online to fully validate the approach. By not wasting time on inefficient tasks, and focusing on the most promising ones, this algorithm results in a faster task selection and a more efficient use of the BCI training session. More precisely, the offline analysis reveals that the use of this algorithm can reduce the time needed to select the most appropriate task by almost half without loss in precision, or alternatively, allow us to investigate twice the number of tasks within a similar time span. Online tests confirm that the method leads to an optimal task selection. This study is the first one to optimize the task selection phase by an adaptive procedure. By increasing the number of tasks that can be tested in a given time span, the proposed method could contribute to reducing 'BCI illiteracy'.

This work has been published in [17].

6.6.4. Enhancing visuospatial attention performance with brain-computer interfaces

Participants: Thomas Brochier [Institut des Neurosciences de La Timone, Marseille], Maureen Clerc, Romain Trachel.

Brain-Computer Interfaces (BCI) can provide innovative solutions beyond the medical domain. In human research, visuospatial attention is often assessed from shifts in head or gaze orientation. However in some critical situations, these behavioral features can be dissociated from covert attention processes and brain features may indicate more reliably the spatial focus of attention. In this context, we investigate whether EEG signals could be used to enhance the behavioral performance of human subjects in a visuospatial attention task. Our results demonstrate that a BCI protocol based on adaptive or warning displays can be developed to shorten the reaction time and improve the accuracy of responses to complex visual targets. We performed offline and online tests demonstrating the validity of this type of approach.

This work was presented at conferences in the HCI community [35] and in the Neural Engineering community [34].

6.6.5. Verbal communication through brain computer interfaces

Participants: Maureen Clerc, Dieter Devlaminck, Claude Desnuelle [CHU de Nice l'Archet], Violaine Guy [CHU de Nice l'Archet], Manu Maby [Centre de Recherche Neurologique de Lyon], Jérémie Mattout [Centre de Recherche Neurologique de Lyon], Théodore Papadopoulo.

Brain Computer Interfaces (BCI) provide a way of communicating directly from brain activity, bypassing muscular control. We report some recent advances in a BCI communication system called the P300 speller, which is a virtual brain-operated keyboard. This system relies on electroencephalographic activity time-locked to the flashing of the desired letters. It requires calibration of the system, but very little training from the user. Clinical tests are being conducted on a target population of patients suffering from Amyotrophic Lateral Sclerosis, in order to confirm the usability of the P300 speller for reliable communication.

This work has been published in [26]. It is also the object of an intensive clinical study on 20 patients which we are currently conducting at Nice University Hospital.

ATLANMOD Project-Team

6. New Results

6.1. Reverse Engineering

Model Driven Reverse Engineering (MDRE), and its applications such as software modernization, is a discipline in which model-driven development (MDD) techniques are used to treat legacy systems. During this year, Atlanmod has continued working actively on this research area. The main contributions are the following:

- In the context of the ARTIST FP7 project, the work has started on reusing (and extending accordingly) MoDisco and several of its components to provide the Reverse Engineering support required within the project. More particularly, the MoDisco Model Discovery + Model Understanding two-step approach is being promoted as an important part of the ARTIST migration methodology and process [35] [19]. Work has also been performed, in the context of the TEAP FUI project dealing with Enterprise Architecture, on how to design and implement a model driven federation approach from heterogeneous data sources (e.g. Excel files, databases, etc.) directly inspiring from these same MoDisco principles [20].
- In order to react to the ever-changing market, every organization needs to periodically reevaluate and evolve its company policies. These policies must be enforced by its Information System (IS) by means of a set of so-called business rules that drive the system behavior and data. Clearly, policies and rules must be aligned at all times but unfortunately this is a challenging task. In most ISs, the implementation of business rules is scattered among the different components of the system, therefore appropriate techniques must be provided for the discovery and evolution of changing business rules. In [39], [25], [26], we describe a MDRE framework and tool aiming at extracting business rules out of COBOL source code. In [27], we describe a Model-based process and tool to extract business rules, expressed as OCL integrity constraints, from relational databases. In these works, the use of modeling techniques facilitate the representation of the rules at a higher-abstraction level which enables stakeholders to understand and manipulate them more easily. A thesis financed by IBM to advance the research on this topic has been completed this year
- In a web context, JSON has become a very popular lightweigth format for data exchange. JSON is human readable and easy for computers to parse and use. However, JSON is schemaless. Though this brings some benefits (e.g. flexibility in the representation of the data) it can become a problem when consuming and integrating data from different JSON services since developers need to be aware of the structure of the schemaless data. We believe that a mechanism to discover (and visualize) the implicit schema of the JSON data would largely facilitate the creation and usage of JSON services. For instance, this would help developers to understand the links between a set of services belonging to the same domain or API. In this sense, we have proposed a model-based approach to generate the underlying schema of a set of JSON documents [22].

6.2. Security

Most companies information systems are composed by heterogeneous components responsible of hosting, creating or manipulating critical information for the day-to-day operation of the company. Securing this information is therefore one of their main concerns, more particularly specifying Access Control (AC) policies. However, the task of implementing an AC security policy (sometimes relying on several mechanisms) remains complex and error prone as it requires knowing low level and vendor-specific facilities. In this context, discovering and understanding which security policies are actually being enforced by the Information System (IS) becomes critical. Thus, the main challenge consists in bridging the gap between the vendor-dependent security features and a higher-level representation. This representation has to express the policies by abstracting from the specificities of the system components, allowing security experts to better understand the policy and to implement all related evolution, refactoring and manipulation operations in a reusable way.

In 2013, we have tackled the aforementioned problems with respect to three key information system components: networks of firewalls, relational database systems and content management systems.

- Firewalls are a key element in network security. They are in charge of filtering the traffic of the network in compliance with a number of access-control rules that enforce a given security policy. In [33] we have described a model-driven reverse engineering approach able to extract the security policy implemented by a set of firewalls in a working network, easing the understanding, analysis and evolution of network security policies. In [17] we have extended this method to cope with a more complex and specific scenario, i.e, the management of stateful packet filtering.
- A similar approach have been successfully used to extract AC information from relational database systems. Concretely, in [32] we contribute a security metamodel and a reverse engineering process that combines standard database access-control rules with the fine-grained access control provided by triggers and stored procedures. The extraction of this comprehensive model helps security experts to visualize and manipulate database security policies in a vendor-independent manner.
- Out-of-the-box Web Content Management Systems (WCMSs) are the tool of choice for the development of millions of enterprise web sites. However, little attention has been brought to the analysis of how developers use the content protection mechanisms provided by WCMSs, in particular, Access-control (AC). We have proposed in [34] a metamodel tailored to the representation of WCMS AC policies, easing the analysis and manipulation tasks by abstracting from vendor-specific details.

6.3. Collaborative development

In the field of Domain-Specific Languages (DSLs), we have focused on the improvement of the DSLs definition process. When developing DSMLs, the participation of end-users is normally limited to providing domain knowledge and testing the resulting language prototypes. Language developers, which are perhaps not domain experts, are therefore in control of the language development and evolution. This may cause misinterpretations which hamper the development process and the quality of the DSML. Thus, it would be beneficial to promote a more active participation of end-users in the development process of DSMLs. While current DSML workbenches are mono-user and designed for technical experts, we have presented a process and tool support for the example-driven, collaborative construction of DSMLs based on Collaboro in order to engage end-users in the creation of their own languages [23], [24].

6.4. MDE Scalability

As Model-Driven Engineering (MDE) is increasingly applied to larger and more complex systems, additional research and development is imperative in order to enable MDE to remain relevant with industrial practice. In [31] we attempt to provide a research roadmap for scalability in MDE and outline directions for work in this emerging research area. As a first result in this roadmap, in [37] we show that rule-based languages like ATL have strong parallelization properties. Parallelization is indeed one of the traditional ways of making computation systems scalable. We describe the implementation of a parallel transformation engine for the current version of the ATL language and experimentally evaluate the consequent gain in scalability. Finally in [28] we compare the improved scalability of the ATL transformation engine with other engines in the community by addressing the task of generating and analyzing very large flow graphs.

6.5. Model Quality

Our work aims to enhance the quality of the modeling activity in the context of software engineering and language engineering. This year, this has translated in the following results:

- A benchmark that facilitates the comparison between the plethora of tools that provide some kind of quality assurance for models. Similarly to what it is done in many other domains, a common set of test benchmarks that new tools can rely on to experiment and evaluate themselves could speed up the advance in the field. Our proposal can be found [30]

- Validation of the feasibility to apply this kind of techniques in industrial settings based on two case studies [12] and [36]
- Advanced on the verification of model transformations using SMT solvers (instead of SAT or CSP-based approaches commonly used before), with some encouraging results [21] and, related to this, [13]
- A method to build models using instance-level information in terms of examples and counterexamples (gathering requirements using these instance scenarios is usually better from a stakeholder's point of view than trying to explain us general rules about the business). So far existing approaches have often focused on the generation of static models from such instance-level information but have omitted the inference of OCL business rules that could complement the static models and improve the precision of the software specification. We propose an approach to automating such inference [29]. The basic idea is based on an incorporation of the problem solving mechanism and getting user feedback: Candidates are generated by a problem solving, and irrelevant ones are eliminated using the user feedback on generated counterexamples and examples. Our approach is realized with the support tool InferOCL and has been applied on several user cases, indicating a possibility to apply this solution prototype in practice.

AVALON Team

6. New Results

6.1. Energy efficiency of large scale distributed systems

Participants: Ghislain Landry Tsafack Chetsa, Mohammed El Mehdi Diouri, Jean-Patrick Gelas, Olivier Glück, Laurent Lefèvre, François Rossigneux.

6.1.1. Analysis and Evaluation of Different External and Internal Power Monitoring Devices for a Server and a Desktop Machine

Large-scale distributed systems (e.g., datacenters, HPC systems, clouds, large-scale networks, etc.) consume and will consume enormous amounts of energy. Therefore, accurately monitoring the power and energy consumption of these systems is increasingly more unavoidable. The main novelty of this contribution [15] is the analysis and evaluation of different external and internal power monitoring devices tested using two different computing systems, a server and a desktop machine. Furthermore, we also provide experimental results for a variety of benchmarks which exercise intensively the main components (CPU, Memory, HDDs, and NICs) of the target platforms to validate the accuracy of the equipment in terms of power dispersion and energy consumption. We highlight that external wattmeters do not offer the same measures as internal wattmeters. Thanks to the high sampling rate and to the different measured lines, the internal wattmeters allow an improved visualization of some power fluctuations. However, a high sampling rate is not always necessary to understand the evolution of the power consumption during the execution of a benchmark.

6.1.2. Your Cluster is not Power Homogeneous

Future supercomputers will consume enormous amounts of energy. These very large scale systems will gather many homogeneous clusters. We analyze the power consumption of the nodes from different homogeneous clusters during different workloads. As expected, we observe that these nodes exhibit the same level of performance. However, we also show that different nodes from a homogeneous cluster may exhibit heterogeneous idle power energy consumption even if they are made of identical hardware. Hence, we propose an experimental methodology to understand such differences. We show that CPUs are responsible for such heterogeneity which can reach 20% in terms of energy consumption. So energy aware (Green) schedulers must take care of such hidden heterogeneity in order to propose efficient mapping of tasks. To consume less energy, we propose an energy-aware scheduling approach taking into account the heterogeneous idle power consumption of homogeneous nodes [20]. It shows that we are able to save energy up to 17% while exploiting the high power heterogeneity that may exist in some homogeneous clusters.

6.1.3. Energy Consumption Estimations of Fault Tolerance protocols

Energy consumption and fault tolerance are two interrelated issues to address for designing future exascale systems. Fault tolerance protocols used for checkpointing have different energy consumption depending on parameters like application features, number of processes in the execution and platform characteristics. Currently, the only way to select a protocol for a given execution is to run the application and monitor the energy consumption of different fault tolerance protocols. This is needed for any variation of the execution setting. To avoid this time and energy consuming process, we propose an energy estimation framework [16], [17], [7]. It relies on an energy calibration of the considered platform and a user description of the execution setting. We evaluate the accuracy of our estimations with real applications running on a real platform with energy consumption monitoring. Results show that our estimations are highly accurate and allow selecting the best fault tolerant protocol without pre-executing the application.

6.1.4. Energy Consumption Estimations of Data Broadcasting

Future supercomputers will gather hundreds of millions of communicating cores. The movement of data in such systems will be very energy consuming. We address the issue of energy consumption of data broadcasting in such large scale systems. To this end, in [19], [7], we propose a framework to estimate the energy consumed by different MPI broadcasting algorithms for various execution settings. Validation results show that our estimations are highly accurate and allow to select the least consuming broadcasting algorithm.

6.1.5. A Smart-Grid Based Framework for Consuming Less and Better in Extreme-Scale Infrastructures

As they will gather hundreds of million cores, future exascale supercomputers will consume enormous amounts of energy. Besides being very important, their power consumption will be dynamic and irregular. Thus, in order to consume energy efficiently, powering such systems will require a permanent negotiation between the energy supplier and one of its major customers represented by exascale platforms. We have designed SESAMES [18], [53], a smart and energy-aware service-oriented architecture manager that proposes energy-efficient services for exascale applications and provides an optimized reservation scheduling. The new features of this framework are the design of a smart grid and a multi-criteria green job scheduler. Simulation results show that with the proposed multi-criteria job scheduler, we are able to save up to 2.32 % in terms of energy consumption, 24.22 % in terms of financial cost and reduce up to 7.12 % the emissions of CO_2.

6.1.6. Clustered Virtual Home Gateway (vHGW)

This result is a joint work between Avalon team (J.P. Gelas, L. Lefevre) and Addis Abeba University (M. Tsibie and T. Assefa). The customer premises equipment (CPE), which provides the interworking functions between the access network and the home network, consumes more than 80% of the total power in a wireline access network. In the GreenTouch initiative (cf Section 7.3), we aim at a drastic reduction of the power consumption by means of a passive or quasi-passive CPE. Such approach requires that typical home gateway functions, such as routing, security, and home network management, are moved to a virtual home gateway (vHGW) server in the network. In our first prototype virtual home gateways of the subscribers were put in LXC containers on a unique GNU/Linux server. The container approach is more scalable than separating subscribers by virtual machines. We demonstrated a sharing factor of 500 to 1000 virtual home gateways on one server, which consumes about 150 W, or 150 to 300 mW per subscriber. Comparing this power consumption with the power of about 2 W for the processor in a thick client home gateway, we achieved an efficiency gain of 5-10x. The prototype was integrated and demonstrated at TIA 2012 in Dallas. In our current work, we propose the Clustered vHGWs Data center architecture to yield optimal energy conservation through virtual machine's migration among physical nodes based on the current subscriber's service access state, while ensuring SLA respective subscribers. Thus, optimized energy utilization of the data center is assured without compromising the availability of service connectivity and QoS preferences of respective subscribers.

6.1.7. Improving Energy Efficiency of Large Scale Systems without a priori Knowledge of Applications and Services

Unlike their hardware counterpart, software solutions to the energy reduction problem in large scale and distributed infrastructures hardly result in real deployments. At the one hand, this can be justified by the fact that they are application oriented. At the other hand, their failure can be attributed to their complex nature which often requires vast technical knowledge behind proposed solutions and/or thorough understanding of applications at hand. This restricts their use to a limited number of experts, because users usually lack adequate skills. In addition, although subsystems including the memory and the storage are becoming more and more power hungry, current software energy reduction techniques fail to take them into account. We propose a methodology for reducing the energy consumption of large scale and distributed infrastructures. Broken into three steps known as (i) phase identification, (ii) phase characterization, and (iii) phase identification and system reconfiguration; our methodology abstracts away from any individual applications as it focuses on the infrastructure, which it analyses the runtime behaviour and takes reconfiguration decisions accordingly.

The proposed methodology is implemented and evaluated in high performance computing (HPC) clusters of varied sizes through a Multi-Resource Energy Efficient Framework (MREEF). MREEF implements the proposed energy reduction methodology so as to leave users with the choice of implementing their own system reconfiguration decisions depending on their needs. Experimental results show that our methodology reduces the energy consumption of the overall infrastructure of up to 24% with less than 7% performance degradation. By taking into account all subsystems, our experiments demonstrate that the energy reduction problem in large scale and distributed infrastructures can benefit from more than "the traditional" processor frequency scaling. Experiments in clusters of varied sizes demonstrate that MREEF and therefore our methodology can easily be extended to a large number of energy aware clusters. The extension of MREEF to virtualized environments like cloud shows that the proposed methodology goes beyond HPC systems and can be used in many other computing environments.

6.1.8. Reservation based Usage for Energy Efficient Clouds: the Climate Architecture

The FSN XLcloud project (cf Section 7.1) strives to establish the demonstration of a High Performance Cloud Computing (HPCC) platform based on OpenStask, that is designed to run a representative set of compute intensive workloads, including more specifically interactive games, interactive simulations and 3D graphics. XLcloud is based on OpenStack, and Avalon is contributing to the energy efficiency part of this project. We have proposed and brought our contribution to Climate, a new resource reservation framework for OpenStack, developed in collaboration with Bull, Mirantis and other OpenStack contributors. Climate allows the reservation of both physical and virtual resources, in order to provide a mono-tenancy environment suitable for HPC applications. Climate chooses the most efficient hosts (flop/W). This metric is computed from the CPU / GPU informations, mixed with real power consumption measurements provided by the Kwapi framework. The user requirements may be loose, allowing Climate to choose the best time slot to place the reservation. Climate will be improved with standby mode features, to shut down automatically the unused hosts. The first release of Climate is planned at the end of January 2014, and we expect an incubation in the next version of OpenStack.

6.2. Simulation of Large Scale Distributed Systems

Participants: Frédéric Desprez, Jonathan Rouzaud-Cornabas, Frédéric Suter.

6.2.1. Toward Better Simulation of MPI Applications on Ethernet/TCP Networks

Simulation and modeling for performance prediction and profiling is essential for developing and maintaining HPC code that is expected to scale for next-generation exascale systems, and correctly modeling network behavior is essential for creating realistic simulations. In [11], we proposed an implementation of a flow-based hybrid network model that accounts for factors such as network topology and contention, which are commonly ignored by other approaches. We focused on large-scale, Ethernet-connected systems, as these currently compose 37.8% of the TOP500 index, and this share is expected to increase as higher-speed 10 and 100GbE become more available. The European Mont-Blanc project that studies exascale computing by developing prototype systems with low-power embedded devices will also use Ethernet-based interconnect. Our model is implemented within SMPI, an open-source MPI implementation that connects real applications to the SimGrid simulation framework (cf Section 5.5). SMPI provides implementations of collective communications based on current versions of both OpenMPI and MPICH. SMPI and SimGrid also provide methods for easing the simulation of large-scale systems, including shadow execution, memory folding, and support for both online and offline simulation. We validated our proposed model by comparing traces produced by SMPI with those from real world experiments, as well as with those obtained using other established network models. Our study shows that SMPI has a consistently better predictive power than classical LogP-based models for a wide range of scenarios including both established HPC benchmarks and real applications.

6.2.2. SimGrid: a Sustained Effort for the Versatile Simulation of Large Scale Distributed Systems

SIMGRID (cf Section 5.5) is a toolkit for the versatile simulation of large scale distributed systems, whose development effort has been sustained for the last fifteen years. Over this time period SIMGRID has evolved from a one-laboratory project in the U.S. into a scientific instrument developed by an international collaboration. The keys to making this evolution possible have been securing of funding, improving the quality of the software, and increasing the user base. We detailed in [55]how we have been able to make advances on all three fronts, on which we plan to intensify our efforts over the upcoming years.

6.2.3. Simulating Multiple Clouds from a Client Point of View: SGCB an AWS Simulator

Validating a new application over a Cloud is not an easy task and it can be costly over public Clouds. Simulation is a good solution if the simulator is accurate enough and if it provides all the features of the target Cloud. In [49], we have proposed an extension of the SIMGRID simulation toolkit to simulate the Amazon IaaS Cloud. Based on an extensive study of the Amazon platform and previous evaluations, we have integrated models into the SIMGRID Cloud Broker and exposed the same API as Amazon to the users. Our experimental results have shown that our simulator is able to simulate different parts of Amazon for different applications.

6.3. Active Data: A Data-Centric Approach to Data Life-Cycle Management

Participants: Gilles Fedak, Anthony Simonet.

Data-intensive science offers new opportunities for innovation and discoveries, provided that large datasets can be handled efficiently. Data management for data-intensive science applications is challenging; requiring support for complex data life cycles, coordination across multiple sites, fault tolerance, and scalability to support tens of sites and petabytes of data. In [28], we argue that data management for data-intensive science applications requires a fundamentally different management approach than the current ad-hoc task centric approach. We propose Active Data, a fundamentally novel paradigm for data life cycle management. Active Data follows two principles: data-centric and event-driven. We report on the Active Data programming model and its preliminary implementation, and discuss the benefits and limitations of the approach on recognized challenging data-intensive science use-cases.

6.4. HPC Component Model

Participants: Zhengxiong Hou, Vincent Lanore, Christian Perez.

6.4.1. Auto-tuning of Stencil Based Applications

We have finished designing a tuning approach for stencil applications on multi-core clusters [25]. We focused in particular on a 2D Jacobi benchmark application as well as memory bandwidth performance. The tuning approach includes data partitioning within one node, the selection of the number of threads within a multi-core node, a data partitioning for multi nodes, and the number of nodes for a multi-core cluster. This model is based on a set of experiments on machines of GRID'5000 and on the Curie supercomputer.

6.4.2. Static 2D FFT Adaptation through a Component Model based on Charm++

Adaptation algorithms for HPC applications can improve performance but their implementation is often costly in terms of development and maintenance. Component models such as Gluon++, which is built on top of Charm++, propose to separate the business code, encapsulated in components, and the application structure, expressed through a component assembly. Adaptation of component-based HPC applications can be achieved through the optimization of the assembly. We have studied such an approach with the adaptation to network topology and data size of a Gluon++ 2D FFT application. Preliminary experimental results obtained on the GRID'5000 platform show the suitability of the proposed approach.

6.4.3. Towards Scalable Reconfiguration in Component Models

Some HPC applications require reconfiguration of their architecture at runtime; examples include adapting to (cloud) resource elasticity, efficient distributed deployment, Adaptive Mesh Refinement (AMR), and load balancing. This class of applications raises challenges such as handling of concurrent reconfigurations and distributed architecture representation at runtime. To our knowledge, no existing programming model addresses those challenges in the general case with both high programmability and scalability. We have identified a list of specific subproblems and use-cases and we have devised a preliminary component model to address some of them.

6.5. Resource Management and Scheduling

Participants: Eddy Caron, Frédéric Desprez, Gilles Fedak, Jose Luis Lucas, Christian Perez, Jonathan Rouzaud-Cornabas, Frédéric Suter.

6.5.1. Resource Management Architecture for Fair Scheduling of Optional Computations

Most High-Performance Computing platforms require users to submit a pre-determined number of computation requests (also called jobs). Unfortunately, this is cumbersome when some of the computations are optional, i.e., they are not critical, but their completion would improve results. For example, given a deadline, the number of requests to submit for a Monte Carlo experiment is difficult to choose. The more requests are completed, the better the results are, however, submitting too many might overload the platform. Conversely, submitting too few requests may leave resources unused and misses an opportunity to improve the results.

In cooperation with IRIT (Toulouse), we have proposed a generic client-server architecture and an implementation in DIET, a production GridRPC middleware, which auto-tunes the number of requests [12]. Real-life experiments show significant improvement of several metrics, such as user satisfaction, fairness and the number of completed requests. Moreover, the solution is shown to be scalable.

6.5.2. Advanced Promethee-based Scheduler Enriched with User-Oriented Methods

Efficiently scheduling tasks in hybrid Distributed Computing Infrastructures (DCI) is a challenging pursue because the scheduler must deal with a set of parameters that simultaneously characterize the tasks and the hosts originating from different types of infrastructure. In [27], we propose a scheduling method for hybrid DCIs, based on advanced multi-criteria decision methods. The scheduling decisions are made using pairwise comparisons of the tasks for a set of criteria like expected completion time and price charged for computation. The results are obtained with an XtremWeb-like pull-based scheduler simulator using real failure traces for a combination of three types of infrastructure. We also show how such a scheduler should be configured to enhance user satisfaction regardless their profiles, while maintaining good values for makespan and cost. We validate our approach with a statistical analysis on empirical data and show that our proposed scheduling method improves performance by 12-17% compared to other scheduling methods. Experimenting on large time-series and using realistic scheduling scenarios lead us to conclude about time consistency results of the method.

6.5.3. Fair Resource Sharing for Dynamic Scheduling of Workflows on Heterogeneous Systems

Scheduling independent workflows on shared resources in a way that satisfy users Quality of Service is a significant challenge. In [37], we described methodologies for off-line scheduling, where a schedule is generated for a set of known workflows, and on-line scheduling, where users can submit workflows at any moment in time. We consider the on-line scheduling problem in more detail and present performance comparisons of state-of-the-art algorithms for a realistic model of a heterogeneous system.

6.5.4. Image Transfer and Storage Cost Aware Brokering Strategies for Multiple Clouds

Nowadays, Clouds are used for hosting a large range of services. But between different Cloud Service Providers, the pricing model and the price of individual resources are very different. Furthermore hosting a service in one Cloud is the major cause of service outage. To increase resiliency and minimize the monetary cost of running a service, it becomes mandatory to span it between different Clouds. Moreover, due to dynamicity of both the service and Clouds, it could be required to migrate a service at run time. Accordingly, this ability must be integrated into the multi-Cloud resource manager, *i.e.* the Cloud broker. But, when migrating a VM to a new Cloud Service Provider, the VM disk image must be migrated too. Accordingly, data storage and transfer must be taken into account when choosing if and where an application will be migrated.

In [47], we have extended a cost-optimization algorithm to take into account storage costs to approximate the optimal placement of a service. The data storage management consists in taking two decisions: where to upload an image, and keep it on-line during the experiment lifetime or delete it when unused. Based on our experimentations, we have shown that the storage cost of VM disk image must not be neglected as done in previous work. Moreover, we have shown that using the accurate combinations of storage policies can dramatically reduce the storage cost (from 90% to 14% of the total bill).

6.6. Security for Virtualization and Clouds

Participants: Eddy Caron, Arnaud Lefray, Jonathan Rouzaud-Cornabas.

6.6.1. Improving Users' Isolation in IaaS: Virtual Machine Placement with Security Constraints

Nowadays virtualization is used as the sole mechanism to isolate different users on Cloud platforms. Due to improper virtualization of micro-architectural components, data leak and modification can occur on public Clouds. Moreover, using the same attack vector (improper virtualization of micro-architectural components), it is possible to induce performance interferences, *i.e.* noisy neighbors. Using this approach, a VM can slow down and steal resources from concurrent VMs. In [43], we have proposed placement heuristics that take into account isolation requirements. We have modified three classical heuristics to take into account these requirements. Furthermore, we have proposed four new heuristics that take into account the hierarchy of the Cloud platforms and the isolation requirements. Finally, we have evaluated these heuristics and compare them with the modified classical ones. We have shown that our heuristics are performing at least as good as classical ones but are scaling better and are faster by a few order of magnitude than the classical ones.

6.6.2. Security for Cloud Environment through Information Flow Properties Formalization with a First-Order Temporal Logic

The main slowdown of Cloud activity comes from the lack of reliable security. The on-demand security concept aims at delivering and enforcing the client's security requirements. In [50], we have presented an approach, Information Flow Past Linear Time Logic (IF-PLTL), to specify how a system can support a large range of security properties. We have presented how to control those information flows from lower system events. We have given complete details over IF-PLTL syntax and semantics. Furthermore, that logic enables to formalize a large set of security policies. Our approach is exemplified with the Chinese Wall commercial-related policy. Finally, we have discussed the extension of IF-PLTL with dynamic relabeling to encompass more realistic situations through the dynamic domains isolation policy.

6.6.3. Security Metrics for the Cloud Computing and Security-aware Virtual Machine Placement

In a classic Cloud Computing scenario, a client connects to a provider platform/service and submits his computation requirements, sometimes known as Service Level Agree- ments (SLAs). Then, the platform executes the computation taking into account, in its allocation algorithms, criteria like data location, CPU usage or duration of a job. As security in Cloud Computing is a main concern, we propose to consider security as another criteria for jobs scheduling. Thus, two questions need to be answered. The first one is how a client can describe his needs in terms of security level and the second one is how the scheduler could leverage the security to satisfy the client requirements? To provide an answer, a system of security metrics is essential. Indeed, with appropriate metrics, we can quantify and compare the security level of our resources. Moreover, a client can easily describe his security requirements and the scheduler can allocate the fitted resources using these metrics. Unfortunately, such system of metrics is not yet available. Consequently, we developed a system of security metrics specific to the Cloud Computing and scheduling algorithms using these metrics for a Security-Aware Virtual Machine (VM) placement.

6.7. Self-healing of Operational Issues for Grid Computing

Participant: Frédéric Desprez.

Many scientists now formulate their computational problems as scientific workflows. Workflows allow researchers to easily express multi-step computational task. However, their large scale and the number of middleware systems involved in these gateways lead to many errors and faults. Fair quality of service (QoS) can be delivered, yet with important human intervention. Automating such operations is challenging for two reasons. First, the problem is online by nature because no reliable user activity prediction can be assumed, and new workloads may arrive at any time. Therefore, the considered metrics, decisions and actions have to remain simple and to yield results while the application is still executing. Second, it is non-clairvoyant due to the lack of information about applications and resources in production conditions. Computing resources are usually dynamically provisioned from heterogeneous clusters, clouds or desktop grids without any reliable estimate of their availability and characteristics. Models of application execution times are hardly available either, in particular on heterogeneous computing resources.

In collaboration with Rafaël Silva and Tristan Glatard, we proposed a general self-healing process for autonomous detection and handling of operational incidents in scientific workflow executions on grids. Instances are modeled as Fuzzy Finite State Machines (FuSM) where state degrees of membership are determined by an external healing process. Degrees of membership are computed from metrics assuming that incidents have outlier performance, e.g. a site or a particular invocation behaves differently than the others. These metrics make little assumptions on the application or resource characteristics. Based on incident degrees, the healing process identifies incident levels using thresholds determined from the platform history. A specific set of actions is then selected from association rules among incident levels. The healing process is parametrized on real application traces acquired in production on the European Grid Infrastructure (EGI).

To optimize task granularity in distributed scientific workflows, we presented a method that groups tasks when the fineness degree of the application becomes higher than a threshold determined from execution traces. Controlling the granularity of workflow activities executed on grids is required to reduce the impact of task queuing and data transfer time. Our method groups tasks when the fineness degree of the application, which takes into account the ratio of shared data and the queuing/round-trip time ratio, becomes higher than a threshold determined from execution traces. The algorithm also de-groups task groups when new resources arrive. Results showed that under stationary load, our fineness control process significantly reduces the makespan of all applications. Under non-stationary load, task grouping is penalized by its lack of adaptation, but our de-grouping algorithm corrects it in case variations in the number of available resources are not too fast [21].

To address unfairness among workflow executions, we proposed an algorithm to fairly allocate distributed computing resources among workflow executions to multi-user platforms. We consider a non-clairvoyant, online fairness problem where the platform workload, task costs, and resource characteristics are unknown and not stationary. We define a novel metric that quantifies unfairness based on the fraction of pending work in a workflow. It compares workflow activities based on their ratio of queuing tasks, their relative durations, and the performance of resources where tasks are running, as information becomes available during the execution. Our method is implemented and evaluated on 4 different applications executed in production conditions on EGI. Results show that our method can very significantly reduce the standard deviation of the slowdown, and the average value of our metric [22].

AVIZ Project-Team

6. New Results

6.1. Hybrid-Image Visualizations

Participants: Petra Isenberg [correspondant], Pierre Dragicevic, Wesley Willett, Anastasia Bezerianos, Jean-Daniel Fekete.

We investigated hybrid-image visualization for data analysis in large-scale viewing environments. Hybrid-image visualizations blend two different visual representations into a single static view, such that each representation can be perceived at a different viewing distance. Our work was motivated by data analysis scenarios that incorporate one or more displays with sufficiently large size and resolution to be comfortably viewed by different people from various distances. Hybrid-image visualizations can be used, in particular, to enhance overview tasks from a distance and detail-in-context tasks when standing close to the display. By taking advantage of humans' perceptual capabilities, hybrid-image visualizations do not require tracking of viewers in front of a display. Moreover, because hybrid-images use a perception-based blending approach, visualizations intended for different distances can each utilize the entire display. In our paper we contributed a design space, discussed the perceptual rationale for our work, provided examples and a set of techniques for hybrid-image visualizations, and described tools for designing hybrid-image visualizations. We will also release software that will help in the construction of hybrid-image visualizations.

../../../../projets/aviz/IMG/hybrid.jpg

Figure 8. Overview of a treemap showing a subset of the tree of life with a hybrid image visualization. Larger structures are clearly visible from far but do not interfere with reading detail when up close (see in-set).

6.2. Visualization for Interactive Displays

Participants: Tobias Isenberg [correspondant], Petra Isenberg.

Because the access to and analysis of information is becoming increasingly important anywhere and at any time, researchers have begun to investigate the role of interactive displays as data analysis platforms. Visualization applications play a crucial role in data analysis and development of dedicated systems and tools for small to large interactive displays to support such application contexts is underway. We contribute a systematic and quantitative assessment of the literature from ten different venues, an open repository of papers, and a code-set that can be used to categorize the research space [22]. We found just over 100 publications at the intersection of interactive surfaces and visualization in our careful examination of 10 different publication venues related to the topic. We found that research has so far largely focused on the development of interaction techniques, for multi-touch tabletop devices, and 2D spatial and abstract visualizations. Together, all publications addressed a wide spectrum of research questions and, given the many possible combinations of interactive surfaces and visualization, the research space is still wide open. While several projects developed applications for data analysis with visualization on interactive surfaces, their availability in practice is still rare. Commercial companies and open-source communities have begun to provide ported versions of their products/tools for tablets and mobile phones (e.g., Tableau Mobile 4 and KiwiViewer 5), showing the need for visualization application on surfaces. Nevertheless, the support for data analysis tasks on these and other interactive surfaces can certainly still be improved—a lot more research with respect to the development and evaluation of the fundamentals of data exploration and analysis is needed for interactive displays. Some example directions of future work in this context were outlined in research agendas published in the journal IEEE Computer [24] as well as in IEEE Computer Graphics and Applications [21].

In a specific project investigated an interaction design concept for exploratory 3D data visualization that marries direct-touch interaction with stereoscopic vision. The design is inspired by the mental mapping that occurs for mouse interaction where the physical control space is mapped through a mental rotation to the display space. Similarly, we explore touch interaction on a monoscopic tablet, mapped through a mental rotation to the stereoscopic display space. Because this mental mapping can become increasingly skewed we show when and how to re-synchronize the views (see Figure 9).

Figure 9. Tablet-based navigation of a stereoscopically displayed 3D dataset.

6.3. Visualization for Soccer Analysis

Participants: Charles Perin, Romain Vuillemot, Jean-Daniel Fekete [correspondant].

A new generation of soccer data is now available, as some companies (http://www.optasports.com/) collect and provide extensive data covering almost all professional soccer championships, with a wealth of multivariate information related to time, player positions, and types of action, to name a few. Currently, most analysis on such data relate to statistics on individual players or teams. For instance, statistics on "team ball possession" and "number of goal attempts for team A or B" are popular on websites, TV and newspapers and often accompanied by bar charts or plots on a soccer field. However, soccer analysts we collaborated with consider that quantitative analysis alone does not convey the right picture of the game, as context, player positions and phases of player actions are the most relevant aspects.

SoccerStories [] (Figure 10) is a visualization interface to support analysts in exploring soccer data and communicating interesting insights that we designed to support the current practice of soccer analysts and to enrich it, both in the analysis and communication stages. Our system provides an overview+detail interface of game phases, and their aggregation into a series of connected visualizations, each visualization being tailored for actions such as a series of passes or a goal attempt. To evaluate our tool, we ran two qualitative user studies on recent games using SoccerStories with data from one of the world's leading live sports data providers. The first study resulted in a series of four articles on soccer tactics, by a tactics analyst, who said he would not have been able to write these otherwise. The second study consisted in an exploratory follow-up to investigate design alternatives for embedding soccer phases into word-sized graphics. For both experiments, we received a very enthusiastic feedback and participants consider further use of SoccerStories to enhance their current workflow. This article received a Best Paper Honorable Mention in VIS 2013.

We also explored how spectators of a live soccer game can collect detailed data while watching the game [46]. Our motivation arouse from the lack of free detailed sport data, contrasting with the large amount of simple statistics collected for every popular games and available on the web. Assuming many spectators carry a smart phone during a game, we implemented a series of input interfaces for collecting data in real time. In a user study, we asked participants to use those interfaces to perform tracking tasks such as locating players in the field, qualifying ball passes, and naming the player with ball while watching a video clip of a real soccer game. Our two main results are 1) the crowd can collect detailed and fairly complex data in real-time with reasonable quality while each participant is assigned a simple task, and 2) a set of design implications for crowd-powered interfaces to collect live sport data. We also discuss the use of such data into SoccerStories, and the design implications coming with the visual communication of missing and uncertain detailed data.

Finally, we presented R2S2 [45] in the SportVis workshop (VIS 2013), a hybrid visualization technique as an intermediate step between Rank Chart and Slope Graph to better understand and analyze team evolutions during soccer championships. Currently used rank tables for soccer are relative (ranked-based) and do not convey the absolute difference between teams. R2S2 provides a way to visualize these differences using the Slope Graph technique (value-based). By interactively setting the parameters of R2S2, we make the distance between teams appear, minimizing the overlaps caused by the Slope Graph technique.

More information about these projects is available at http://www.aviz.fr/soccer.

6.4. Interaction Model for Visualizations Beyond the Desktop

Participants: Yvonne Jansen [correspondant], Pierre Dragicevic.

We introduced an interaction model for beyond-desktop visualizations that combines the visualization reference model with the instrumental interaction paradigm. Beyond-desktop visualizations involve a wide range of emerging technologies such as wall-sized displays, 3D and shape-changing displays, touch and tangible input, and physical information visualizations. While these technologies allow for new forms of interaction, they are often studied in isolation. New conceptual models are needed to build a coherent picture of what has been done and what is possible. We described a modified pipeline model where raw data is processed into a visualization and then rendered into the physical world. Users can explore or change data by directly manipulating visualizations or through the use of instruments. Interactions can also take place in the physical world outside the visualization system, such as when using locomotion to inspect a large scale visualization. Through case

../../../../projets/aviz/IMG/soccer.png

Figure 10. Using SoccerStories: (a) navigating among soccer phases of a game; (b) mapping a phase on a focus soccer field; (c) exploring the phase by grouping actions into tailored visualizations; and (d) communicating using Sportlines embed into text.

studies we illustrated how this model can be used to describe both conventional and unconventional interactive visualization systems, and compare different design alternatives.

6.5. Network Visualization

Participants: Benjamin Bach [correspondant], Basak Alper, Andre Spritzer, Emmanuel Pietriga, Nathalie Henry-Riche, Tobias Isenberg, Jean-Daniel Fekete.

Although much research has been done on finding efficient ways to visualize different kinds of networks (social networks, computer networks, brain networks, etc), many problems are still open. Rather than trying to find optimal layouts, we focus on novel representation and navigation techniques to explore such networks. Our research focusses on three major problems: (i) heterogeneous networks, (ii) comparison of graphs, (iii) dynamic networks, and (iv) generating networks for controlled user evaluations.

Heterogeneous Networks: Heterogeneous networks are networks with multiple node and edge types, such as ontologies in the Semantic Web. Ontologies also provide means to describe type hierarchies on node and edge types as well as other set theoretic relations between these types. Such heterogeneous networks can be explored on two levels, the model (types and possible relations between types), and instances (actual nodes and their connections). In order to allow for interactive exploration of such data, we extended the NodeTrix technique [67] to visualize ontologies. Our prototype is called *OntoTrix* [11] (Figure 11 (a), which allows for reorganizing matrices by splitting and merging them, traversing node and edge hierarchies and visualize different types of connection between nodes.

Graph Comparision: Analyzing brain networks, which can represent anatomical fibers as well as functional correlation between brain regions, is complex in many ways. By analysing brain scientists tasks we concluded that many user tasks can actually performed by comparing two networks. In [28](Figure 11 (b)) we design and discuss several ways to compare two weighted graphs and finally compare the two most promising designs in a controlled user study. We found that our encoding for adjacency matrices outperforms the one for node-link diagrams, even for sparse networks. The implications for brain analysis tools are manyfold and our results generalize to other domains that are concerned with comparing (dense and weighted) networks.

Dynamic Networks: A very common technique to explore dynamic networks are animations and small multiples, each of which being supporting different tasks, while falling short on others. With *GraphDiaries* [10](Figure 11 (c)), we design an interface based on the combination of both techniques while offering flexible temporal navigation techniques as well as enhanced perceptive feedback to understand changes between time steps. *GraphDiaries* supports further navigation techniques such as temporal aggregation, direct difference views and layout adjustment. While *GraphDiaries* is highly extensible, its techniques are designed to be integrated in existing visualization tools.

While animations and the techniques in *GraphDiaries* are useful for many networks, dense dynamic networks are still an important open problem. We hence generalized the idea of matrices to visualize temporal networks, by describing a visualization and interaction model based on the space time cube metaphor (Figure 11 (d)). In analogy with the physical world, this *Matrix Cube* can be manipulated and decomposed in order to explore the network, while the cube model serves as a consistent visual and mental model of the data and visualization. We implemented an interface called *Cubix* that allows us to perform simple view switches and decomposition operations in the cube. Cubix and the Matrix Cube was evaluated with two experts, an astronomer and brain scientist, exploring their own real world data. With the Matrix Cube and its decomposition operations, we are able to visualize and navigate within very dense dynamic networks such as brain networks, trading flows and technical networks. The design space of possible visualizations that the Matrix Cube and its operations offer is both, huge but structured at the same time. It allows us to explore many future designs.

As part of the effort of visualizing publications and work of Jean-Daniel Fekete, we designed a visualization to show his collaborations over the past years, relating his papers and his collaborators [29](Figure 11 (e)). A poster was presented at part of a poster submission to IEEE Vis, 2013 in Atlanta. To the best of our knowledge, no such technique to visualize any sort of dynamic ego networks have been published so far.

../../../../projets/aviz/IMG/networks.png

Figure 11. Examples for network visualizations

Network Generation: As any controlled user study, evaluating network visualizations requires control over the data. However, it is hard to found real world data with the desired properties and in reasonable amount for a controlled user study. Synthetic data can help but the output of random graph generators is hard to control and hardly resembles actual real-world data. With GraphCuisine [56] we present and interactive approach to generate graphs. In an iterative process, the computer generates suggestions while the user selects her preferred graphs and graph measures.

6.6. GridVis: Visualisation of Island-Based Parallel Genetic Algorithms

Participants: Waldo Cancino [correspondant], Hugo Gilbert, Benjamin Bach, Evelyne Lutton, Pierre Collet.

Island Model parallel genetic algorithms rely on various migration models and their associated parameter setting. A fine understanding of how the islands interact and exchange informations is an important issue for the design of efficient algorithms. GridVis, is an interactive tool that has been developed for visualising the exchange of individuals and the propagation of fitness values between islands. GridVis has been developed in Java, to monitor how the islands communicate: when and how much individuals of which fitness they effectively exchange during a run. We model the computer cluster that is running the island model, as dynamic network and use an adjacency matrix to show the relations (exchange between individuals) between nodes (computers) in the cluster (Figure 12 (a)). Several experiments have been performed on a grid and on a cluster to evaluate GridVis' ability to visualise the activity of each machine and the communication flow between machines. Experiments have been made on the optimisation of a Weierstrass function using the EASEA language, with two schemes: a scheme based on uniform islands and another based on specialised islands (Exploitation, Exploration and Storage Islands).

../../../../projets/aviz/IMG/figure1_labelled.png

Figure 12. Visualisation of a grid with 20 machines: Each computer in the cluster appears twice in the matrix, once as row and once as column. Cells inside the matrix indicate information about the interaction of computers during evolution, for example, the amount of individuals exchanged (read from row to column). Similar to heat maps activity (exchange of individuals) is mapped to darkness (dark cells indicate high exchange, bright cells show low exchange).

AXIS Project-Team

6. New Results

6.1. Introduction

Our new results are split into our three sub-objectives as described in Section 3.1 :

- **Sub-Objective 1: Mining for Knowledge Discovery in Information Systems:**
  This year we obtained ten main results (cf. Section 6.2): five on Clustering methods, four on how to apply these clustering methods on real data and finally one related to the use of ontology for Multi-View KDD process.

  Let us note that two 2011 results have been published this year as book chapters [34], [31]. Chongsheng Zhang published also his work conducted during his Explore programm at UCLA (USA) when, as AxIS PhD student, he was visiting the WIS team of Prof. Carlo Zaniolo at UCLA in 2010 [26].
- **Sub-Objective 2: Information and Social Networks Mining for Supporting Information Retrieval:**
  This year, we pursued our two main works on this topic (cf. Section 6.3):
  - the detection of communities in a social network (detection of graphs extracted from relational data) (cf. Section 6.3.1),
  - the multi view clustering of relational data (cf. SEction 6.3.2).
- **Sub-Objective 3: Interdisciplinary 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 [70] [11] , experimentations take place out of labs with large numbers of heterogeneous people instead of carefully controlled panels of users etc.

  All these deep changes required improvements of existing practices, methods and tools for the design/evaluation of information systems as well as for usage analysis. This evolution called also for a structured user-centred methodology (methods and ICT tools) to deal with open innovation. Various different disciplines and trends are dedicated in understanding user behaviour on Internet and with Digital Technologies, notably Human Computer Interaction community (HCI), Computer Supported Cooperative Work (CSCW), Workplace Studies, Service Design, Distributed Cognition and Data Mining.

  Our contribution to open innovation research related to ICT-based services or products keeps its focus on usage analysis and user experience measurement for design, evaluation and maintenance of information systems and our activities from 2011 have been conducted both breadth wise and in depth with two main objectives :
  - Improving design and evaluation support tools and methods for user driven driven innovation,
  - Development of the FocusLab platform

This year, our research was conducted along three focus:

- Extension of usability methods and models (cf. Section 6.4). First we pursued our work on User Evaluation and Tailoring of Personal Information in the context of the ANR project PIMI. Second a paper related to our strategy and heuristics for rural tourist web sites benchmarking elaborated in the context of the Pacalabs project HOTEL-REF-PACA is written for submission in 2014;
- Designing and evaluating user experience in the context of a living lab: this year five results came from ELLIOT project (cf. Section 6.5) such as an environmental data platform based on citizen sensing, low-cost sensor, user experience measurement, user behaviour change analysis, studies of persuasive technologies and gamification in Energy economy and green services.
- FocusLab Platform (cf. Section 6.6).

6.2. Mining for Knowledge Discovery in Information Systems

6.2.1. Fuzzy Clustering on Multiple Dissimilarity Matrices

Participants: Yves Lechevallier, Francisco de Carvalho.

During 2013 we introduce fuzzy clustering algorithms [18] and [27] that can partition objects taking into account simultaneously their relational descriptions given by multiple dissimilarity matrices. The aim is to obtain a collaborative role of the different dissimilarity matrices to get a final consensus partition. These matrices can be obtained using different sets of variables and dissimilarity functions. These algorithms are designed to furnish a partition and a prototype for each fuzzy cluster as well as to learn a relevance weight for each dissimilarity matrix by optimizing an adequacy criterion that measures the fit between the fuzzy clusters and their representatives. These relevance weights change at each algorithm iteration and can either be the same for all fuzzy clusters or different from one fuzzy cluster to another.

A new algorithm [19] based on a non-linear aggregation criterion, weighted Tchebycheff distances, more appropriate than linear combinations (such as weighted averages) for the construction of compromise solutions is proposed.

Experiments with real-valued data sets from the UCI Machine Learning Repository (http://archive.ics.uci.edu/ml/) as well as with interval-valued and histogram-valued data sets show the usefulness of the proposed fuzzy clustering algorithms.

6.2.2. Clustering of Functional Boxplots for Multiple Streaming Time Series

Participant: Yves Lechevallier.

We introduced a micro-clustering strategy for Functional Boxplots [30]. The aim is to summarize a set of streaming time series split in non overlapping windows. It is a two step strategy which performs at first, an on-line summarization by means of functional data structures, named Functional Boxplot micro-clusters; then it reveals the final summarization by processing, off-line, the functional data structures. Our main contribution consists in providing a new definition of micro-cluster based on Functional Boxplots and, in defining a proximity measure which allows us to compare and update them. This allows us to get a finer graphical summarization of the streaming time series by five functional basic statistics of data. The obtained synthesis will be able to keep track of the dynamic evolution of the multiple streams.

This work is done in collaboration with the laboratory of Political Science "Jean Monnet", Second University of Naples, Caserta, Italy.

6.2.3. Web Page Clustering based on a Community Detection Algorithm

Participant: Yves Lechevallier.

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.

During 2013 we pursued our previous work on our approach for extracting data based on the modularity function. This approach discovers the existing communities by modeling the data obtained in the pre-processing operation as a weighted graph. The method discriminates the communities through their subject of interest and extract relevant knowledge.

This work is done in collaboration with Yacine Slimani from the LRIA laboratory at the Ferhat Abbas University, Setif, Algerie and will be submitted to an international journal.

6.2.4. Normalizing Constrained Symbolic Data for Clustering

Participants: Marc Csernel, Francisco de Carvalho.

Clustering is one of the most common operation in data analysis while constrained is not so common. During 2013 we presented a clustering method [31] in the framework of Symbolic Data Analysis (S.D.A) which allows us to cluster Symbolic Data. Such data can be constrained relations between the variables, expressed by rules which express the domain knowledge. But such rules can induce a combinatorial increase of the computation time according to the number of rules. The algoritm presented a way to cluster such data in polynomial time. This method is based first on the decomposition of the data according to the rules, then we can apply to the data a clustering algorithm based on dissimilarities.

6.2.5. Dynamic Clustering Method for Mixed Data

Participants: Yves Lechevallier, Marc Csernel, Brigitte Trousse.

For ELLIOT project purposes (cf. Section 7.3.1), a new version of MND method (Dynamic Clustering Method for Mixed Data) has been elaborated. It determines iteratively a series of partitions which improves at each step the underlying clustering criterion. All the proposed distance functions for p variables are determined by sums of dissimilarities corresponding to the univariate component descriptors Y_j. The most appropriate dissimilarities have been suggested above according to the type of variables.

In practice, however, data to be clustered are typically described by different types of variables. An overall dissimilarity measure is obtained by a linear combination of the dissimilarity measures computed with respect to the different kinds of variables.

A new release of MND algorithm based on past work [80] has been developed for ELLIOT purposes, providing some default configuration parameters for non experts.

In this version two types of distances are proposed:

- **Quantitative distance**: the choice is type L1 distance or Euclidean distances when the types of variables are quantitative or continuous.
- **Boolean distance**: the choice is Khi2, type L1 distance or Euclidean distances when the type of variables is categorical or discrete.

This algorithm has been applied to cluster answers at questionnaires issued from a diary tool within the ELLIOT Green Services use case (cf. Section 6.5.4).

6.2.6. Applying a K-means clustering method for districts clustering according to Pollution

Participants: Brigitte Trousse, Yves Lechevallier, Guillaume Pilot, Caroline Tiffon.

Our motivation was to provide citizen a comparative analysis at the district level related to pollution data from Azimut stations (ozone O3 and nitrogen dioxide NO2). To achieve this, the Nice Côte d'Azur territory was discretized into small areas. IoT Data are preprocessed for each district and period of time before applying clustering. The temporal and spatial units were clustered into 5 and then into 6 clusters. The partition into 5 clusters was selected, then the temporal units for each area were counted. For the partition in 5 clusters, for each area the percent of each cluster was counted. Around 30 areas with more than 10 temporal units were found. We improved this to classify different districts of the city based on their IoT data (Azimut data O3-NO2) for each hour/day in order to provide a new functionnality in the second version of MyGreenServices.

This work is partially funded by ELLIOT project (see Section 7.3.1).

6.2.7. *Summarizing Dust Station IoT Data with REGLO, a FocusLab web service*

Participants: Yves Lechevallier, Brigitte Trousse, Guillaume Pilot, Xavier Augros.

Within ELLIOT, we applied the GEAR (or REGLO in French) method [57], [58], [59] on the evolution of dust data issued from one citizen sensor.

Our motivation was to summarize IoT data in order to have a pollution context for each user. Such IoT summaries constitute interesting individual contextual data for supporting the living lab manager to better interpret the user behavior and finally the user experience.

REGLO summarised IoT data with isolated points and line segments.

The goal now is to carry out an analysis of these summaries to automatically determine the characteristics of the curve.

We selected only segments. For each segment we calculated four variables that characterize it:

- The slope of the segment,
- The midpoint of the segment (average of this segment),
- The length of the segment,
- The duration of the segment (the time interval between the start time and the end time of the segment).

From these four values we can achieve an interpretation of the previous curve, taking into account only two variables and constructing a 2D representation.

This work is partially funded by ELLIOT project (see Section 7.3.1).

6.2.8. *Clustering of Solar Irradiance*

Participants: Thierry Despeyroux, Francisco de Carvalho, Yves Lechevallier, Thien Phuc Hoang Nguyen.

The development of grid-connected photovoltaic power systems leads to new challenges. The short or medium term prediction of the solar irradiance is definitively a solution to reduce the storage capacities and, as a result, authorizes to increase the penetration of the photovoltaic units on the power grid. We present the first results of an interdisciplinary research project which involves researchers in energy, meteorology and data mining, addressing this real-world problem. The objective here is to show interest and disadvantages of two approaches for classifying curves.

In Reunion Island from December 2008 to March 2012, solar radiation measurements has been collected, every minutes, using calibrated instruments. Prior to prediction modelling, two clustering strategies has been applied for analysis the data base of 951 days.

During 2013 we continued our research and obtained many results [28].

Our methodology is based on two clustering approaches. The objective here is to show interest and disadvantages of two approaches for classifying curves.

The first approach combines the following proven data-mining methods. Principal Component Analysis was used as a pre-process for reduction and de-noising and the Ward Hierarchical and K-means methods to find a partition with a good number of classes.

The second approach [78],[20] uses a clustering method that operates on a set of dissimilarity matrices. Each cluster is represented by an element or a subset of the set of objects to be classified. The five meaningfully clusters found by the two clustering approaches are compared.

6.2.9. Understanding of Cooking User's Recipes by Extracting Intrinsic Knowledge

Participants: Damien Leprovost, Thierry Despeyroux, Yves Lechevallier.

On community web sites, users share knowledge, being both authors and readers. We present a method to build our own understanding of the semantics of the community, without the use of any external knowledge base. We perform this understanding by knowledge extraction from analysed user contributions. We propose an evaluation of the trust attributable to that deduced understanding to assess the quality of user content, on cooking recipes provided by users on sharing web sites. This work is partially funded by FIORA project (see Section 7.2.2). Two articles have been accepted in early 2014 [25], [29].

6.2.10. Knowledge Modeling for Multi-View KDD Process

Participant: Brigitte Trousse.

We pursued our supervision (with our colleagues H. Behja and A. Marzark from Morocco) of E.L. Moukhtar Zemmouri's PhD thesis (Morocco) on a Viewpoint Model in the context of a KDD process, topic we initiated during Behja's PhD thesis [40]). E. Zemmouri defended his thesis at the end of this year [75]. Below is the summary of his PhD thesis.

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 involves several actors (domain experts, data analysts, KDD experts â€¦) 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 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 re-usability in terms of viewpoint of successful data mining experiences within an organization. An article selected from the international conference NGNS 2012 [74] will be published in the on-line Journal of Mobile Multimedia, Volume 9 No.3 &4 March 1, 2014.

6.3. Information and Social Networks Mining for Supporting Information Retrieval

6.3.1. Clustering of Relational Data and Social Networks Data: Graph Aggregation

Participant: Yves Lechevallier.

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 we updated the program developed by Louati Amine in 2011. A book chapter [34] proposes a new aggregation criteria.

This work is done by Louati Amine (AxIS) in collaboration with Marie-Aude Aufaure, head of the Business Intelligence Team, "Ecole Centrale de Paris", MAS Laboratory.

6.3.2. Multi-View Clustering of Relational Data

Participants: Thierry Despeyroux, Francisco de Carvalho, Yves Lechevallier.

In the work reported in [47] in collaboration with Francisco de A.T. de Carvalho, we introduce an improvement of a clustering algorithm described in [78] 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 use 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. This work has been published this year as a chapter in "Advances in Knowledge Discovery and Management" [32].

6.4. Extension of Usability Methods and Tools

6.4.1. 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) an ergonomic evaluation was conducted on the initial prototype, in its PC version [49] and its mobile version [48]. In addition, an experiment was conducted on the usability of the new improved PIMI prototype. The goals were to evaluate its usability, and to assess user tailoring as an evaluation technique. Thirty users participated to the study: a first part consisted in a standard user test (SUT) and a second part was a usability test with tailoring (UTT). Overall, a total of 51 usability problems were diagnosed. Among those, 32 resulted from SUT, and 19 from UTT. Part of the latter (11) are additional to the ones identified during SUT, and to those diagnosed previously by usability inspection (UI with Ergonomic Criteria). The active involvement of users through customization scenarios appear to provide additional cues for usability assessment, and for design, with new generic usability recommendations [23],[22].

6.5. Designing and Evaluating User Experience and Methods for Open Innovation

6.5.1. MyGreenServices: a Pollution Collective-Awareness Platform based on Citizen Sensing

Participants: Brigitte Trousse, Guillaume Pilot, Xavier Augros, Florian Bonacina, Caroline Tiffon, Anne-Laure Negri, Bernard Senach.

Adopting a living lab approach and following an experiential design process [63], we co-created with users and implemented a Pollution Collective-Awareness platform based on Citizen Sensing called "MyGreenServices" [38]. This deployment was very rich in terms of a better understanding of research problems to be addressed in this context in order to lead to user behaviour changes: citizen sensing, environmental crowdsourcing platform and user experience in the context of IoT.

MyGreenServices (http://mygreenservices.inria.fr) which was very robust offers various green services such as the visualization of environmental data collected by citizen, the alert services, the ability to download data, the forum for sharing ideas and best practices in terms of eco-responsible behaviors. MyGreenServices provides access to citizen measures (stations and electric vehicles) for any registered user. Moreover, citizens who host a station can trace the time history of the data sensed. The priority was to provide to users all the IoT data by them. Two ways to represent data have been chosen as shown in Figure 1 :

- The use of maps with measures coming from environmental sensors and based on a colour scale indication;
- The pollution curves that support the cartography and allow the access to the detailed data for the user.

A pollution alert service has been created considering two points of view:

- The first consists of localising a person (with his agreement) and indicating via email or text message the passage through a polluted area;
- The second allow the user to define an area to follow and the user will be advised of pollution alerts for the area by email or text message.

An important effort has been done in designing, testing and improving user interfaces based on pre-test with the usability testing software named Morae and experiments in real situations.

Two experiments have been carried out in February and in June 2013, with the aim to test the platform MyGreenServices by two user profiles (consumers and producers of data) and to measure User experience. The aim of the experiments is to assess the user experience and experiential learning related to MyGreeenServices; this includes experience related to the IoT devices, to the measures and services as well as air quality awareness and behaviour changes monitoring. See Section 6.5.3) for more details on the used model and measurement methodology.

For supporting Citizen Sensing, we elaborated IoT installation guides for our three Pollution stations (based on user feedbacks): Pollux station for dust from CKAB [7], Azimut stations for Ozone and Nitrogen dioxide from Azimut Monitoring [8] and AxISbox stations for dust (Inria Cf. Section 6.5.2).

In order to ensure a proper data analysis, log and usage analytics were structured and gathered in an admin tool designed by the AxIS team at Inria. This tool is a component of the MyGreenServices portal.

6.5.2. AxISbox, a Prototype of a Low-Cost Dust Arduino-based Station

Participant: Guillaume Pilot.

In order to provide more citizen sensors during our Elliot experiments, we developed a first prototype of a new low cost dust (PM10) station (with Rasburry and Arduino) called AxISbox (cf. Figure 2) which we tested for research purposes. This prototype was validated during the second ELLIOT experiment in June.

6.5.3. Modelling and Measuring User Experience for Green IoT-based Services

Participants: Brigitte Trousse, Anne-Laure Negri, Caroline Tiffon, Xavier Augros, Guillaume Pilot.

[7] CKAB URL: http://ckab.com/polluxnz-city
[8] Azimut Monitoring URL: http://www.azimut-monitoring.com/

../../../../projets/axis/IMG/MGSsynthesis.png

Figure 1. MyGreenServices Platform

../../../../projets/axis/IMG/MGSsensors.png

Figure 2. Citizen sensors: Pollux station, AxISbox and Azimut mobile station

In accordance with the overall objective of MyGreenServices, we provided an UX modelling and measurement methodology for Green IoT-based services we applied on MyGreenServices. In our ELLIoT context, we focused on the level of awareness/experiential learning raised after usage of MyGreenServices (awareness pollution, awareness of citizen dissemination and change of behaviors), the ease of use and diffusion aspects (as being a tool provided to the citizen). Two objects of the learning were considered: IoT via myGreenServices portal and Air quality. We used a differential between a pre-profile and post-profile. Our UX methodology in the context of ELLIOT project is lying on the five steps we applied on the two versions of MyGreenServices:

- Instantiation of the holistic UX model elaborated within ELLIOT [63] (cf. the first three columns in Figure 3),
- Choice of types of UX momentary, episodic, cumulative) depending on the moment of the measurement (cf. Figure 4),
- Identification of relevant data to be collected and UX indicators (cf. the last two columns in Figure 3),
- Definition of UX metrics for indicators and rules (see Section 6.5.4 for the example of the Usefulness property),
- and finally data pre-processing and UX indicators/properties computation (via for some properties FocusLab 6.6).

../../../../projets/axis/IMG/MGS-UX-model.jpg

Figure 3. MyGreenServices UX Model

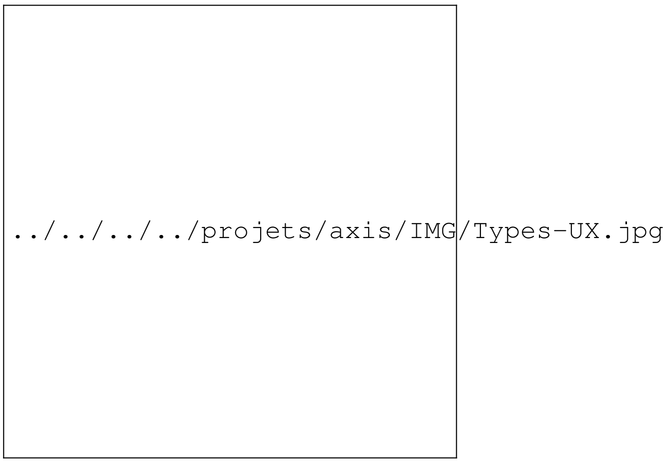


Figure 4. UX Types extracted from [70]

The two experiments clearly indicate both good results in terms of user experience with better result for the second experiment due to the improvement of MyGreenServices (v2) and better community management. A comparative analysis has been made for our two experiments, showing better quantitative value of UX indicators for the second version which was based on User feedback.

6.5.4. Evaluating User Behaviour Changes For MyGreenServices Usefulness Measurement

Participants: Brigitte Trousse, Yves Lechevallier, Xavier Augros, Caroline Tiffon.

The Usefulness UX property of our UX model [38] is calculated by aggregating the analysis of two questions related to a change of behaviours during (4 times) and/or after the experiment in terms of: transportation, aeration, outgoing, sport, aeration or others. We used the web service MNDClustering_Sequence (based on our MND clustering method [45]) to classify the answers to these questions and to provide a sequence of clusters by each user. See Section 6.6.2 related to this new web service.

A data table was built with all the answers for each (user, timestamp) and is analyzed to generate a partition in 3 clusters for the experiment by calling the Focuslab MND webservice (cf. Section 6.2.5) which has been improved this year. The Output via MNDClusterSequence web service is a csv data file with for each user the sequence of 5 clusters obtained during the experiment.

Then we identified the users having changed their behaviour. We use the following UX rules to conclude on this property:

- If % users declaring a change of behaviour $>$ 5% then high
- If % users declaring a change of behaviour $<$ 5% and $>$ 1% then medium
- If % users declaring a change of behaviour $<$ 1% then low

The result is "high" related to our two experiments. Note that other questions related to the usefulness of some MyGreenServices functionalities (alerts, forum, data synthesis, etc.) could be integrated in a more global rule for Usefulness.

6.5.5. Persuasive Technologies in Energy Economy

Participants: Bernard Senach, Anne-Laure Negri.

The ECOFFICES project [51] was for AxIS project team our first step towards eco-behaviour study. This research was complemented in 2012 with a literature review aiming at a deeper understanding of breaks and levers to eco behavior adoption. The work in this topic lead to a presentation [9]in the mobility context during the GreenCode Forum (see the video on YouTube) and to an internal seminar for Axis members. A draft of an Inria research report on this topic has been started.

The two research lines "Energy Economy" and "Persuasive Technology" have been merged and an analysis of the Ecoffices challenge has been engaged in the light of works in the fields of Persuasive Technologies and Game Design. In this analysis, the Ecoffices Energy challenge is considered as an hybrid system combining gamification and persuasive principles. Using available models of each field, the experimental device used in the Ecoffices project is deconstructed and evaluated. The persuasive quality analysis relies on the Persuasive System Design model [62]. Concerning the gaming quality of Ecoffices, a first model (Octalysis http://www.yukaichou.com/gamification-examples/octalysis-complete-gamification-framework/) was discarded and we are now using the gamification principles from the literature for the analysis [76].

At the end of 2012, we joined the work group PISTIL (Persuasive Interaction for SusTainablLity) and engaged several actions within this group and two papers are planned for the JIPS 2014 Special Issues on Persuasive Technologies [10]: one on an analysis of the ECOFFICES challenge (under writing) and another on the design and evaluation of persuasive systems.

6.5.6. Persuasive Technologies in Green Services

Participants: Brigitte Trousse, Anne-Laure Negri, Mylène Leitzelman, Florian Bonacina, Caroline Tiffon.

The ELLIOT project was for AxIS project team our second step towards eco-behaviour study. It provided us a very rich context to study behaviour changes related to pollution awareness. Our experimental results showed a very promising tendancy in terms of user behaviour changes and the impact of MyGreenServices on leading user eco-behaviours [38].

Persuasive technologies and gamification were used in the context of green Services use case. A specific focus was on gamification for the two customised Ideastream-based tools we developed for the co-creation step and mainly for the one used inside MyGreenServices platform (see Figure 5).

6.6. FocusLab Platform

6.6.1. New Graphical Charter and New Functionalities

Participants: Xavier Augros, Florian Bonacina, Brigitte Trousse.

This year we implemented a new version of the Focuslab platform (v1.3) (http://focuslab.inria.fr) with a new graphical charter, the addition of the documentation part (books, articles, thesis, reports, etc.) and new functionalities such as cross references between the hardware/software parts with the documentation part, the opportunity of reserving hardware, hardware+software or documentation and a new administration interface. This new version has been tested internally in the team at the end of the year.

6.6.2. FocusLab Generic Web Service: MNDCluster_Sequence

Participants: Xavier Augros, Yves Lechevallier, Brigitte Trousse.

This year for Elliot purposes, we built a new FocusLab generic Web Service called MNDCluster-Sequence. This web service uses the new release of MND clustering method [80] (cf. Section 6.2.5) which computes the best partition based on all data for each (user, timestamp). Then it builds for each user the sequence of 5 clusters taking into account the five user time stamp in our case. The resulting sequences are then added for each user as new qualified data in the dataset of Green Services.

This web service is added to those already integrated in FocusLab (See for more details our 2012 activity report http://raweb.inria.fr/rapportsactivite/RA2012/axis/uid116.html)

[9]URL: http://www-sop.inria.fr/axis/papers/2012/GreenCode_2012
[10]On-line journal : Journal d' Interaction Personne-Système, Journal of "Association Francophone d'Interaction Homme-Machine".

../../../../projets/axis/IMG/MGSforum.jpg

Figure 5. "Gamified Forum" page (including AxISbox)

AYIN Team

6. New Results

6.1. Markov Random Fields

6.1.1. Hierarchical multitemporal and multiresolution classification in remote sensing imagery

Participants: Ihsen Hedhli, Josiane Zerubia [contact].

This activity was conducted in collaboration with Dr. Gabriele Moser and Prof. Sebastiano B. Serpico (Department of Electrical, Electronic, and Telecommunications Engineering and Naval Architecture, DITEN, University of Genoa) [http://www.unige.it] with partial financial support from CNES [http://www.cnes.fr].

Markov random field (MRF), hierarchical classification, satellite image time series

The capability to monitor the Earth's surface, and especially urban and built-up areas, for environmental disasters such as floods or earthquakes, and to assess the ground impact and damage caused by such events, play important roles from multiple social, economic, and human viewpoints. Current and forthcoming satellite missions for Earth observation (EO; e.g., Pleiades, COSMO-SkyMed, TerraSAR-X, Sentinel) possess huge potential for such applications, as they allow a spatially distributed and temporally repetitive view of the monitored area at the desired spatial scales. In this framework, accurate and time-efficient classification methods using time series are especially important tools for supporting rapid and reliable assessment of the ground changes and damage induced by a disaster, in particular when an extensive area has been affected. Given the huge amount and variety of data available, the main difficulty is to find a classifier that takes into account multi-band, multi-resolution, multi-date, and possibly multi-sensor data.

This research addresses the problem of supervised classification at multiple spatial resolutions for multiple dates. The approach is based on the extension of recent methods proposed by DITEN and/or AYIN [4], [5], [6]. These methods focus on a supervised Bayesian classifier that combines joint class-conditional statistical modeling and a hierarchical Markov random field. The key idea of the proposed method is to combine the multiresolution modeling capabilities of this previous technique with a model for the temporal correlation among distinct images in a time series. For this purpose, a hierarchical spatio-temporal Markov random field model has been proposed that is aimed at fusing the pixel-wise, neighborhood, multiresolution, and temporal information associated with the input time series. Pixel-wise information is characterized through separate statistical modeling for each target class (e.g., vegetation, urban, etc.) by using a finite mixture model, estimated using a modified stochastic expectation maximization algorithm. Such a model is well suited to dealing with heterogeneous classes, and each mixture component may reflect the contribution of the different materials contained in a given class. At each considered resolution, the different input bands are statistically combined by using multivariate copulas, and the resulting statistical pixel-wise model is integrated in a hierarchical Markov random field based on a quad-tree structure. Among the different algorithms employed in the literature, we chose to use an exact estimator based on the marginal posterior mode (MPM). Specifically, a new formulation of MPM is developed to formalize, within the aforementioned hierarchical model, a 'cascade' multi-date decision rule. Such a classifier is sufficiently flexible to take into account different types of data (e.g., multispectral, panchromatic, synthetic aperture radar). The method is being experimentally validated with data acquired over a given area at different resolutions (e.g., multiresolution Pleiades images), directly integrated at the different levels of the cascade hierarchical model. An example of a classification result is illustrated in Fig. 1 . Here, Pleiades multiresolution images (panchromatic resolution: 50 centimeters and multispectral resolution: 2 meters) acquired over Port-au-Prince quay (Haiti) on two different dates are considered. Spatially disjoint training areas were manually annotated. The classification has been performed with respect to 5 main classes: urban areas, natural landscape, sand, containers, and wet areas. A visual analysis of the resulting map suggests that the proposed approach achieves remarkable accuracy.

../../../../projets/ayin/IMG/fig1.png

Figure 1. Right: Initial optical image of Port-au-Prince (Haiti) (©CNES, 2013). Left: Classification map obtained with the proposed multi-temporal hierarchical method for the 5 classes (blue: wet areas; green: vegetation; red: urban areas; yellow: sand; purple: containers).

6.1.2. A multi-layer Markov model for change detection in temporally separated aerial image pairs

Participants: Praveer Singh, Josiane Zerubia [contact].

This work was carried out in collaboration with Prof. Zoltan Kato from Institute of Informatics, University of Szeged, Hungary. [http://www.inf.u-szeged.hu/~kato/].

Multilayer Markov Random Fields (MRF), Histogram of Gradients (HOG), change detection, graph-cut optimization, aerial / satellite images.

In the proposed approach, we have tried to include both texture as well as pixel level information to build a three layer Markov model using the *Histogram of Oriented Gradients (HOG)* and the *Gray Level Difference* features on the topmost and bottommost layer respectively. Using a ground truth (GT) mask defined manually by an expert for each of the image pairs in the data set (obtained from the Hungarian Institute of Geodesy, Cartography and Remote Sensing), we employ a supervised technique to mark the initial set of pixels / sites as foreground or background. On the basis of the *HOG difference* and the *Gray level difference* feature vector corresponding to all the pixels in the image pair, a probability density function is fitted individually for the binary label set comprising of foreground and background labels using the GT. The probabilistic estimate is calculated using one training image pair for each data set. Using this probabilistic measure, a negative log likelihood is computed for each pixel (for both the features as well as the binary label set) which is then passed to the energy function of the proposed 3-layer MRF model. The final segmentation is obtained by minimizing the energy using a graph-cut algorithm, and subsequently a final foreground and background labelling is obtained over the combined layer. Figure 2 , shows aerial image pairs, one of them captured in 1984 by FOMI, Hungary (a) and the other one by GoogleEarth in 2007 (c). (b) is the ground truth and (d) is a combination of the hierarchical MRF based change detection (in red), ground truth (in green) and changes detected correctly (in yellow).

6.1.3. Graph-cut model for spectral-spatial classification of hyperspectral images

Participants: Aakanksha Rana, Yuliya Tarabalka [contact].

Hyperspectral images, graph cut, multi-label alpha expansion, contextual information, energy minimization

The very high spatial and spectral resolution of the last generation of remote sensors provides rich information about every pixel in an image scene, hence opening new perspectives in classification, but also presenting the challenge of analysing high data volumes. While pixel-wise classification methods analyze each pixel independently, classification results can be significantly improved by including spatial information in a classifier.

In this work, we proposed a spectral-spatial method for hyperspectral image classification based on a graph cut. The classification task is expressed as an energy minimization problem on the spatio-temporal graph of image pixels, and is solved by using the graph-cut α-expansion approach. The energy to optimize is computed as a sum of data and interaction energy terms, respectively. The data energy term is computed using the outputs of the probabilistic support vector machines classification. The second energy term, which expresses the interaction between spatially adjacent pixels in the eight-neighborhood, is computed by using dissimilarity measures between spectral vectors, such as vector norms, spectral angle map, or spectral information divergence. The performance of the proposed method was validated on hyperspectral images captured by the ROSIS and the AVIRIS sensors. Figure 3 compares classification results obtained by applying support vector machines and the proposed approach for the ROSIS hyperspectral image acquired over the University of Pavia. The new method yields higher classification accuracies when compared to the recent state-of-the-art approaches.

6.2. Marked point processes

6.2.1. Marked point process models for boat extraction from high resolution remotely sensed optical images

Participants: Paula Craciun, Josiane Zerubia [contact].

../../../../projets/ayin/IMG/HMRF.png

Figure 2. Change detection in an aerial image pair using a hierarchical MRF. a) Aerial image captured in 1984 by ©FOMI; b) Ground truth; c) Aerial image captured by ©GoogleEarth in 2007; d) Combination of the hierarchical MRF based change detection (in red), ground truth (in green), and changes detected correctly (in yellow).

../../../../projets/ayin/IMG/pavia_png.png

Figure 3. Hyperspectral image of the University of Pavia. (a) Ground-truth (b) Support vector machines classification map. (c) Graph-cut classification map.

This work was done in collaboration with Dr. Mathias Ortner (ASTRIUM EADS) [http://www.astrium.eads.net] and Prof. Pierre del Moral (ALEA team, Inria Bordeaux).

Stochastic geometry, Markov model, detection, parallel algorithm

Marked point process models have been successfully applied to object extraction problems in high resolution optical images, ranging from tree crown or road extraction to flamingo or crowd counting. We try to model the problem of boat detection and counting in harbors. The difficulty of this problem resides in the particular distribution of the objects. The model consists of two energy terms: a data term, which reflects the model's fidelity to the input image, and a prior term containing knowledge about the objects to be extracted. The model relies on a high number of parameters and is computationally intensive. The purpose of this research is to extend a previously developed marked point process model of ellipses and make it more computationally manageable. In particular, we add a preprocessing step in which we determine the global and local direction of the objects [8], [17]. Additionally, segmentation of land and water areas is implemented as a preprocessing step. Boat extraction results are shown in Figure 4 . Finally, we implement an improved parallel sampler, thereby drastically improving computation times.

../../../../projets/ayin/IMG/Extraction.png

Figure 4. Boat extraction in a harbor using a marked point process model (a) harbor image ©CNES; (b) extraction results.

6.2.2. Parameter estimation for automatic object detection in very high resolution optical images

Participants: Aurélie Boisbunon, Josiane Zerubia [contact].

This work was partially funded by the French Space Agency CNES [http://www.cnes.fr].

Markov model, Monte Carlo method, evolutionary algorithm, optimization, image processing, detection

The main goal of this work is to study parameter estimation for several marked point processes. Currently, the parameters of such models are estimated by a Stochastic Expectation and Minimization (SEM) algorithm, which is computationally expensive. We will investigate and propose new parameter estimation techniques, based on Randomized Quasi-Likelihood and evolutionary algorithms, for the parameters of the probability density of a marked point process. The goal is to improve computation times with respect to SEM while maintaining similar accuracy. The first application envisioned is boat detection for harbor activity monitoring (see Figure 5).

../../../../projets/ayin/IMG/boatsmiddle.png

Figure 5. Harbor activity monitoring. ©CNES

6.2.3. Wrinkle detection using a marked point process

Participants: Seong-Gyun Jeong, Yuliya Tarabalka, Josiane Zerubia [contact].

Skin image processing, wrinkle detection, line detection, marked point process, RJMCMC

We developed a novel wrinkle detection algorithm using a marked point process (MPP). Since wrinkles are the most important visual features of aging, automatic wrinkle detection algorithm can have many applications, such as the evaluation of cosmetic products, age estimation, and aging synthesis. In order to detect wrinkles of arbitrary shape, we represent wrinkles as a set of small line segments. Note that each line segment consists of a length and an orientation. A stochastic wrinkle model density exploits the local edge profile and constrains the spatial placement of adjacent lines. To maximize the model density, we employ a reversible jump Markov chain Monte Carlo (RJMCMC) sampler. A state of the Markov chain corresponds to a wrinkle configuration, and it is updated according to the acceptance ratio of sub-transition kernels: line segment births and deaths, and an affine transformation kernel. The transition kernels perturb the Markov chain by adding, removing, or modifying a wrinkle segment in the current configuration. In addition, an acceleration scheme has been developed for the RJMCMC sampler that enforces the connectivity of line segments. RJMCMC with acceleration reduces mixing time and improves detection accuracy as well.

Figure 6 compares wrinkle detection results simulated by random walk and the proposed acceleration scheme. The proposed algorithm faithfully detects wrinkles as smoothly connected lines. In addition, Figure 6 (d) plots the energy as a function of the number of iterations. It shows that the proposed acceleration method reaches a lower energy more rapidly than the random walk method.

6.3. Shapes and contours

6.3.1. Shape reconstruction from lidar data

Participant: Ian Jermyn [contact].

This work is being done in collaboration with Dr. Stuart Jones, Dr. Jochen Einbeck, and PhD student Thomai Tsiftsi of Durham University, UK [https://www.dur.ac.uk].

sand body, petroleum, shape, submanifold,

The cross-sectional shapes of 'sand bodies', ancient underground river channels filled with sediment, are of great interest in geology, and to the petroleum industry, because the shape is strongly correlated with the nature of the sediment, and in particular with its porosity, which in turn helps determine the volume fraction of crude oil contained in the sand body. The geological literature, however, only discusses simple characterizations of these shapes, and there is much room for improvement. This project aims to build probabilistic models of the cross-sectional shapes of sand bodies based on lidar point cloud data gathered from surface-projecting sand bodies by geologists in the field. Such models, when built, can be used to test the current geological classification of sand bodies, to generate new and geologically relevant classes, and to build functional models of the connection between sand body shape and oil yield.

Current work is focused on extracting reliable cross-sectional shapes from the lidar data (see Figure 7), a difficult task in itself since the sand bodies are frequently occluded or otherwise incomplete. Bayesian inference based on parameterized models of shape suggested by the current geological classification are used for this purpose. Since sand body shapes are concentrated near a low-dimensional submanifold of shape space, these models will later be extended using techniques such as mixtures built on principal curves, adapted to curved manifolds, in order to find and characterize this submanifold.

6.3.2. Riemannian metrics on spaces of curves and surfaces

Participant: Ian Jermyn [contact].

This work is being done in collaboration with Prof. Anuj Srivastava of Florida State University [http://www.fsu.edu].

Shape, Riemannian metric, elastic, curve, surface, functional data, alignment

../../../../projets/ayin/IMG/wrinkle_detection.png

Figure 6. Comparison of wrinkle detection results using different simulation procedures: (b) random walk and (c) the proposed acceleration scheme. Energy as a function of the number of iterations is plotted in (d).

../../../../projets/ayin/IMG/sandbodies.png

Figure 7. Left: a point cloud containing a sand body extracted from a larger cloud. Right: cross-sectional shape derived from the point cloud.

Statistical shape modelling has many applications in image processing and beyond. One of the key problems in this area is to develop and understand measures of shape similarity. One approach uses Riemannian metrics defined on ‘shape space’, the quotient of spaces of sphere or disc embeddings by similarities or other geometric group, and the diffeomorphism group of the sphere or disc. These metrics are defined by Riemannian metrics on the space of embeddings on which the transformation groups act by isometries, and so attention is focused on understanding such metrics and their properties.

Current work is focused on two areas. The first is on classifying and describing the diffeomorphism-invariant metrics on function spaces (shapes in one dimension) that satisfy additional desiderata useful in different applications, with particular application to function alignment. The second is on generalizing to surfaces the elastic metric much used in the case of curves, and in finding surface representations that permit analytic results to be derived, or that simplify computations, in the same way that the square root velocity representation simplifies computations involving the elastic metric on curves (see Figure 8).

../../../../projets/ayin/IMG/surfacegeodesic.png

Figure 8. Top: interpolating surfaces based on a previous Riemannian metric. Bottom: interploating shapes based on the generalized elastic metric.

6.3.3. Sampling methods for random field models of shape

Participant: Ian Jermyn [contact].

Part of this work is being done in collaboration with Prof. Zoltan Kato and PhD student Csaba Molnar of the University of Szeged, Hungary [http://www.inf.u-szeged.hu/~kato/], and part in collaboration with PhD student Michael Racovitan of Durham University, UK [https://www.dur.ac.uk].

Shape, long range interaction, Markov random field, phase field, contour, learning, wavelet

The detection and segmentation of objects from images is a problem with innumerable applications in many domains. Probabilistic models of shape, used as prior distributions in the inference process, are a necessity in solving any nontrivial instance of this problem. In many cases of importance, the shapes to be modelled cannot be treated efficiently, or at all, with current techniques, for example when multiple instances of an object must be segmented. The overall goal of this project is to develop a general shape modelling methodology capable of dealing with these difficult cases, as well as more traditional instances of the problem.

Recent algorithmic work has focused on developing efficient sampling methods for the models, for use in parameter and model learning. The models, whether expressed in terms of shape boundaries, phase fields, or binary fields, contain many long-range frustrated interactions, and hence are not amenable to standard techniques. Simplifications of the interaction structure using adapted wavelet bases, and re-expressions of the models using varieties of Hubbard-Stratanovich transformation are two directions being explored.

../../../../projets/ayin/IMG/sampleresult.png

Figure 9. A typical result on an image of lipid cells.

6.3.4. Multiple-instance object detection via a third-order active contour shape model

Participants: Ikhlef Bechar, Ian Jermyn, Josiane Zerubia [contact].

This work was funded by the EADS Foundation [http://www.fondation.eads.com].

Object detection, multiple objects, shape, invariance, prior, higher-order active contour (HOAC), energy minimization

Recent modelling work has focused on generalizing the higher-active contour methodology to families of shapes whose members consist of an arbitrary number of object instances, each of which is similar to a given reference shape. This means finding energies on the space of regions that possess low-energy local minima corresponding to an arbitrary number of instances of the reference shape. To this end, we have studied a family of fourth-order energy functionals on regions based on a kernel given in closed form as a function of the reference region. The energy has, amongst its global minima, regions consisting of an arbitrary number of well-separated instances of the reference shape, each under an arbitrary Euclidean transformation, thereby eliminating the need to estimate group-valued 'pose' parameters. It may be combined with a likelihood energy, and the result minimized using gradient descent, speeded up by use of the Fourier domain. Although problems still remain, a series of experiments on both synthetic and real images has demonstrated the feasibility of the approach (see Figure 10).

../../../../projets/ayin/IMG/shape_prior.png

Figure 10. Detection of a shape in a noisy infrared image (SNR = 5dB): (top row) without using prior shape knowledge, and (bottom row) using the proposed fourth-order prior shape model. First column: initialization; second column: intermediate contour; third column: final contour; fourth column: segmentation.

6.4. Shapes in time

6.4.1. Graph-based model for multitemporal segmentation of sea ice floes from satellite data

Participants: Claudio Price González, Yuliya Tarabalka [contact].

This work has been done in collaboration with Dr. Ludovic Brucker (NASA GSFC, USA) [http://www.nasa.gov].

Multitemporal segmentation, region growing, MODIS, sea ice floes

Automated segmentation of the evolution of sea ice from satellite images would allow scientists studying climate change to build accurate models of the sea ice meltdown process, which is a sensitive climate indicator. In this work, we proposed a new method which uses shape analysis and graph-based optimization to segment a multiyear ice floe from time series of satellite images [13]. The new approach combines data from two instruments onboard the NASA Aqua satellite, enabling several measurements per day over the Earth's polar regions: Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E); and Moderate-Resolution Imaging Spectroradiometer (MODIS). The method performs best merge region growing, followed by energy minimization on the image graph, where the energy consists of two terms describing the floe shape (shape term) and the gradient between the floe and the background (data term), respectively. We validated the performance of the proposed method for segmentation of a shrinking ice floe from a sequence of AMSR-E and MODIS images acquired in August–October 2008 (see Figure 11). The results obtained showed both the effectiveness of the proposed approach and its robustness to low-contrast data.

6.4.2. Enforcing monotonous shape growth or shrinkage in video segmentation

Participant: Yuliya Tarabalka [contact].

This work has been done in collaboration with Dr. Guillaume Charpiat (STARS team, Inria-SAM), Dr. Bjoern Menze (Computer Vision Laboratory at ETH Zurich and Asclepios team at Inria-SAM), and Dr. Ludovic Brucker (NASA GSFC, USA) [http://www.nasa.gov].

Video segmentation, graph cut, shape analysis, shape growth

Automatic segmentation of objects from video data is a difficult task, especially when image sequences are subject to low signal-to-noise ratio or low contrast between the intensities of neighboring structures. Such challenging data are acquired routinely, for example, in medical imaging or satellite remote sensing. While individual frames can be analyzed independently, temporal coherence in image sequences provides a lot of information not available for a single image. In this work, we focused on segmenting shapes that grow or shrink monotonically in time, from sequences of extremely noisy images.

We proposed a new method for the joint segmentation of monotonically growing or shrinking shapes in a time sequence of images with low signal-to-noise ratio [15]. The task of segmenting the image time series is expressed as an optimization problem using the spatio-temporal graph of pixels, in which we are able to impose the constraint of shape growth or shrinkage by introducing unidirectional infinite-weight links connecting pixels at the same spatial locations in successive image frames. The globally-optimal solution is computed with a graph-cut algorithm. The performance of the proposed method was validated on three applications: segmentation of melting sea ice floes; of growing burned areas from time series of 2D satellite images [16]; and of a growing brain tumor from sequences of 3D medical scans. In the latter application, we imposed an additional inter-sequences inclusion constraint by adding directed infinite-weight links between pixels of dependent image structures. Figure 12 shows a multi-year sea ice floe segmentation result. The proposed method proved to be robust to high noise and low contrast, and to cope well with missing data. Moreover, it showed linear complexity in practice.

6.5. Other detection approaches

6.5.1. Illumination modeling and chromophore identification in dermatological images for skin disease analysis

Participants: Zhao Liu, Josiane Zerubia [contact].

../../../../projets/ayin/IMG/Claudio.png

Figure 11. Comparison of results for the MODIS image sequence acquired in August-October 2008. Manual segmentation of the ice floe contour is shown in green, hierarchical step-wise optimization result in red, and the new graph-based approach in blue.

../../../../projets/ayin/IMG/GC.png

Figure 12. Top: MODIS images for four time moments (days 230, 233, 235 and 267 of 2008, respectively). Bottom: corresponding aligned images with segmentation contours (in red). Manual segmentation is shown in green.

This work is part of the LIRA Skin Care Project, which includes four key partners: Philips R&D [http://www.research.philips.com], CWI (Netherlands) [http://www.cwi.nl], Inria (France), and Fraunhofer Institutes (Germany) [http://www.fraunhofer.de/en.html].

Chromophore identification, illumination modeling, skin disease analysis, dermatology

Skin color is an important characteristic for the accurate diagnosis and grading of cutaneous lesions by experienced dermatologists in clinical practice. However, the visual perception of skin color is not only a function of the major chromophores (melanin and hemoglobin) underneath the skin surface, but is also affected by external illumination and the spectral responses of imaging detectors. Skin color representation in a specific color space (e.g. RGB and its transformations) is not a genuine physical quantity. It sometimes fails to provide precise information about the concentrations of cutaneous chromophores, and is easily influenced by external imaging factors. As a result, conventional colorimetry may not properly describe the underlying histological content of skin, and hence tends to yield less trustworthy results when applied directly for skin disease analysis.

Building on a previous study that considered human skin as a diffuse reflectance surface, our work models human skin as having specular and diffuse reflectance, leading to a novel illumination correction method. Based on this method, we have developed a new scheme for chromophore identification from dermatological photographs. The algorithm has three steps. First, specular reflectance is separated from diffuse reflectance in the original skin images through specular pixel localization and image interpolation using a nonlinear weighted averaging process. Second, the resultant diffuse reflectance component is decomposed into a base layer and a detail layer. The base layer, representing low-frequency illumination and shading effects, is approximated by polynomial curve fitting using an initial illumination map using an adaptive bilateral filter as a prior. The detail layer, primarily containing high-frequency chromophore reflectance, can then be calculated by subtracting the base layer from the corresponding diffuse spectral band in logarithmic form. Finally, by incorporating knowledge of chromophore absorption characteristics, melanin and hemoglobin densities are identified using the detail layers from different spectral channels [11].

For algorithm evaluation, the method was applied to two skin disease analysis problems: computer-aided melanoma diagnosis [11] and automatic acne detection [12]. For melanoma diagnosis, 201 conventional RGB skin lesion images (62MMs, 139 benign nevi (BN)) were collected from free public databases (http://www.dermquest.com/, http://www.dermis.net/) to form an experimental data set. Figure 13 -(I) shows an example of a superficial spreading melanoma with obvious horizontal shading effects, and the corresponding experimental results. It is clear that the proposed algorithm successfully removed the imaging artifacts from the original skin lesion photographs.

For acne detection, a set of 50 challenging images were tested as a qualitative evaluation to demonstrate the usefulness of the proposed method. Automatic acne segmentation is performed using an MRF model based on chromophore descriptors. Figure 13 -(II) shows one acne example captured in an uncontrolled environment from a free public database (http://www.dermnetnz.org/). The detected acne areas are highly consistent under visual inspection, and the inflammatory acne can be distinguished from hyperpigmentation by comparing the average values of the melanin and hemoglobin indices.

../../../../projets/ayin/IMG/Joint.jpg

Figure 13. Examples of different types of skin disease requiring shading removal and chromophore identification: (I-a) Original melanoma image; (I-b) Corrected melanoma image; (I-c) Melanin index map of image (I-a); (I-d) Hemoglobin index map of image (I-a); (I-e) Two-class segmentation results from Otsu's method on the original melanoma image (black line) and the corrected melanoma image (blue line), respectively; (II-a) Original acne image; (II-b) Melanin index map of image (II-a); (II-c) Hemoglobin index map of image (II-a); (II-d) Acne segmentation result using an MRF model, highlighting inflammatory acne (blue line) and hyperpigmentation (black line), respectively.

BACCHUS Team

6. New Results

6.1. Residual distribution schemes for steady problems

Participants: Rémi Abgrall [Corresponding member], Mario Ricchiuto, Dante de Santis, Algiane Froehly, Cécile Dobrzynski, Pietro Marco Congedo.

We have understood how to approximate the advection diffusion problem in the context of residual distribution schemes. The third order version of the schemes has been validated for both laminar and turbulent flows. It is uniformly accurate with respect to the local Reynolds number. The turbulent version makes used of extension to the Spalart Allmaras model. We have studied the (iterative) convergence issues using Jacobian Free techniques or the LUSGS algorithm. Tests in two and three dimensions have been carried out. This work is submitted to J.Comput.Phys. We are now able to handle two and three dimensional laminar and turbulent flows on hybrid and high order (curved boundaries) meshes. Moreover, we have extended the scheme to the use of complex equations of state, and we perform high-order computations with real-gas effects. This work is submitted to Computers&Fluids.

6.2. Curved meshes

Participants: Rémi Abgrall, Cécile Dobrzynski [Corresponding member], Algiane Frohely.

One of the main open problems in high order schemes is the design of meshes that fit with enough accuracy the boundary of the computational domain. If this curve/surface is not locally straight/planar, the elements must be curved near the boundary, and their curvature need to be propagated to the interior of the domain to have valid elements. When the mesh is very stretched, this can be quite challenging since, in addition, we want that the mesh keep a structure, in particular for boundary layers. Using tools explored in iso-geometrical analysis, we have been able to construct a software computing curved meshes (in 2D and 3D), while keeping as much as possible the structure of the mesh and guaranteeing that the generated mesh is suitable to CFD simulation (all elements have a positive Jacobian). This software is being used for high order computations with the IDIHOM project. The full paper has been accepted in IJNMF and will be published in 2014.

6.3. Hypoelastic models

Participants: Rémi Abgrall [Corresponding member], Pierre-Henri Maire.

In collaboration with CEA (P.H. Maire), we have developed and tested a new finite volume like algorithm able to simulate hypo-elastic and plastics problems on unstructured meshes. This has been published in [19].

6.4. Penalisation methods using unstructured meshes

Participants: Rémi Abgrall, Heloise Beaugendre [Corresponding member], Cécile Dobrzynski, Leo Nouveau, Quentin Viville.

In Computational Fluid Dynamics the interest on embedded boundary methods for Navier-Stokes equations increases because they simplify the meshing issue, the simulation of multi-physics flows and the coupling of fluid-solid interactions in situation of large motions or deformations. Nevertheless an accurate treatment of the wall boundary conditions remains an issue of these methods. In this work we develop an immersed boundary method for unstructured meshes based on a penalization technique and we use mesh adaption to improve the accuracy of the method close to the boundary. The idea is to combine the strength of mesh adaptation, that is to provide an accurate flow description especially when dealing with wall boundary conditions, to the simplicity of embedded grids techniques, that is to simplify the meshing issue and the wall boundary treatment when combined with a penalization term to enforce boundary conditions. The bodies are described using a level-set method and are embedded in an unstructured grid. Once a first numerical solution is computed mesh adaptation based on two criteria the level-set and the quality of the solution is performed. The full paper will be published in the Journal of Computational Physics in january 2014.

6.5. Unsteady problem

Participants: Rémi Abgrall, Mario Ricchiuto [Corresponding member].

A higher order version of the explicit multi-stage RD schemes we have designed has been obtained in one dimension, and its extension to two space dimensions is in the works. A moving mesh ALE formulation of the multistage explicit schemes developed [58] (paper submitted to J.Sci.Comp.) as a basis for adaptive mesh movement, in development in collaboration with Pr. A. Guardone. We have also started work on new formulations based on different time stepping schemes of the multi-step type.

Concerning implicit schemes, the work on higher order space time formulations in collaboration with the Leeds university and with A. Larat of Ecole Centrale Paris. The advantage of this formulation in terms of efficiency has been shown for shallow water problems [24], while the extension to higher than second order is still in development (Inria RR-7843).

6.6. Uncertainty Quantification

Participants: Rémi Abgrall, Pietro Congedo [Corresponding member], Gianluca Geraci, Maria Giovanna Rodio, Kunkun Tang, Julie Tryoen, Mario Ricchiuto, Thierry Magin.

We developed an unified scheme in the coupled physical/stochastic space. Basing on the Harten multiresolution framework in the stochastic space, we proposed a method allowing an adaptive refinement/derefinement in both physical and stochastic space for time dependent problems (aSI scheme). As a consequence, an higher accuracy is obtained with a lower computational cost with respect to classical non-intrusive approaches, where the adaptivity is performed in the stochastic space only. Performances of this algorithm are tested on scalar Burgers equation and Euler system of equations, comparing with the classical Monte Carlo and Polynomial Chaos techniques [6], [7]. We have also coupled the aSI scheme withe the DEM method for building an accurate stochastic scheme for multiphase flows. A paper is submitted to the Journal of Computational Physics on this topic.

Concerning non-intrusive methods, we proposed a formulation in order to compute the decomposition of high-order statistics. The idea is to compute the most influential parameters for high orders permitting to improve the sensitivity analysis. Second objective is to illustrate the correlation between the high-order functional decomposition and the PC-based techniques, thus displaying how to compute each term from a numerical point of view. This method has been proposed in both classical and Anchored ANOVA representation. Two papers are actually under revision on this topic. Moreover, a bayesian-based method has been used within a Polynomial Chaos framework for rebuilding the freestream conditions, starting from wall measurements during the atmospheric reentry of a space vehicle. See [16] for more details. Moreover, an uncertainty propagation method has been applied to the robust analysis of cavitating flows in a Venturi tube, displaying very interesting results concerning the influence of inlet conditions and the multiphase model parameters (see[23] for more details).

Uncertainty propagation studies are actually underway for assessing the influence of boundary conditions and model parameters for the simulation of a tsunami.

6.7. Robust Design Optimization

Participants: Pietro Congedo [Corresponding member], Gianluca Geraci, Gianluca Iaccarino.

The Simplex-Simplex approach, that has been proposed in 2011, has been further developed. In particular, the algorithm has been improved yielding an evolved version of the Simplex2 approach, and the formulation has been extended to treat mixed aleatory/epistemic uncertainty. The resulting SSC/NM (Simplex Stochastic Collocation/Nelder-Mead) method, called Simplex2, is based on i) a coupled stopping criterion and ii) the use of an high-degree polynomial interpolation of the optimization space. Numerical results show that this method is very efficient for mono-objective optimization and minimizes the global number of deterministic evaluations to determine a robust design. This method is applied to some analytical test cases and to a realistic problem of robust optimization of a multi-component airfoil (see [17] for more details).

Morevoer, we proposed a strategy for multi-objective robust design optimization, with a stochastic dimension reduction based on ANOVA analysis. The developed strategy has been applied to the robust optimization of dense-gas turbines (see [15] for more details).

6.8. Multiphase flows

Participants: Rémi Abgrall [Corresponding member], Pietro Congedo, Maria-Giovanna Rodio.

We developed the numerical solver based on a DEM formulation modified for including viscous effects and a more complex equation of state for the vapor region. The method used is the DEM for the resolution of a reduced five equation model with the hypothesis of pressure and velocity equilibrium , without mass and heat transfer. This method results in a well-posed hyperbolic systems, allowing an explicit treatment of non conservative terms, without conservation error (see [8] for more details). The DEM method directly obtains a well-posed discrete equation system from the single-phase conservation laws, producing a numerical scheme which accurately computes fluxes for arbitrary number of phases. We considered two thermodynamic models , i.e. the SG EOS and the Peng-Robinson (PR) EOS. While SG allows preserving the hyperbolicity of the system also in spinodal zone, real-gas effects are taken into account by using the more complex PR equation. The higher robustness of the PR equation when coupled with CFD solvers with respect to more complex and potentially more accurate multi-parameter equations of state has been recently discussed. In this paper, no mass transfer effect is taken into account, thus the PR equation can be used only to describe the vapor behavior, while only the SG model is used for describing the liquid [22].

Another topic covered by Bacchus is about the numerical approximation of non conservative systems. One very interesting example is obtained by the Kapila model, for which shock relations can be found from physical principles. Most, if not all, the know discretisation are at best stable, but do not converge under mesh refinement. We have proposed a way to do so by using some modifications of a Roe-like solver.

6.9. Depth averaged free surface modeling

Participants: Mario Ricchiuto [Corresponding member], Philippe Bonneton, Andrea Filippini, Stevan Bellec.

We have improved the modeling capabilities of our codes by an efficient implementation of residual based dicretizations of a non-hydrostatic enhanced Boussinesq system [21]. in particular, we have demonstrated how residual based stabilization terms do not pollute the accuracy of the underlying centered discretization, and lead very low dispersion error, while allowing to handle in a stable manner the hyperbolic (hydrostatic) limit. In the framework of the internship of P. Bagicaluppi, this has been used to construct a non-oscillatory model including wave breaking effects (paper in preparation).

In parallel, we have started an in depth study of the improvement of the dispersion operators, which control the position and height of the waves. This has allowed to highlight the existence of new form of weakly nonlinear models [62]. A paper is in preparation.

6.10. Self healing composites modeling

Participants: Mario Ricchiuto [Corresponding member], Gérard Vignoles, Gregory Perrot.

This year we have started the coupling between `COCA` and the structural solver of the LCTS lab. The coupling is done for the moment using simple scripting, and providing the structural solver with an equivalent fiber-surface in contact with oxygen at a given time. A simplified potential flow model (classical source potential) for the oxide has also been developed and is being tested.

6.11. Parallel remeshing

Participants: Cécile Dobrzynski, Cédric Lachat, François Pellegrini [Corresponding member].

All the work of the elapsed year on `PaMPA` concentrated on the design and implementation of parallel remeshing algorithms (see Section 5.6 for more details about the software itself). These algorithms are based on several steps: (i) identification of the areas to remesh; (ii) splitting of these areas into zones of prescribed size and/or estimated workload; (iii) redistribution and centralization of as many zones as possible on the processors; (iv) sequential remeshing of the zones; (v) reintegration of the zones to their original locations; (vi) identification of the remaining areas and loop to step (ii) when work remains.

Several splitting algorithms have been designed and evaluated, so as to provide zones with adequate aspect ratios to the sequential partitioners. Load imbalance is still a concern, since zones must not be too small, while they must be numerous enough so as to maximize concurrency across all of the available processors.

As of December, `PaMPA` has been able to remesh a coarse mesh of 27 millions of tetraedra up to a fine mesh comprising more than 600 millions of tetraedra, in 34 minutes, on 240 processors of the Avakas cluster at MCIA Bordeaux, using the `MMG3D` sequential remesher. Remeshing up to a finer mesh of above one billion of elements is the next milestone to reach, to evidence the capabilities of the software.

Cédric Lachat defended his PhD last December. A first abstract has been submitted, and two more journal papers are in preparation.

Figure 4. Cut of a 3D cube made of tetrahedra showing the effect of parallel remeshing by `PaMPA`.

6.12. Graph remapping

Participants: Sébastien Fourestier, François Pellegrini [Corresponding member].

The work on remapping mostly took place in the context of the PhD of Sébastien Fourestier, who defended last June. This work concerned the coding and evaluation of the parallel graph repartitioning and remapping algorithms that were designed last year. Indeed, the sequential version of these algorithms had been integrated in version `6.0` of `Scotch`, released in the beginning of December 2012. The implementation of the parallel algorithms, which started last year, took place in the `6.1` branch of `Scotch`, to be released once the `6.0` branch is made stable.

The evaluatin of our algorithms showed that the diffusion-based optimization algorithm, which behaves well in the context of partitioning, exhibits an unwanted behavior when adapted to the repartitioning and remapping cases. Typically, when the mapping changes too much, external constraints to the flows that represent the different parts may prevent them from meeting, thus reducing the quality of the frontier they should create by flooding one against the others. These algorithm have to be improved.

A journal paper summing-up all the work done during the past years in the context of process mapping, within the Joint Laboratory for Petascale Computing (JLPC) between Inria and UIUC, has been submitted.

6.13. Sparse matrix reordering for ILU solvers

Participants: Astrid Casadei [HIEPACS project-team, Inria Bordeaux - Sud-Ouest], Sébastien Fourestier, François Pellegrini [Corresponding member].

In the context of ANR `PETALh`, our task is to find ways of reordering sparse matrices so as to improve the robustness of incomplete LU factorization techniques. The path we are following is to favor the diagonal dominance of the matrices corresponding to the subdomains of the Schur complement. Our studies aim at injecting some information regarding off-diagonal numerical values into nested dissection like reordering methods, so as to favor the preservation of high off-diagonal values into either the subdomains or the separators of Schur complement techniques.

The experimental framework had been set-up last year. It consisted in a modified version of the `Scotch` sparse matrix ordering library for computing orderings and of the `HIPS` iterative sparse linear system solver for evaluating them. The text cases used were provided by the industrial partners of the `PETALh` project.

In order to improve diagonal dominance, several cut-off methods have been proposed in order to carve the matrix pattern and speed-up computations towards convergence. These cut-off methods were based on either linear or logarithmic scales, with cut-off values selected according to various distributions.

While some of these methods improve convergence on some restricted classes of matrices, as our first experiments showed last year, no method was able to provide overall improvement on a wide range of matrices. This research path is consequently considered as inefficient. A research report has been written.

6.14. Numerical methods for high altitude aerodynamics and rarefied gas flows

Participants: Luc Mieussens [Corresponding member], Florent Pruvost [IMB, engineer], G. Dechristé [IMB, PhD], N. Hérouard [CEA-CESTA, PhD], Stéphane Brull [IMB], L. Forestier-Coste [IMB, Post Doc].

This activity involves many developments for rarefied gas flow simulations for very different applications, and the design of numerical schemes for high altitude aerodynamics based on some kinetic model :

- the simulation code `CORBIS` (rarefied gases in 2 space dimensions on structured meshes) has been re-engineered : modular form, use of the `git` version control system, modification to use unstructured meshes, MPI/OpenMP hybrid parallelization. Very good performance in terms of scalability and efficiency have been obtained, up to 700 cores.
- a new method to generate locally refined grids in the space of velocities has been proposed and shown to provide CPU time gains of the order of 30 (w.r.t the existing approach). This work has been published in (Baranger *et al., J. Comput. Phys* 257(15), 2014) ;
- the second order Discontinuous Galerkin method has been shown to be more accurate and faster than higher order finite volume methods (up to fourth order) for one-dimensional rarefied gases problems. We have analytically proved that this method is Asymptotic Preserving for the Stokes regime ;
- a new kinetic model for multispecies reacting flows for re-entry applications has been proposed. In this model, the mixture oxygen-nytrogen is described by a kinetic equation, while the minor species (O, NO, N) are described by reaction diffusion equations. The implementation of this model in a full 3D code is under way ;

- we have presented one of the first numerical simulation of the Crookes radiometer ever. This has been obtained with a Cartesian grid approach, with a cut-cell techniques allowing a simplified treatment of moving solid boundaries. This work has been published in the proceedings of the 28th Symposium on rarefied Gas Dynamics ;
- We have proposed a new method to discretize kinetic equations based on a discretization of the velocity variable which is local in time and space. This induces an important gain in term of memory storage and CPU time, at least for 1D problems (this work has been presented in a paper submitted for publication). Two-dimensional extensions are under development ;
- We have shown that the recent method “Unified Gas Kinetic Scheme”, proposed by K. Xu to simulate multi-scale rarefied gas flows, can be extended to other fields, like radiative transfer. This approach, based on a simple finite volume technique, is very general and can be easily applied to complex geometries with unstructured meshes. This work has been published in (Mieussens, *et al., J. Comput. Phys* 253(15), 2013).

BAMBOO Project-Team

6. New Results

6.1. Symbiont genome evolution and dynamics

The objective of this part of our work was to analyse genome rearrangements and dynamics. The results obtained were both algorithmic and biological.

In terms of algorithms, we developed a new method for repeat identification (RIME) [12], as well as an algorithm for finding the minimum number of three constrained versions of inversions that transform one given genome into another [25]. The constrained versions concerned symmetric, almost-symmetric and unitary inversions. The genome rearrangement algorithm is not exact: it is based on a greedy randomized search procedure to find such minimum number of constrained inversions.

The main set of biological results [4], [14] concerned trypanosomatids of the genera *Angomonas* and *Strigomonas* that live in a mutualistic association characterised by extensive metabolic cooperation with obligate endosymbiotic Betaproteobacteria. In contrast to their counterparts lacking symbionts, such trypanosomatids exhibit lower nutritional requirements and are autotrophic for essential amino acids and vitamins. Phylogenetic analyses showed that the cooperation in the first case is complemented by multiple horizontal gene transfers, from bacterial lineages to trypanosomatids, that appear to have occurred several times in the course of evolution. In contrast, but for three exceptions, such transfers are absent as concerns vitamin biosynthesis.

The above work was made possible in part because of the sequencing and annotation of the genomes whose metabolic pathways could then be inferred. We participated in these for some of the genomes involved in the above study [17].

6.2. Host-symbiont metabolic dialog

The methodological work done has covered one main question concerning what we called metabolic stories. Given a subset of metabolites representing those monitored as being under- or over-produced in some condition (e.g., interaction with a parasite) and a metabolic network represented as a compound graph, metabolic stories are maximal directed acyclic graphs (DAGs) that cover all the metabolites in the subset of interest, and have all sources and targets among these metabolites. One exact algorithm (TOUCHÉ, [24]) was developed to enumerate all metabolic stories that improved on our previous method (GOBBOLINO).

The algorithm above was validated on biological data [16] in a study of the response of yeast to cadmium exposure. We used this system as a proof of concept for our method and we showed that we are able to find a story that reproduces very well the current knowledge about the yeast response to cadmium. We further showed that this response is mostly based on enzyme activation. We also provided a framework for exploring the alternative pathways or side effects this local response is expected to have in the rest of the network. Finally, we discussed several interpretations for the changes we see and we suggest hypotheses which could in principle be experimentally tested.

6.3. Host-symbiont genetic dialog

Two sets of problems were addressed: (i) the development of algorithms for analysing NGS data especially RNAseq, and (ii) the development of algorithms for identifying small RNAs, notably microRNAs, and their targets.

The computational work on NGS is described in another section.

Computational work on small RNAs, initially miRNAs, led to the development of a new algorithmic method. This builds upon previously developed approaches, one which was applied to *Anopheles darlingi* for inferring miRNAs that however had a high rate of false positives, and a second that provided a way for navigating among all the candidates found. Recently however, we arrived at a better model for such inference in the double sense that the rate of false positives is smaller without losing in sensitivity, while the method is much faster. The paper presenting this work and the algorithm (MIRINHO) was submitted and is currently in revision.

6.4. Symbiont-host co-cladogenesis and co-evolution at the sequence and network levels

The problem here was to: (i) study the co-evolution of a set of hosts and their symbionts, and (ii) to understand the genetic architecture of a parasitic invasion by investigating the different phenotypes such invasion produces in the host.

Work on the first point took longer than initially planned but two papers are now submitted. In the first, titled “Co-phylogeny Reconstruction via an Approximate Bayesian Computation", we describe an algorithm (COALA) for estimating the frequency of co-evolutionary events based on a likelihood-free approach. The benefits of this method are twofold: (1) it provides more confidence in the set of costs to be used in a reconciliation, and (2) it allows to estimate the frequency of the events in cases where the dataset consists of trees with a large number of taxa. We evaluate our method on simulated and on real datasets. We show that in both cases, for a same pair of host and parasite trees, different sets of frequencies for the events constitute equally probable solutions. Moreover, sometimes these sets lead to different parsimonious optimal reconciliations, in the sense of presenting a different number of the events. For this reason, it appears crucial to take this into account before attempting any further biological interpretation of such reconciliations. More generally, we also show that the set of frequencies can vary widely depending on the input host and parasite trees. Indiscriminately applying a standard vector of costs may thus not be a good strategy.

In the second submitted paper related to the study of co-evolution and titled “EUCALYPT: Efficient tree reconciliation enumerator", we present a polynomial-delay algorithm for enumerating all optimal reconciliations. We show that in general many optimal solutions exist. We give an example where, for two pairs of host-parasite trees having each less than 40 leaves, the number of solutions is 2309, even when only time feasible solutions are kept. To facilitate their interpretation, those solutions are also classified in terms of the number of each event that they contain. This often enables to reduce considerably the number of different classes of solutions to examine further, but the number may remain high enough (16 for the same example). Depending on the cost vector, both numbers may increase considerably (for the same instance, to respectively 4080384 and 275).

Concerning the second question (genetic architecture of a parasitic invasion), one such phenotype is called “cytoplasmic incompatibility” (CI). Briefly, when a parasite invades a male host, it induces the death of the host’s offspring unless the female is also infected. This has been explained by a toxin/antitoxin model that involves a toxin deposited by the parasites in the male’s sperm inducing the death of the zygote unless neutralised by an antidote produced by the parasites in the egg. One toxin/antitoxin pair is usually linked to one genetic factor. Given a set of observed CIs, the question is how many genetic factors explain it. In its simplest form, this mathematically translates into, given a bipartite graph, finding its minimum biclique edge cover. One biclique corresponds to one factor. We had previously analysed the complexity of the problem and proposed an algorithm that was this year applied to a set of CI data from *Culex pipiens* [18].

6.5. NGS for biodiversity

In collaboration with the Laboratoire d’Écologie Alpine (LECA) at Grenoble where there is a strong expertise on DNA meta-barcoding, we had devised several tools for barcode design and analysis. ECOPRIMERS thus identified new barcode markers and their associated PCR primers within a DNA meta-barcoding approach. The algorithm was optimised two quality indexes measuring taxonomical range and discrimination to select the most efficient markers from a set of reference sequences, according to some experimental constraints such as marker length or specifically targeted taxa. We had also devised assembler algorithms directed to organelles (mitochondria or chloroplasts). This year, in collaboration with the Inria project-team MISTIS, we developed a statistical modelling approach to investigate the spatial cross-correlations between different taxa identified by meta-barcoding of soil sample from French Guiana (this was selected as a conference paper at the "45ème Journées de Statistiques" 2013 that took place at Toulouse and is organised by the Société Française de Statistique. This approach allows to visualise the co-occurrence pattern as a "species interaction graph", and to study the mutual exclusion (competition) or inclusion (symbiosis) of different plant species.

6.6. NGS for genotypic variation detection

The computational work on NGS data concerned both algorithmic design and complexity analysis.

Based on the idea that each genotypic variation will correspond to a recognisable pattern in a de Bruijn graph constructed from a set of sequence reads, we had proposed a generic model for SNPs in DNA data, and then generalised it to the analysis of RNA. In this case, not only SNPs are present but also alternative splicing (AS) events, which, once again, generate a recognisable pattern in the de Bruijn Graph. We had therefore proposed a general model for all these variations (SNPs, indels and AS events) and introduced an exact algorithm (KISSPLICE) to extract all alternative splicing events. The algorithm also outputs candidate SNPs and indels. This year, we improved the algorithm [26]. As the problem relates to an old one in algorithmics (cycle enumeration), we also revisited it from a theoretical point of view [23].

The improved version of KISSPLICE [26] was used to analyse RNAseq data from two lines of *Asobara tabida* exhibiting different ovarian phenotypes in the absence of its endosymbiont *Wolbachia*. Although infected individuals of the two lines have similar phenotypes, numerous genes are differentially expressed between the two infected conditions. This could mean that two divergent strategies of tolerance have evolved. Preliminary results on the analysis of polymorphisms between these two lines suggest that differentially expressed genes tend to accumulate more variation. We are currently, via experiments done by the biologists in our team, testing the hypothesis that such genes are under strong selection pressure and may evolve through mutation accumulation, a process that could be related to assimilation.

A preliminary analysis of human data from the ENCODE project performed with KISSPLICE showed that an assembly-based method (without reference genome) is able to recover AS events that are missed by mapping-based methods (with a reference genome). Some of these events were experimentally validated, which represents the best type of proof we can provide to the biologists. The experimental part is made by our collaborator from the Inserm, Didier Auboeuf, in his team at the Centre National de Cancérologie of Lyon (CNCL), with whom we had an Inserm project, EXOMIC, funded for three years starting from 2012.

The identification of SNPs is also getting renewed interest even in the presence of a reference genome thanks to the possibility of re-sequencing many times the genome of a same or of very closely related species. The difficulty in the case of SNPs is to distinguish them from sequencing errors and from inexact repeats. We proposed a statistical test enabling to identify variations that are condition-specific, which enables to greatly enrich the list of potential SNP candidates. The paper on this test is in preparation. Its results as applied to the RNAseq data from two lines of *Asobara tabida* (see above) and to Drosophila species having diverged very recently were validated by, respectively, Fabrice Vavre and Cristina Vieira, both members of BAMBOO.

We also started addressing the problem that repeats (such as transposable elements for instance but not only) represent more in general for both local and global assemblers. We are thus developing a method that would enable to identify, in a de Bruijn graph built from RNAseq data, the vertices potentially corresponding to the borders of a repeated sequence. Preliminary results on simulated and real data show that the approach is promising (paper in preparation).

BANG Project-Team

6. New Results

6.1. Proliferation dynamics and its control

6.1.1. Proliferation dynamics in cell populations

Participants: José Luis Avila Alonso [DISCO project-team, Inria Saclay IdF], Annabelle Ballesta, Gregory Batt [CONTRAINTES project-team], François Bertaux, Frédérique Billy, Frédéric Bonnans [Commands project-team, Inria Saclay IdF], Catherine Bonnet [DISCO project-team, Inria Saclay IdF], Jean Clairambault, Marie Doumic, Xavier Dupuis [Commands project-team], Ján Eliaš, Germain Gillet [IBCP, Université Cl. Bernard Lyon 1], Pierre Hirsch [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], 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, Szymon Stoma [CONTRAINTES project-team], 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, after a rich harvest of publications (most of them in 2013 and even 2014), is awaiting new developments, since of the main two young researchers on this theme, F. Billy has concluded her 2-year Inria postdoc at Bang, leaving for an industrial company in November 2012, while 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 (J.L. Avila Alonso, C. Bonnet, Hitay Özbay, S. Niculescu), 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 January 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. In a book chapter summing up the PhD work of J.L. Avila Alonso [26], and in two submitted conference papers [28], [29], a new model of haematopoiesis for AML is presented, including phases of the cell division cycle and maturation stages, with targets for therapeutic control.
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 apoptosis.**

   With G. Gilllet (professor at IBCP/Lyon), A. Ballesta and M. Doumic have designed a mathematical ODE model for the mitochondrial pathway of apoptosis, focused on the early phase of apoptosis (before the cytochrome C release). This model has been validated by experimental data carried out in G. Gillet's lab and applied to propose new therapeutic strategies against cancer [6].
5. **Molecular model of the activity of the p53 protein.** This work, firstly the object of Luna Dimitrio's PhD thesis [37], who left in 2012 for the pharmaceutic industry (SANOFI), has been continued since a new PhD student, Ján Eliaš, has taken over this theme in September 2012 in a new PhD thesis at UPMC, under the supervision of J. Clairambault and B. Perthame. His work has given rise in 2013 to 2 publications [14], [32].

6. **TRAIL - induced apoptosis in HELA cells** Explaining cell-to-cell variability is a major step towards understanding how cancer cells escape action of chemotherapeutic drugs. We set up and studied an integrated model of stochastic gene expression, deterministic translation and protein degradation capable of explaining fractional killing and reversible resistance in Hela cells in response to treatment with TNF-Related Apoptosis Inducing Ligand, TRAIL (Bertaux, Stoma, Drasdo, and Batt, submitted). The results of the model suggests that stochastic fluctuations are a fundamental determinant in understanding cell-to-cell variability, and identified relations between the characteristic time scales of the processes at which stochasticity should play a particular important role.

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 [9], [27], [7], [8], [31]. Whereas transition kernels between cell cycle phases without control have been experimentally identified in cell cultures by FUCCI imaging [9], 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. A 12-year collaboration work with Francis Lévi on (circadian) chronotherapeutic optimisation in cancer is reported in [30].
2. *Intracellular pharmacokinetic-pharmacodynamic (PK-PD) models for anticancer drugs.* This theme has continued to be developed with new publications for the drugs irinotecan [5], 5-fluorouracil and oxaliplatin [31], and with a recent mini-review by A. Ballesta and J. Clairambault on mathematical models of treatment of metastatic colorectal cancer [4].

6.1.3. Optimisation of cancer chemotherapy and cancer radiotherapy

Participants: Juan Carlos Alfonso [University Complutense, Madrid, Spain], Annabelle Ballesta, Frédérique Billy, Frédéric Bonnans [Commands project-team], Rebecca Chisholm, Jean Clairambault, Sandrine Dulong [INSERM Villejuif (U 776)], Xavier Dupuis [Commands project-team], Alexandre Escargueil [INSERM and UPMC, St Antoine Hospital], Olivier Fercoq [MaxPlus project-team], Stéphane Gaubert [MaxPlus project-team], Miguel Angel Herrero [University Complutense, Madrid, Spain], Michael Hochberg [ISEM, CNRS, Montpellier], Dirk Drasdo, Nick Jagiella, Francis Lévi [INSERM U 776, Villejuif], Thomas Lepoutre [Dracula project-team], Tommaso Lorenzi, Alexander Lorz, Luis Núñez [University Complutense, Madrid, Spain], Benoît Perthame, Emmanuel Trélat [LJLL, UPMC].

1. **Limiting unwanted toxic side effects: age-structured models of the cell cycle.** Optimising cancer chemotherapy, in particular chronotherapy, is the final aim of these activities. A classical numerical method of optimization under the constraint of limiting toxicity to healthy tissues has been applied to the McKendrick model of the cell cycle divided in phases, endowed with physiologically based targets for both internal (circadian) and external (pharmacological) control. This model has been partly biologically identified on continuous FUCCI recordings of proliferating NIH3T3 cells in culture media; these data were made available to us within the C5Sys consortium, an ERASYSBIO+ European project. Then additional theoretical characteristics establishing hypothetical differences between healthy and cancer cell populations, relying on different responses to physiological circadian clock influences on gating by Cyclin-Cdk complexes between cell cycle phases, have been used to solve the optimization problem, proposing an optimal drug infusion regimen [7], [8], [27], [9]. Using an even more complex McKendrick-like model of the cell cycle, a connection with previously established PK-PD ODE models of the anticancer drugs 5-Fluorouracil and Oxaliplatin has

been established, proposing optimized combined drug delivery flows to solve the same optimization problem [31].

2. **Limiting drug resistance in cancer cell populations: cell Darwinism.** This theoretical activity has been continued 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 [22]. 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 [7], [18], [35]. The time scales under investigation, much shorter than in ecology, are however much longer than in microbiology, and are those of clinical treatments. Theoretical optimization of external controls representing combined cytotoxic and cytostatic treatments on these models with the aim to limit the emergence of drug resistance are presently under assessment, in collaboration with Emmanuel Trélat (LJLL, UPMC), paper in preparation.
3. **Molecular aspects: ABC transporters.** 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 [4], [5].
4. **Optimisation of cell kill in AML.** Underway is also the use of methods of optimal control methods developed by the Commands project-team (Frédéric Bonnans, Xavier Dupuis) to optimise therapies in the treatment of Acute Myeloblastic Leukaemia (AML). X. Dupuis has lately produced a paper [40], accepted for publication in Math. Mod. Phys. Phenom, on optimisation of a combined treatment using a cytotoxic drug (representing Aracytin) and a cytostatic drug (representing AC220, an antagonist of Flt-3 receptors). This work is led in conjunction with the DISCO team, cf. supra "Modelling haematopoiesis with applications to AML").
5. **Estimating dose painting effects in radiotherapy: a mathematical model.** Tumor heterogeneity is widely considered to be a determinant factor in tumor progression and in particular in its recurrence after therapy. Unfortunately, current medical techniques are unable to deduce clinically relevant information about tumor heterogeneity by means of non-invasive methods. As a consequence, when radiotherapy is used as a treatment of choice, radiation dosimetries are prescribed under the assumption that the malignancy targeted is of a homogeneous nature. In this work we discuss the possible effects of different radiation dose distributions on heterogeneous tumors by means of an individual cell-based model. To that end, a case is considered where two tumor cell phenotypes are present, which strongly differ in their respective cell cycle duration and radiosensitivity properties. We show herein that, as a consequence of such differences, the spatial distribution of such phenotypes, as the resulting tumor heterogeneity, can be predicted as growth proceeds. As a consequence, heterogeneous dosimetries can be selected to enhance tumor control by boosting radiation in the region occupied by the more radioresistant tumor cell phenotype. It is also shown that, when compared with homogeneous dose distributions as those being currently delivered in clinical practice, such heterogeneous radiation dosimetries fare always better than their homogeneous counterparts (Alfonso et. al., PLoS One accepted [3]).

6.1.4. Protein polymerisation and application to amyloid diseases

Participants: Annabelle Ballesta, Vincent Calvez [ENS Lyon], Marie Doumic, 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 team of biologists led by H. Rezaei [44], 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 [36] investigated the dependency of the principal eigenvector and eigenvalue upon its parameters. We exhibited possible nonmonotonic dependency on the parameters, opposite 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, 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 [38] 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 [12], in collaboration with N. Krell, M. Hoffmann and L. Robert. Such methods are indeed successfully applied to investigate bacterial growth, in collaboration with L. Robert (INRA and UPMC), see Figure 1 .

In [13], we generalised the inverse techniques proposed previously in [39], [43], 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, using refined estimates the Mellin transform of the equation, has just been accepted for publication in *Inverse Problems* [10].

6.2. Tissue growth, regeneration and cell movements

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

Participants: Luís Lopes Neves de Almeida, Nikolaos Bournaveas [Univ. Edinburgh], Axel Buguin [UPMC, Institut Curie], Vincent Calvez [ENS Lyon], Casimir Emako-Kazianou, 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 for which numerical simulations have been performed. Behaviour of solutions can be understood by performing a hydrodynamical limit of the kinetic equation. It leads to aggregation type equations for which finite time blow up is observed [42]. Then measure solutions for this system should be considered. A theoretical framework for the existence of weak solutions has then been developed [17], [34] where duality solutions for such system has been investigated which are equivalent to gradient flow solutions [33].

../../../../projets/bang/IMG/FigureLydia.jpg

Figure 1. Age and Size Distribution of a bacterial culture (E. coli): comparison between the experimental distribution (A) and the best-fit simulation (B). The methods developed in [38] and [12] allowed us to discriminate between a size-dependent and an age-dependent division rate.

Our understanding of traveling waves has progressed considerably in three directions: fitting continuous models and IBMs [21], fitting precisely models with experiments based on known biological values of parameters, and opening new paradigms: traveling waves can connect a dynamically unstable state to a Turing unstable state, certainly the stable wave connects the unstable state to a pulsating state.

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 [21]. Besides this work, a direct comparison with agent-based and continuum models has been performed, showing very good agreement over a large parameter range.

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

Participants: Gregory Batt [CONTRAINTES project-team], François Bertaux, Noémie Boissier, Kai Breuhahn [German Cancer Centre, Heidelberg], Petru Bucur [Hopital Paul Brousse, Paris], 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 [Research Associate, University of Leipzig], Elmar Heinzle [University of Saarbrücken and NOTOX consortium], Nick Jagiella, Ursula Klingmüller [German Cancer Centre, Heidelberg and LungSys Consortium], Pierre Nassoy [Institut Curie, Paris and Univ. of Bordeaux], Johannes Neitsch, Benoît Perthame, Jens Timmer [University of Leipzig, Germany], Irène Vignon-Clémentel [REO project-team], Paul Van Liedekerke, Eric Vibert [Hôpital Paul Brousse, Villejuif], Ron Weiss [MIT, USA].

1. **Ammonia metabolism in healthy and damaged liver** The model on ammonia detoxification in liver, integrating a compartment model for the glutamine synthetase-active peri-central and the glutamine-inactive peri-portal liver lobule compartment (see Bang report 2012) with the spatial - temporal model of liver regeneration after drug-induced peri-central damage [41] has been extended to include the mass balance of other body compartments. The analysis shows that some body compartments that in the healthy liver produce ammonia, in the damaged liver detoxify blood from ammonia. The detoxification model of liver in combination with the body ammonia balance can be found in ref. (Schliess et. al., Hepatology, accepted [20]).
2. **Drug metabolism in hepatocytes** Since the begining of 2013 animal experiments for testing of cosmetics are forbidden within the EU. This has triggered initiatives towards how modeling may help to investigate drug toxicity, circumventing animal testing. The basic conceptual idea is to test drugs (cosmetics, perspectively also other drugs) in in-vitro systems such as monolayers, sandwich cultures, or multi-cellular spheroids, and use the emerging data to infer the expected toxicity in-vivo using novel experimental and computational approaches [16]. We have integrated an intracellular mathematical model of paracetamol drug metabolism in a mathematical agent-based cell model for monolayer and multi-cellular spheroids and compared simulation results with experimental findings in the same systems. We find that cell-to-cell variabilty can largely explain the experimentally observed cell population survival fractions. The mathematical model is now refined based on measurements of intermediate drug metabolites.

3. **Cell mechanics and its impact on cell proliferation** A novel numerical methodology has been developed to simulate the mechanics of cells and tissues using a continuum approach. Analogously to the Center Based Models, particles are used to represent (parts of) the cells but rather than discrete interactions they represent a continuum. This approach can be used for tissue mechanics simulations in where the individual cell-cell interactions are discarded but instead a constitutive law is proffered [23].
Moreover, a new model in where cell adhesion dynamics is addressed. The cell model is constructed by a triangulated surface and a coarse-grained internal scaffolding structure. A model cell can adapt to realistic cell shapes, and is able to interact with a substrate or other cells. The parameters in this model can be determined by canonical experiments performed on cells informing about cell deformation, compression and cell-cell adhesion [19].
A computational model for the confined growth of cells in a capsule has been developed. This model represents a realistic simulation tool for a novel experimental system (Institut Curie, Prof P. Nassoy) in where cells are grown in an elastic environment to mimic the effects of mechanical stress on cells and while monitoring their fate. Model parameter calibration is now ongoing to reproduce the correct quantitative behavior of the cells in order to unravel the relationship between cell mechanical stress and cell behavior.
4. **Playing the game of life with yeast cells** Within a collaboration with a synthetic biology lab at MIT, multicellular modelling of engineered yeast cell populations is performed. Those cells secrete a messenger molecule (IP) which diffuse in the medium, bind to other cells, and trigger a signalling cascade, which finally induces expression of lethal genes. A model has been established based on our single-cell-based model framework associated with PDE simulations, and it is currently used to explain and guide experiments conducted at the MIT. In 2013, the project has achieved significant progress on several aspects. First, we were able to quantitatively reproduce newly produced, rich data on the signaling cascade behavior with a kinetic model describing signaling reactions. Second, comparison between simulations and data allowed to identify key characteristics of the death module, which is positioned downstream of the signaling cascade: there is a rapid and stochastic commitment to death, followed by a deterministic and long delay (2-4 cell generations) needed before cells actually die. Finally, data production and analysis iterations with our collaborators alloweded to optimize the procedures for experimental measurements and the quantitative analysis of data in a synergistic manner.
5. **Other projects in short** Further progress have been achieved on the reconstruction of lung cancer micro-architecture from bright field micrographs. In partial hepatectomy (PHx), pig data on the changes of microarchitecture during regeneration after PHx have been generated and stained now being processed. The image processing chain for liver architecture reconstruction has been refined and extensive analysis has been performed on the architecture of the bile canaliculi network in healthy liver and in disease states of liver. Moreover, non-small-cell lung cancer cell invasion pattern have been analyzed leading to interesting observations now being studied by modelling.
For multi-scale modeling of liver regeneration after drug-induced pericentral damage, integration of a molecular model of hepatocyte growth factor signalling with an agent-based model of liver regeneration has been extended to include blood flow in the lobule, as well as the contributions of the body compartment to the degradation and production of hepatocyte growth factor (HGF).

6.2.4. *Modelling flows in tissues*

Participants: Noémie Boissier, 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].

1. **Flow and perfusion scenarios in cancer.** We started reconstruction of the blood vessel system of lung cancers removed by surgery. For this purpose, patients underwent DCE-MRI prior to surgery. Part of the tumors after surgery was sliced and stained for nuclei, proliferation and endothelial cells. The slice data were recorded (Mevis, Luebeck) to allow identification of the position of the indiviudal structures in 3D space. The structures were then segmented. The work turned out to be particularly challenging because of staining artifacts for which image algorithms had to correct for. Nevertheless, last results look promising so that at least the network formed by larger vessels can be segmented and reconstructed in 3D. The so emerging data will be used for modeling of blood flow using the models developed in 2012.
2. **Flow in liver lobules.** We integrated blood flow in the new software CellSys (see above under software) and refined the algorithms. Moreover, we increased the resolution of the capillaries by triangulating them from high resolution confocal scanning micrographs.

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

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

In 2013 we continued to investigate the dependence of physical and biological mechanisms of actomyosin cable formation and wound closure depending on the geometry of the wound, with particular emphasis on the effect of the wound edge curvature.

When the actomyosin cable starts to contract and the wound starts to close we have noticed that the behavior of the cable is related with the local curvature of the wound edge. This led us to study the curves evolving by positive part of their curvature in a Euclidean framework. A model where we consider viscous behavior and friction in the tissue plus boundary terms associated to cable and lamellipodial forces is under development. The numerical simulations obtained using this model are in good agreement with the previous experimental results and we are pursuing the model development by challenging it with new experiments.

6.3. Neurosciences

Participants: Jonathan Touboul, Gilles Wainrib, Tanguy Cabana, Mathieu Galtier, Luis Garcia Del Molino, Khashayar Pakdaman.

We pursued our studies of disordered networks of the brain and collective phenomena in neuroscience. We have been more interested this year in the role of disorder in the spontaneous emergence of synchronized activity. In order to study these phenomena, we have been establishing limit equations for randomly coupled networks [11], and the analysis of this equation reveal a number of transitions due to the level of disorder in the connectivity. A universal transition observed in such randomly coupled networks is a transition to chaotic activity for large levels of noise. These transitions were investigated [24] and were shown to be related to an explosion of complexity at the edge of chaos, i.e. the number of equilibria is exponentially large with the network size at the phase transition, and the exponential factor was related to the Lyapunov exponent. These large-scale limits give rise to nonlinear reduced equations that we have been introducing in [15]. Eventually, when considering that the network is structured into different populations and that the connectivity weights satisfy a balance condition, which is postulated as a natural scaling of the synaptic input, we have shown that the network shows random transitions to periodic activity depending on the spectrum of the random connectivity matrix [25], yielding up and down states or synchronized oscillations depending on the eigenvalue of larger real part of the connectivity matrix.

BEAGLE Project-Team

6. New Results

6.1. Stochastic dynamics of gene expression

A number of studies have established that stochasticity in gene expression may play an important role in many biological phenomena but the molecular mechanisms at stake are still poorly understood. By joint experimental and computational approaches, we explored the role played by chromatin dynamics in the regulation of stochastic gene expression in higher eukaryotic cells [31]. For this purpose, our biological partner generated isogenic chicken-cell populations expressing a fluorescent reporter integrated in one copy per clone. Although the clones differed only in the genetic locus at which the reporter was inserted, they showed markedly different fluorescence distributions, revealing different levels of stochastic gene expression. Use of chromatin-modifying agents then showed that direct manipulation of chromatin dynamics had a marked effect on the extent of stochastic gene expression. We then fitted the experimental data to a two-state model describing the opening/closing process of the chromatin. The model showed that the differences between clones seemed to be due mainly to the duration of the closed state, and that the agents we used mainly seem to act on the opening probability. These results highlight the importance of chromatin dynamics in stochastic gene expression. They shed a new light on the mechanisms of gene expression in higher eukaryotic cells, and argues in favor of relatively slow dynamics with long (hours to days) periods of quiet state.

This work was part of Gaël Kaneko's PhD and results from our long-lasting collaboration with Olivier Grandrillon and his BM2A team in the CGphyMC (Centre de Génétique et de Physiologie Moléculaire et Cellulaire, Lyon).

6.2. The impact of anomalous diffusion on cell signaling

This year, we published two papers describing the impact of diffusion and clustering in membrane domains for ubiquitous biological pathways. In the first paper [18], we showed that clustering of receptors in membrane domains affect severely the activation of 'hit and run' type of pathways (eg. IRS1, G proteins). This is a pure diffusion result and it is obtained without modifying molecular affinity. This impairment is dramatically important when receptors are highly clustered such are the cases for insulin and adrenergic receptors. In the same direction, we studied the impact of modified diffusion on the other ubiquitous pathway: enzyme/substrate equilibrium [10]. In that case diffusion was modified either using subdiffusion - obstacles and Continuous Time Random Walk - or using space-based inhomogeneous diffusion. We showed that while impairing diffusion all three mechanisms behave differently in the stationary regime. Therefore it it not possible to assume simple space-dependent diffusion for subdiffusion at the equilibrium limit (as it is always assumed). Furthermore, we showed that in the case of space-dependent diffusion - the shape of the diffusion profile can drastically affect the equilibrium and modify the pathway. Since these three phenomenons are thought to occur either in the membrane or the cytosplasm, our results show they can have non trivial effect of all chemical reaction occurring within these medium.

This research will be carried on in the group in 2014 and expanded by the more mathematical approaches initiated in 2013 by collaborations with V Calvez (Numed Inria Lyon), T Lepoutre (Dracula, INIA Lyon) and S. Fedotov (Univ Manchester, UK).

6.3. Localization of protein aggregates in E. coli

Aggregates of misfolded proteins are a hallmark of many age-related diseases. Recently, they have been linked to aging of *Escherichia coli* (*E. coli*) where protein aggregates accumulate at the old pole region of the aging bacterium. Because of the potential of *E. coli* as a model organism, elucidating aging and protein aggregation in this bacterium may pave the way to significant advances in our global understanding of aging. A first obstacle along this path is to decipher the mechanisms by which protein aggregates are targeted to specific intercellular locations. Here, using an integrated approach based on individual-based modeling, time-lapse fluorescence microscopy and automated image analysis, we show that the movement of aging-related protein aggregates in *E. coli* is purely diffusive (Brownian). Using single-particle tracking of protein aggregates in live *E. coli* cells, we estimated the average size and diffusion constant of the aggregates. Our results provide evidence that the aggregates passively diffuse within the cell, with diffusion constants that depend on their size in agreement with the Stokes-Einstein law. However, the aggregate displacements along the cell long axis are confined to a region that roughly corresponds to the nucleoid-free space in the cell pole, thus confirming the importance of increased macromolecular crowding in the nucleoids. We thus used 3D individual-based modeling to show that these three ingredients (diffusion, aggregation and diffusion hindrance in the nucleoids) are sufficient and necessary to reproduce the available experimental data on aggregate localization in the cells. Taken together, our results strongly support the hypothesis that the localization of aging-related protein aggregates in the poles of *E. coli* results from the coupling of passive diffusion-aggregation with spatially non-homogeneous macromolecular crowding. They further support the importance of “soft” intracellular structuring (based on macromolecular crowding) in diffusion-based protein localization in *E. coli*.

This work is a collaboration with the microbiology group led by A. Lindner (INSERM U1001, Cochin Med School, Paris). It has been published in [3] as part of A.S. Coquel’s PhD (defended Nov 2012, co-supervision H. Berry-A. Lindner).

6.4. The molecular signaling basis of neuronal plasticity

Many of the cell-level properties of the neurons vary as a function of the signals from other neurons or past activity. These modifications are often maintained in the long term, giving rise to cell memory. We have developed models of how the implicated signaling networks self-organize to support a memory and how this leads to cell-level responses such as changes of the firing threshold [24] or the spike-timing dependence [34]. The latter, for instance, corresponds to the observation that the probability of transfer of an electrical signal (spike) between two connected neurons (the synaptic weight) adapts depending on the timing between previous consecutive presynaptic and postsynaptic spikes. Combining a model of the implicated signaling networks with experimental measurements, we have uncovered the molecular mechanisms supporting this memory.

This work is developed in collaboration with both with applied mathematicians (B. Cessac, Inria Neuromath-comp, Sophia-Antipolis) and experimental neurobiologists (L. Venance, Collège de France, Paris).

6.5. A model for adipocyte size based on size-dependent lipid fluxes

We proposed in a paper published this year [28] a novel model that explains some of the peculiarities in the fat tissue storage cells. Indeed, adipocytes, as they are called, come in various size – with up to one order of magnitude in amplitude – but do not possess any characteristic size. The cellularity, the cell size distribution, is bimodal. We showed that a simple model of size-dependent lipid fluxes (using data from Carmen Lab) can explain this bimodality and allow us to retrieve any target cell distribution. Our result also provides an elegant and testable hypothesis for the triggering of adipocytes proliferation. The amount of unstored free fatty acid is actually a marker that the population has reached its maximal volume. This amount could serve as an index to start the proliferation.

This was a joint work with experimentalists from the CARMEN Institute (INSERM UMR1060, Lyon), namely C. Soulage and A. Géloën, and was part of H. Julienne master’s thesis.

6.6. Evolution of antibiotic resistance

The emergence of antibiotic resistant bacteria is a major threat to public health and there is a constant need for education to limit dangerous practices. Here, we propose to use alife software to develop training media for the public and the physicians. On the basis of the Aevol model we have been developing for more than six years, we built a game in which players fight bacterial infections using antibiotics. In this game the bacteria can evolve resistance traits, making the infection more and more difficult to cure. The game has been tested with automatic treatment procedures, showing that it behaves correctly. It was demonstrated during the French "Nuit des Chercheurs" in October 2012 and was published in 2013 in the ECAL conference [2].

This is a joint work with Dominique Schneider from the Laboratoire Adaptation et Pathogénie des Microorganismes (LAPM, UMR CNRS 5163, Grenoble).

6.7. Spontaneous dynamics of genome size

Even though numerous genome sequences are now available, evolutionary mechanisms that determine genome size, notably their fraction of non-coding DNA, are still debated. In particular, although several mechanisms responsible for genome growth (proliferation of transposable elements, gene duplication and divergence, etc.) were clearly identified, mechanisms limiting the overall genome size remain unclear. By using a matrix population model, we showed that genome size can be simply limited by the spontaneous dynamics of duplications and large deletions, which tends to make genomes shrink even if the two types of rearrangements occur at the same rate. In the absence of Darwinian selection, we proved the existence of a stationary distribution of genome size even if duplications are twice as frequent as large deletions. To test whether selection can overcome this spontaneous dynamics, we also simulated our model numerically and chose a fitness function that directly favors genomes containing more genes, while keeping duplications twice as frequent as large deletions. In this scenario where, at first sight, everything seems to favor infinite genome growth, we showed that genome size remains nonetheless bounded. As a result, our study reveals a new pressure that could help limiting genome growth.

This work was part of Stephan Fischer's PhD thesis, which was defended in December 2013. A manuscript is currently under review. Stephan's PhD was co-supervised by Samuel Bernard (Inria Dracula team and Institut Camille Jordan, UMR CNRS 5208, Lyon).

6.8. Inference of evolutionary molecular events at different scales

We have progressed in the integration of several evolutionary events at different scales of genomes in a single model used for inference of ancient events from the observation of extant genomes. We handle nucleotide substitutions, gene duplications, losses, lateral transfers and rearrangements. We have tested the framework on 36 cyanobacteria species, reconstructing up to 80% of ancestral chromosomes in some clades [8]. The inference algorithm is still mainly sequencial, in the sense that it first accounts for nucleotide substitutions, gene duplications, losses, lateral transfers [29], and then for rearrangements. But we also developed a way to provide a feedback of the result on rearrangements to the inference of substitutions by correcting gene trees [22], [38]. We have used these methods to reconstruct a nucleotide-scale sequence of the genome of the medieval black death agent [9],[44]. It includes a chromosome and three plasmids, and is different in structure from any extant strain. We follow the first ancient bacterial genome sequencing in 2011 and complete and order the genome with computational predictions. We then dispose of a complete view of the molecular evolution in the *Yersinia pestis* clade.

This work was part of Murray Patterson's post-doctoral fellowship. It also involved collaborations with L. Gueguen and V. Daubin from the Laboratoire de Biométrie et Biologie Evolutive (UMR CNRS 5558, Lyon), with Nadia El-Mabrouk from the Département d'Informatique et de Recherche Opérationnelle in Montréal (Canada), with Cédric Chauve from the Department of Mathematics of Simon Fraser University (Burnaby, Canada), and with G. Szollosi from the Biophysics Research Group in Budapest (Hungary).

BIGS Project-Team

6. New Results

6.1. Modern methods of data analysis

Participants: R. Bar, S. Ferrigno, B. Lalloué, J-M. Monnez, A. Muller-Gueudin, S. Tindel

6.1.1. Help to medical decision and telemedecine in the monitoring of heart failure

We describe here a project started in 2013, for which we expect some concrete output in 2014. This project fits in the general framework of telemedecine and more precisely in the monitoring of heart failure patients. From measurements performed automatically and daily on a patient at home through a new process under development at the Pluri-Thematic Clinical Investigation Center of the University Hospital of Nancy, the aim is to propose therapeutic adjustments to improve the prognosis of patient in order to increase his chances of survival or to avoid his rehospitalization.

The patient's condition and its evolution are determined by the initial values of his biological or clinical parameters as well as those collected throughout his follow-up. The treatments are intended to stabilize or change the values of parameters in order to avoid the occurrence of adverse events, in particular the death of the patient. This is why the first part of the study will consist in building survival scores or rehospitalization scores according to the values of biological or clinical parameters.

In a second part, we will seek to build models of the evolution of the values of biological or clinical parameters depending on treatments (average or cumulative drug doses, drug combinations) and patients' characteristics. This will allow to predict the potential effect of an adjustment proposal or modification of treatment and then predict a new survival score to conclude the relevance or not of the proposed medication.The physician will have this help to confirm or change his decision which belongs finally to him.

We will use to carry out this study a wide range of classic and recent methods of data analysis, in particular discriminant analysis, without a priori: several methods will be used, compared and selected according to their performance in the treated applications.

6.1.2. Online factorial data analysis methods

Nowadays data analysts are often faced with the problem of dealing with a rapid and infinite flow of data. Examples include web, telecommunications, process control or financial data. We made first the assumption that the data are generated at random according to a stationary distribution, but in many cases this assumption does not hold true. We developed in [13] the online adaptation of principal component analysis and other dimension reduction statistical algorithms by using stochastic approximation. An R package was developed by Romain Bar.

6.1.3. Data analysis techniques and Bayesian models applied to the context of social inequalities and environmental exposures

The aim of [10] is to improve the knowledge about and apply data mining techniques and some Bayesian model in the field of social and environmental health inequalities. The health event considered is infant mortality. We try to explain its risk with socio-economic data retrieved from the national census and environmental exposures such as air pollution, noise, proximity to traffic, green spaces and industries. The data mining part details the development of a procedure of creation of multi-dimensional socio-economic indices and of an R package that implements it.

6.1.4. A simultaneous stepwise variable selection and clustering algorithm to discriminate a class variable with numerous levels

In supervised learning the number of levels of a categorical variable to explain can be high. When some of its levels are of low frequency, clustering them in order to reduce the number of classes can be useful to perform relevant discriminant analyses. On the other hand selecting relevant predictors is a crucial step to build robust and efficient classification rules, especially when too many variables are available in regard to the overall sample size. We are currently carrying out an extension of an algorithm we had devised to solve both these problems using an alternate minimization of Wilks' Lambda. We show through simulations the interest of adding Akaike Information Criterion as another optimality criterion. We also moved forward to stepwise selection and applied this new version of our algorithm to real allergology datasets.

6.1.5. Local polynomial regression. Application to the estimation of the fetal growth.

This topic is an ongoing collaboration with M. Maumy-Bertrand, for which we expect a publication in 2014. We have established exact rate of strong uniform consistency for the local linear estimator of the conditional distribution function. We want to extend our results to obtain exact rates of strong uniform consistency for the local linear estimator of other conditional quantities: the conditional mean $\mathbb{E}(Y|X)$, and the conditional quantiles $q_\alpha(x) = \inf\{y : F(y|x) \geq \alpha\}$, for $\alpha \in (0, 1)$.

Another crucial problem with the non parametric regression methods is the choice of the bandwidth parameter h. It is common in practice to choose $h > 0$ so to minimize asymptotically the mean square error (MSE) or the mean integrated square error (MISE). This minimization leads to an optimal choice of h of the form $h_n = C(X_1, ..., X_n)n^{-1/5}$, where n is the sample size, and $X_1, ..., X_n$ are the n independent copies of the random variable X. This bandwidth is called a *data-driven bandwidth* to enhance its dependence to the data. Our current project in this direction consists in establishing the consistency of the local linear estimator when the bandwidth h is allowed to range in a small interval which may decrease in length with the sample size. Such a result would be immediately applicable to prove uniform consistency of the local linear estimator when the bandwidth is a data-driven bandwidth $h_n = C(X_1, ..., X_n)n^{-1/5}$.

Turning to applications, note that we have a contact with Professor Bernard Foliguet at the Maternité Régionale de Nancy. We will continue to collaborate with him, to estimate growing curves of the fetal weight, and other fetal quantities thanks to the techniques mentioned above.

6.1.6. Cohort analysis

In an ongoing work with the INSERM team of P. Guéant, we aim at describing the complex interactions between genetic, phenotypic and biologic variables that are available in medical cohorts, in different contexts (cognitive decline; inflammatory intestinal diseases; liver cancer).

A firt step in our analysis, which should be finished in 2014, consists in giving an overview of the existing methods given in the literature, for the analysis of qualitative and quantitative data. Indeed, we have to describe links between qualitative and quantitative data:

1. with exploratory methods, or factorial models,
2. with regression models to predict qualitative variable by the use of qualitative or quantitative factors.

In the sequel, we will test non association or independence between the variables. The objective is to develop new methods, adapted to the studied cohorts (matching cases/controls, high number of individuals, high number of explicative variables, missing data problem). The particularity of our work is to combine statistical and symbolical methods.

After having identified and choice the relevant variables, we will have to give a model for classifying the data. The proposed models will allow us to identify subgroups of invidious, with common genetic, biologic and phenotypic characteristics.

6.1.7. Local polynomial estimation and goodness-of-fit tests

We describe here an ongoing work with Marie-José Martinez, assistant professor at the IUT of Grenoble and member of the Inria MISTIS team. A related publication should be finished at the end of 2014. Many clinical trials and other medical studies involve responses that might be considered to have a normal distribution. However, this is not invariably the case and models based on this distribution are often indiscriminately applied to data which might be better handled otherwise. This is especially true for discrete data. An approach which may yields models that are more biologically reasonable in many situations is to use generalized linear models (GLM).

In statistical theory, generalized linear models were formulated by John Nelder and Robert Wedderburn (1972) as a way of unifying various other statistical models including for examples linear regression, logistic regression and poisson regression. Such a technique was developed by McCullagh and Nelder (1989). It is an extension of the linear model, in the sense that its satisfies a relation of the form $Y = g(X) + \epsilon$ where:

- The stochastic component ϵ follows other distributions than the Gaussian.
- The function g can be non linear.

Notice that those models are well-suited to analyze dependences between variables following distributions in the so-called exponential family, like Poisson, Binomial and Gamma distributions. In practice, link functions are chosen such that the inverse link, $\mu = g^{-1}(\eta)$ is easily computed. For instance, for binomial data, logit and probit link functions are commonly used.

Our aim in this project is to use generalized linear models in order to extend our global test of goodness-of-fit to a wide range of models used in biological and medical applications. We wish to use the cumulative conditional distribution $F(y|X = x)$ again, which embodies all the information about the joint behavior of two random variables. The expected outcome is a global goodness of fit test for the relationship between two random variables in the exponential family. The test will compare a nonparametric estimator of the cumulative distribution function with the value of the cumulative distribution function under the null hypothesis.

6.1.8. Model selection for SVM

Support vector machines provide a very powerful method of data classification, for which model selection is one of the key issues. For a support vector machine, model selection consists in selecting the kernel function, the values of its parameters, and the amount of regularization. To set the value of the regularization parameter, one can minimize an appropriate objective function over the regularization path. A priori, this requires the availability of two elements: the objective function and an algorithm computing the regularization path at a reduced cost. The literature provides us with several upper bounds and estimates for the leave-one-out cross-validation error of the ℓ_2-SVM. However, no algorithm was available so far for fitting the entire regularization path of this machine. In our contribution [3], we introduce the first algorithm of this kind. It is involved in the specification of new methods to tune the corresponding penalization coefficient, whose objective function is a leave-one-out error bound or estimate. From a computational point of view, these methods appear especially appropriate when the Gram matrix is of low rank. A comparative study involving state-of-the-art alternatives provides us with an empirical confirmation of this advantage.

6.2. Estimation for complex and biological systems

Participants: T. Bastogne, C. Lacaux, S. Mézières, S. Tindel, P. Vallois

6.2.1. Tumor growth modeling

This project is an extension of our article [15], which will be written in 2014. A cancer tumor can be represented for simplicity as an aggregate of cancer cells, each cell behaving according to the same discrete model and independently of the others. Therefore to measure its size evolution, it seems natural to use tools coming from dynamics of population, for instance the logistic model. This deterministic framework is well-known but the stochastic one is worthy of interest. We are currently working on a model in which we suppose that the size V_t at time t of the tumor is a diffusion process of the type :

$$\begin{cases} dV_t = r\,V_t\,\left(1 - \frac{V_t}{\kappa}\right) - c\,V_t + \beta\,V_t\,dB_t \\ V_0 = v > 0 \end{cases} \tag{1}$$

where $(B_t)_{t\geq 0}$ is a standard brownian motion starting from zero.Then (i) We define a family of time continuous Markov chains which models the evolution of the rate of malignant cells and approximate (under some conditions) the diffusion process (V_t). (ii) We study in depth the solution to equation (1). This diffusion process lives between two frontiers : 0 and κ. In this stochastic setting, the role of κ is not so clear and we contribute to understand it. We describe the asymptotic behavior of the diffusion according to the values of the parameters. The tools we resort to are boundary classification criteria and Laplace transform of the hitting time to biological worthwhile level. We believe we are able in particular to express the mean of the hitting time.

The next step in this project can be summarized as follows: at this point in our investigations on tumor growth modelization, we have identified a pertinent and consistent model. Nevertheless our study remains theoretical. A statistical estimation of the parameters $r, \kappa, c\beta$ is thus in order. This would permit to apply our model to real data. A further objective could be to consider a more complex form for the logistic term, see e.g. Schurtz (2007).

6.2.2. Local score associated with long biological sequences

Statistical properties of the distribution of the local score is largely used by molecular biologists to extract important features in biological sequences and in particular to determine the most significant one among a collection of biological sequences. The probabilistic model which is commonly used is the following. Associated with a sequence $(\epsilon_i)_{i\geq 1}$ of independent, centered and reduced random variables, consider $S_n = \epsilon_1 + \cdots + \epsilon_n$ and

$$\underline{S}_n = \min_{0\leq i\leq n} S_i, \quad U_n = S_n - \underline{S}_n = S_n - \min_{i\leq n} S_i \quad n \geq 0.$$

In biological sequence analysis, (ϵ_i) can for example correspond to the physical or chemical properties of the i-th amino acid or nucleotid of the sequence ; it can also reflect the similarity between components of two sequences. The local score $\overline{U}_n$ is the supremum of (U_n) up to time n. Molecular biologists are interested by this supremum as it highlights the best part of the studied sequence, the eventual segment of DNA transmitted by a common ancestor for sequence comparison or the best hydrophobic segment of a protein that would thus naturally move in a transmembrane place. It is clear that the trajectory of (U_n) can be decomposed in a succession of 0 and excursions above 0. These excursions have an important biological interpretation and in particular the highest one corresponds to the best segment due to the physico chemical property or similarity scores that have been chosen. Note that the local score $\overline{U}_n$ can be viewed as the maximum of the heights of all the excursions up to time n. In the article [22], we are interested in complete excursions up to a fixed time. This leads us to introduce the maximum U_n^* of the heights of all the excursions up to time n. The second variable which will play an important role is θ_n^* the time necessary to reach its maximal height U_n^*.
We believe that the knowledge of the joint distribution of the pair (U_n^*, θ_n^*) would permit to get more efficient statistical tests than the ones only based on the local score. This point should be developed in a forthcoming paper.

However, it seems difficult to determine explicitly the law of $(U_n^*,\ \theta_n^*)$ for a fixed n. This difficulty can be overcome considering biological sequences which have a large number of bases and approximating the initial random walk (S_n) by a Brownian motion (B_t) started at 0. Using the functional theorem of convergence of Donsker, the process (U_k) can be compared to

$$\widehat{U}(t) := B(t) - \inf_{0\leq s\leq t} B(s), \quad t \geq 0. \tag{2}$$

This leads us to consider:

1. the local score $\overline{U}(t)$ which is the maximum of the heights of all the excursions of $U(s)$ up to time t,
2. the maximum $U^*(t)$ of the heights of all the complete excursions up to time t,
3. the time $\theta^*(t)$ taken by $U(s)$ starting from the beginning of the largest excursion to hit the maximal level $U^*(t)$.

The approximation of (U_n) by $(\widehat{U}_t)$ permits to prove that the asymptotic distribution of $\left(\frac{U_n^*}{\sqrt{n}}, \frac{\theta_n^*}{n}\right)$ as $n \to \infty$ is the one of $(U^*(1), \theta^*(1))$. Consequently, our initial problem in the discrete setting reduces to determine the joint law of $(U^*(t), \theta^*(t))$, where $t > 0$ is given. We determine in [22] the distribution and the density functions of $(U^*(t), \theta^*(t))$.

6.2.3. Bacteriophage therapies

In the last years Bacteriophage therapies are attracting the attention of several scientific studies. They can be a new and powerful tool to treat bacterial infections or to prevent them applying the treatment to animals such as poultry or swine. Very roughly speaking, they consist in inoculating a (benign) virus in order to kill the bacteria known to be responsible of a certain disease. This kind of treatment is known since the beginning of the 20th century, but has been in disuse in the Western world, erased by antibiotic therapies. However, a small activity in this domain has survived in the USSR, and it is now re-emerging (at least at an experimental level). Among the reasons of this re-emersion we can find the progressive slowdown in antibiotic efficiency (antibiotic resistance). Reported recent experiments include animal diseases like hemorrhagic septicemia in cattle or atrophic rhinitis in swine, and a need for suitable mathematical models is now expressed by the community.

At a mathematical level, whenever the mobility of the different biological actors is high enough, bacteriophage systems can be modeled by a kind of predator-prey equation. Namely, set S_t (resp. Q_t) for the bacteria (resp. bacteriophages) concentration at time t. Then a model for the evolution of the couple (S, Q) is as follows:

$$\begin{cases} \mathrm{d}S_t &= [\alpha - k\,Q_t]\,S_t \mathrm{d}t + \varepsilon\, S_t\, \mathrm{d}W_t^1 \\ \mathrm{d}Q_t &= \left[d - mQ_t - k\,Q_t\,S_t + k\,b\,e^{-\mu\zeta}\,Q_{t-\zeta}\,S_{t-\zeta}\right] \mathrm{d}t + \varepsilon\,Q_t\,\mathrm{d}W_t^2, \end{cases} \tag{3}$$

where α is the reproducing rate of the bacteria and k is the adsorption rate. In equation (3), d also stands for the quantity of bacteriophages inoculated per unit of time, m is their death rate, we denote by b the number of bacteriophages which is released after replication within the bacteria cell, ζ is the delay necessary to the reproduction of bacteriophages (called latency time) and the coefficient $e^{-\mu\zeta}$ represents an attenuation in the release of bacteriophages (given by the expected number of bacteria cell's deaths during the latency time, where μ is the bacteria's death rate). A given initial condition (S_0, Q_0) is also specified, and the term $\varepsilon\,\mathrm{d}W_t$ takes into account a small external noise standing for both uncertainties on the measures and the experiment conditions. One should be aware of the fact that the latency time ζ (which can be seen as the reproduction time of the bacteriophages within the bacteria) cannot be neglected, and is generally of the same order (about 20mn) as the experiment length (about 60mn).

With this model in hand, our main results in this direction (see [1]) have been the following:

- Quantification of the exponential convergence to a bacteria-free equilibrium of equation (3) when d is large enough.
- Use of the previous result plus concentration inequalities in order to study the convergence of the noisy system to equilibrium in a reasonable time range.
- Simulation of the stochastic processes at stake in order to observe the convergence to equilibrium.

6.2.4. Light transport in tissues with probabilistic methods

Photodynamic therapy (PDT) is a type of phototherapy used for treating several diseases such as acne, bacterial infection, viruses and some cancers. The aim of this treatment is to kill pathological cells with a photosensitive drug that is absorbed by the target cells and that is then activated by light. For appropriate wavelength and power, the light beam makes the photosensitizer produce singlet oxygen at high doses and induces the apoptosis and necrosis of the malignant cells. Our project focuses on an innovative application: the interstitial PDT for the treatment of high-grade brain tumors. This strategy requires the installation of optical fibers to deliver the light directly into the tumor tissue to be treated, while nanoparticles are used to carry the photosensitizer into the cancer cells. In order to optimize the intra-cerebral position of our optical fiber, two fundamental questions have to be answered:

1. What is the optimal shape and position of the light source in order to optimize the damage on malignant cells?
2. Is there a way to identify the physical parameters of the tissue which drive the light propagation?

Notice that we are obviously not the first ones to address these issues, and there is nowadays a consensus in favor of the algorithm proposed by L. Wang and S. L. Jacques for the simulation of light transport in biological tissues. However, our starting point is the observation that the usual methods slightly lack of formalism and miss formal representations that answer the questions of identifiability. In [25], in the framework of homogeneous biological tissues, we propose an alternative MC method to Wang's algorithm. Then we also propose a variance reduction method. Interestingly enough, our formulation also allows us to design quite easily a Markov chain Monte Carlo (MCMC) method based on Metropolis-Hastings algorithm and to handle the inverse problem (of crucial importance for practitioners), consisting in estimating the optical coefficients of the tissue according to a series of measurements. We have compared the proposed MC and MCMC method and Wang's algorithm: we see that our MC method is much more consistent. However, MCMC methods induce quick mutations, which paves the way to very promising algorithms in the inhomogenous case. To handle the inverse problem, we derive a probabilistic representation of the variation of the fluence with respect to the absorption and scattering coefficients. This leads us to the implementation of a Levenberg-Marquardt type algorithm that gives an approximate solution to the inverse problem.

6.2.5. System identification of gap junctional intercellular communication channels of two cancer cell lines.

The main challenge addressed in this work [12], [14] was to assess the relevance of a proposed model-based approach to detect differences between gap junctional intercellular communication channels of two cancer cell lines. To that aim, KB and FaDu, two human head and neck carcinoma cell lines, were used. The former expresses connexin proteins (positive line) while the latter does not (negative line). Moreover, each cell line was tested on spheroid (3-D) and monolayer (2-D) slices and *in vitro* assays were repeated six times. Continuous-time system identification algorithms of the Matlab System Identification and CONTSID toolboxes are tested and applied to a set of *in vitro* data. Results firstly show an acceptable fit of the biological responses but they mainly emphasize the possibility to use several model parameters as statistics to discriminate different cancer cell lines. So, this study exemplifies the potential contribution of dynamic system identification methods and tools to the discovery of new diagnostic biomarkers in systems biology.

6.2.6. Photodynamic therapy modeling and control.

We have also carried on the development of methodological and technological innovations for the realtime control of the therapeutic efficiency in photodynamic therapy (Tylcz:2013). One part of the innovation has been protected by a patent submitted in 2012 (No.1261339, INPI) and extended in 2013. A demonstration platform is currently in development.

6.2.7. Bio-inspired system reliability method.

Based on previously developed works (Keinj, 2011, 2012), we have also proposed in [15] an extension of the target theory in biology applied to system reliability. In this development, we consider rough products produced by a factory. Each product coming from the plant has m vital elements and some elements can be damaged. To obtain a perfect product (i.e. all the constitutive m elements are safe) all the damaged elements are repaired and a test phase follows. The result of this two-steps procedure is random. We suppose that the number Z_k of non-damaged elements is a Markov chain valued in the set $\{0, 1, \cdots, m\}$, where k is the number of applied repairing-test phases. We have a qualitative result which says that if the repair phase is efficient then $P(Z_k = m)$ is close to 1. As for production of a large number n of products, the former result allows us to give conditions under which either the n elements or a fraction of these n elements are (is) safe after the application of k previous maintenance phases.

6.2.8. Dynamical Global Sensitivity Analysis as an Early Warning for System's Critical Transition.

In biology, systems with bifurcations may experience abrupt irreversible and often unwanted shifts in their performance, called critical transitions. For many systems like climate, economy, ecosystems it is highly desirable to identify indicators serving as early warnings of such regime shifts. Several statistical measures were recently proposed as early warnings of critical transitions including increased variance, autocorrelation and skewness of experimental or model-generated data. The lack of automatized tool for model-based prediction of critical transitions led to designing DyGloSA, a Matlab toolbox for dynamical global parameter sensitivity analysis (GPSA) of ordinary differential equations models. One part of our research activity in 2013 was focused on the implementation of a global sensitivity analysis method developed in (Dobre, 2011, 2012) into DyGloSA for dynamical global parameter sensitivity analysis (GPSA) of ordinary differential equations models. This work has been carried out in the context of a collaboration with the University of Luxembourg and more precisely the Thomas Sauter's team. We have shown in [2] that tools developed in this toolbox are efficient to analyze several models with bifurcations and predict the time periods when systems can still avoid going to a critical transition by manipulating certain parameter values, which is not detectable with the existing SA techniques.

6.3. Inference for gaussian systems

Participants: C. Lacaux, S. Tindel

6.3.1. Inference for dynamical systems driven by Gaussian noises.

As mentioned at the *Scientific Foundations* Section, the problem of estimating the coefficients of a general differential equation driven by a Gaussian process is still largely unsolved. To be more specific, the most general ($\mathbb{R}$-valued) equation handled up to now as far as parameter estimation is concerned is of the form:

$$X_t^\theta = a + \theta \int_0^t b(X_u)\, du + B_t,$$

where θ is the unknown parameter, b is a smooth enough coefficient and B is a one-dimensional fractional Brownian motion. In contrast with this simple situation, our applications of interest (see the *Application Domains* Section) require the analysis of the following $\mathbb{R}^n$-valued equation:

$$X_t^\theta = a + \int_0^t b(\theta; X_u)\, du + \int_0^t \sigma(\theta; X_u)\, dB_t, \tag{4}$$

where θ enters non linearly in the coefficient, where σ is a non-trivial diffusion term and B is a d-dimensional fractional Brownian motion. We have thus decided to tackle this important scientific challenge first.

To this aim, here are the steps we have focused on in 2013:

- A better understanding of the underlying rough path structure for equation (4), carried out in [6]. This step allows a proper definition of stochastic integrals with respect to fractional Brownian motion in a wide range of contexts.
- Extension of pathwise stochastic integration to processes indexed by the plane in [19], which helps to the definition of noisy systems such as partial differential equations.
- Gaussian type bounds for equations driven by a fractional Brownian motion, obtained in [18], [7]. This is an important preliminary step for likelihood estimates for stochastic processes. Also notice the interesting central limit theorems exhibited in [24], in a context which is similar to our equation of interest.
- Numerical aspects of a maximum likelihood type procedure for an equation of the form (4), expressed in terms of Malliavin calculus tools (see [4]).

6.3.2. LAN property for fractional Brownian motion

We have first focused on an important statistical property of fractional Brownian paths on their own. Indeed, the local asymptotic normality (LAN) property is a fundamental concept in asymptotic statistics, which gives the asymptotic normality of certain estimators such as the maximum likelihood estimator for instance. In [5], we focus on the LAN property for the model where we observe a sample of n observations $\mathbf{X_n} = (X_1, ..., X_n)$ of a Gaussian stationary sequence. The sequence $(X_n)_{n\in\mathbb{N}}$, whose spectral density f_θ is indexed by a parameter θ, can admit antipersistence, long memory or short memory and be noninvertible. To be more specific, our main assumption is:

$$f_\theta(x) \sim_{x\to 0} |x|^{-\alpha(\theta)} L_\theta(x)$$

with L_θ a slowly varying function and $\alpha(\theta) \in (-\infty, 1)$. We prove the LAN property by studying an asymptotic expansion of the log likelihood and using some results on Toeplitz matrices. In particular, our assumptions are fulfilled by fractional Gaussian noises and autoregressive fractionally integrated moving average processes (ARFIMA(p, d, q)). We also obtain the LAN property for fractional Brownian motion.

6.3.3. Self-similarity properties and stable or Gaussian random fields

In 2009, C. Lacaux and H. Biermé carried on the study of some sample paths properties for an important class of anisotropic random fields called operator scaling random fields, which had been previously introduced by H. Biermé, M. Meerschaert and P. Scheffler (2007). To be more specific, the classical self-similarity property is replaced by the following operator scaling property:

$$\forall c > 0,\ \left(X(c^E x)\right)_{x\in\mathbb{R}^d} \overset{(d)}{=} c\left(X(x)\right)_{x\in\mathbb{R}^d}, \tag{5}$$

where $c^E := \exp\left(E\ \ln(c)\right)$. In particular, the Hölder regularity properties of operator scaling Gaussian or stable harmonizable random fields have been expressed in terms of the matrix E. The method they used can be applied to study the modulus of continuity of many stable or Gaussian random fields. As example in 2011, with P. Scheffler, they have followed it to study multi-operator harmonizable stable random fields, which satisfy a local version of the operator scaling property and enjoy a regularity which may vary along the trajectories. In [20], it has been developed in the more general framework of conditionally sub-Gaussian random series. This allows to also study for example some multistable random fields, which have been introduced in (Falconer & al, 2009); for such a field X, the marginal $X(x)$ is a stable random variable whose index of stability can depend on x. In this paper, some conditions have been proposed to establish the uniform convergence of the series (on an eventually random ball), an upper bound for the modulus of continuity of its limit, an uniform control of the partial series ones and an explicit rate of convergence. Focusing on LePage random series, upper bounds of the modulus of continuity of some harmonizable stable or multistable random fields are provided. In the conference paper [11], [20] has then been applied to study the class of linear multifractional multistable motions. In particular, the upper bound obtained for the modulus of linear multifractional stable motion is the sharpest available.

We are also interested in self-similar processes indexed by manifolds in [8]. This study is motivated by the fact various spatial data are indexed by a manifold and not by the Euclidean space $\mathbb{R}^d$ in practical situations such as image analysis.

BIOCORE Project-Team

6. New Results

6.1. Mathematical methods and methodological approach to biology

6.1.1. Mathematical analysis of biological models

6.1.1.1. Mathematical study of semi-discrete models

Participants: Jean-Luc Gouzé, Frédéric Grognard, Ludovic Mailleret, Pierre Bernhard, Elsa Rousseau, Nicolas Bajeux.

Semi-discrete models have shown their relevance in the modeling of biological phenomena whose nature presents abrupt changes over the course of their evolution [99]. We used such models and analysed their properties in several practical situations that are developed in Section 6.2.2 , most of them requiring such a modeling in order to take seasonality into account. Such is the case when the year is divided into a cropping season and a 'winter' season, where the crop is absent, as in our analysis of the sustainable management of crop resistance to pathogens [53] or in the co-existence analysis of epidemiological strains [21]. Seasonality also plays a big role in the semi-discrete modeling required for the analysis of consumers' adaptive behavior in seasonal consumer-resource dynamics, where only dormant offspring survive the 'winter' [61].

6.1.1.2. Model design, identification and validation

Participants: Olivier Bernard, Francis Mairet.

One of the main families of biological systems that we have studied involves mass transfer between compartments, whether these compartments are microorganisms or chemical species in a bioreactor, or species populations in an ecosystem. We have developed methods to estimate the models of such systems [79]. These systems can be represented by models having the general structure popularized by [78], [84], and based on an underlying reaction network:

$$\frac{d\xi}{dt} = K\, r(\xi,\psi) + D(\xi_{in} - \xi) - Q(\xi)$$

We address two problems: the determination of the pseudo-stoichiometric matrix K and the modelling of the reaction rates $r(\xi,\psi)$.

In order to identify K, a two-step procedure has been proposed. The first step is the identification of the minimum number of reactions to be taken into account to explain a set of data. If additional information on the process structure is available, we showed how to apply the second step: the estimation of the pseudo-stoichiometric coefficients.

This approach has been applied to various bioproduction processes, among which activated sludge processes [77], anaerobic digestion [92], [106] and anaerobic digestion of microalgae [100]. Recently it was also used to reduce the ADM1 model in the case of winery effluent wastewater [24].

6.1.2. Metabolic and genomic models

Participants: Jean-Luc Gouzé, Madalena Chaves, Alfonso Carta, Ismail Belgacem, Olivier Bernard, Caroline Baroukh, Rafael Muñoz-Tamayo, Jean-Philippe Steyer.

Global stability for metabolic models and full Michaelis-Menten equations

With techniques of monotone and compartmental systems, we studied full (i.e. not reduced by any time-scale argument) Michelis-Menten reactions or chains of reactions: we prove global stability when the equilibrium exists, and show that it may not exist. This fact has important consequences for reduction of metabolic systems in a coupled genetic-metabolic system [17].

Structural principles for the existence of limit cycles in two-dimensional piecewise affine models

Using concavity and continuity properties of Poincaré maps, we have derived some structural principles which link the topology of the transition graph to the existence, number and stability of limit cycles in a class of two-dimensional piecewise affine biological models [13].

Probabilistic approach for predicting periodic orbits in piecewise affine models

In the state transition graph, a *transition probability* between two nodes can be defined in terms of the parameters of the piecewise affine models. For a cyclic transition graph, this approach can be used to predict the most likely periodic orbit for a given a set of parameters [22].

Growth rate models in bacteria: piecewise affine systems with a dilution term

We have extended the class of piecewise affine systems to deal with dynamics dependent on dilution due to cell growth rate. Considering that growth rate is determined by two limiting factors (RNA polymerase and ribosomes), in [42] we propose and analyze a switched system with two piecewise quadratic modes. This is part of the PhD thesis of Alfonso Carta, and done in collaboration with IBIS project-team.

Transcription and translation models in bacteria

We study detailed models of transcription and translation for genes in a bacterium. With techniques of monotone systems, and time scale hypotheses, we can show the stability of the fast part of these systems, and reduce them to much smaller models [40], [39]. We also study other models of the global cellular machinery. This is part of the PhD theses of Ismael Belgacem, Alfonso Carta, and done in collaboration with IBIS project-team. Moreover, in collaboration with IBIS, we studied and experimentally validated the time scale reduction of the classical two-step model for gene expression [51]

Analysis of circadian rhythms in cyanobacteria

A hierarchy of models (from Boolean to continuous) was used in [23] to successively characterize the wiring structure, qualitative dynamical properties, and then perform parameter estimation on a model describing the system responsible for the circadian rhythm of cyanobacteria.

Interconnections of Boolean modules: asymptotic and transient behaviour

The asymptotic dynamics of high-dimensional networks (e.g., genetic) can be obtained from the interconnection of two input/output Boolean subnetworks, and the analysis of their attractors. This computational cost reducing method is described in [34]. Some extensions include the characterization of the attractors of the interconnected system in terms of invariant sets.

Structure estimation for Boolean models of gene regulation networks

The problem of estimating Boolean models of gene networks from few and noisy measurements is addressed in [41], joint work with C. Breindl and F. Allgöwer from the University of Stuttgart. The class of unate or canalizing Boolean functions is considered and represented by multi-affine polynomials, leading to a reformulation of the estimation problem as a mixed integer linear program.

Analysis of dynamical systems by combining discrete and continuous formalisms

The work reviewed in the HDR of M. Chaves [11] highlights methods of analysis that use and combine techniques from discrete and piecewise affine modeling formalisms, such as construction of the transition graph and its association with the parameters of the system. Some basic methods for generating a discrete transition graph from a given continuous system are described in the internship project of F. Todoran [75].

State estimation for gene networks

We address state estimation for gene regulatory networks with intrinsic and extrinsic noise at the level of single cells. We take the Chemical Master Equation (CME) with random parameters as a reference modeling approach, and investigate the use of stochastic differential model approximations for the construction of practical real-time filters (based on non-linear Kalman filtering) [43]. This is a collaboration with Ibis team.

Modelling the metabolic network in non balanced growth conditions

On the basis of the knowledge of the metabolic network, we propose a new methodology to go beyond the “balanced growth paradigm”(assuming that there is no storage within the cell). We have therefore a tool to represent the possible storage of some key biochemical compounds. This approach was applied to describe the effect of both a light cycle and a nitrogen starvation on the lipid accumulation [37]. The first stage of the approach consists in splitting the metabolic network into sub-networks, which are assumed to satisfy balanced growth condition. The left metabolites interconnecting the sub-networks are allowed to behave dynamically. Then, thanks to Elementary Flux Mode analysis, each sub-network is reduced to macroscopic reactions, for which simple kinetics are assumed. This approach was applied to the accumulation of lipids and carbohydrates of the microalgae *Tisochrysis lutea* under day/night cycles. The resulting model described accurately experimental data obtained in day/night conditions; it efficiently predicts the accumulation and consumption of lipids and carbohydrates.

6.2. Fields of application

6.2.1. Bioenergy

6.2.1.1. Modelling of microalgae production

Participants: Olivier Bernard, Antoine Sciandra, Frédéric Grognard, Philipp Hartmann, Rafael Muñoz-Tamayo, Ghjuvan Grimaud, David Demory, Frédéric Chazalon, Hubert Bonnefond, Jean-Philippe Steyer, Francis Mairet.

Experimental developments

Experiments have been carried out to study the effects of nitrogen limitation on the lipid production in microalgae and support model development. These experiments have been carried out in the Lagrangian simulator, under constant or periodic light and temperature, varying the total amount of light dose in the day. The response in terms of storage carbon (triglycerides and carbohydrates) has been observed.

Other experiments were carried out to reproduce the light percept by a cell in a raceway pond [74], that is a large-scale raceway-track shaped open-air photobioreactor with circulating medium. An electronic platform was developed to reproduce the flashing light which, from the hydrodynamical studies, is likely to happen in a raceway at the cell scale. The experiments show that the microalgae adapt their pigments to the average light that they have received.

The effect in the cell cycle of both the light periodic signal and a nitrogen limitation were studied. The strong interactions of the interactions between the different phases of the cell cycle through checkpoints was highlighted [104].

Finally, we have tested the effect of cement flue gas on microalgae growth and demonstrated that this CO_2 source can be used to feed microalgal industrial cultures [33].

These works have been carried out in collaboration with A. Talec, S. Rabouille, E. Pruvost and C. Combe (CNRS/UPMC -Océanographic Laboratory of Villefranchesur-Mer).

In collaboration with the IFREMER-PBA team (Nantes) we contributed to a study (within the Symbiose project) of the possible associations between microalgae and bacteria to enhance overall productivity [27].

Metabolism of carbon storage and lipid production

A macroscopic model for lipid production by oleaginous microalgae [10] has been previously proposed. This model describes the accumulation of neutral lipids (which can be turned into biofuel), carbohydrates and structural carbon. We now start to progressively dig deeper in the metabolism, with the objective to better predict carbohydrate and lipid accumulation[37], [64].

Modeling the coupling between hydrodynamics and biology

In collaboration with the Inria ANGE team, a model coupling the hydrodynamics of the raceway (based on multilayer Saint Venant system) with microalgae growth was developed [86]. This model is supported by the work of ANGE aiming at reproducing the hydrodynamics of the raceway, with a specific attention to the effect of the paddle wheel on the fluid.

Modeling the photosynthesis response to fast fluctuating light

The impact of the hydrodynamics on the light percept by a single cell was studied thanks to fluid dynamics simulations of a raceway pond [26]. The light signals that a cell experiences at the Lagrangian scale, depending on the fluid velocity, were then estimated. A Droop-Han model was used to assess the impact of light fluctuation on photosynthesis. A new model accounting for photoacclimation was also proposed [46]. Single cell trajectories were simulated by this tool, and the effect on photosynthesis efficiency was assessed using models of photosynthesis [94]. These results were compared to experimental measurements where the high frequency light was reproduced [74].

Modeling a microalgae production process

The integration of different models developed in the group [81], [101], [10] was performed to represent the dynamics of microalgae growth and lipid production in raceway systems, on the basis of the dynamical model developed to describe microalgal growth in a photobioreactor under light and nitrogen limitations. The strength of this model is that it takes into account the strong interactions between the biological phenomena (effects of light and nitrogen on growth, photoacclimation ...), temperature effect [85],[31] and the radiative transfer in the culture (light attenuation due to the microalgae).
Using these approaches, we have developed a model which predicts lipid production in raceway systems under varying light, nutrients and temperature [30]. This model is used to predict lipid production in the perspective of large scale biofuel production.
Finally, we provide guidelines for the design of experiments with high informative content that allows an accurate parameter estimation of this model, concerning the effect of temperature and light on microalgae growth. The optimal experiment design problem was solved as an optimal control problem. E-optimal experiments were obtained by using two discretization approaches namely sequential and simultaneous. Simulation results showed the relevance of determining optimal experimental inputs for achieving an accurate parameter estimation [50].

Nitrogen fixation by nitrogenotrophs

The fixation of nitrogen by *Croccosphera watsonii* was represented with a macro metabolic model [44]. The main fluxes of carbon and nitrogen are represented in the cell. The accumulation of starch during the day to fuel the nitrogenase working in the absence of oxygen during the night was the key process to explain the nitrogen fixation. The strong influence of the cell cycle was also included in the model. Finally, the model was calibrated and validated with the data of 3 experiments carried out with different duration of the light period and daily dose. The model succeded to efficiently reproduce the experimental data.

This work is done in collaboration with Sophie Rabouille (CNRS-Océanographic Laboratory of Villefranche-sur-Mer).

Including phytoplankton photoadaptation into biogeochemical models

The complexity of the marine ecosystem models and the representation of biological processes, such as photoadaptation, is very challenging to tackle so that their representation remains an open question. We compared several marine ecosystem models with increasing complexity in the phytoplankton physiology representation in order to assess the consequences of the complexity of photoadaptation models in biogeochemical model predictions. Three models of increasing complexity were considered, and the models were calibrated to reproduce ocean data acquired at the Bermuda Atlantic Time-series Study (BATS) from in situ JGOFS (Joint Global Ocean Flux Study) data. It turns out that the more complex models are trickier to calibrate and that intermediate complexity models, with an adapted calibration procedure, have a better prediction capability [15].

This work is done in collaboration with Sakina Ayata (UPMC-Océanographic Laboratory of Villefranche-sur-Mer).

6.2.1.2. Control and Optimization of microalgae production

On-line monitoring

Interval observers give an interval estimation of the state variables, provided that intervals for the unknown quantities (initial conditions, parameters, inputs) are known [7]. Several developments were carried out in this direction to improve the design and performances of interval observers. The approach has been applied to estimation of the microalgae growth and lipid production within a production process [28].

Optimization of the bioenergy production systems

Based on simple microalgae models, analytical optimization strategies were proposed. We first focused on the optimal operating conditions for the biomass productivity under day/night cycles using Pontryiagin's maximum principle (assuming a periodic working mode) [25].
On the other hand, we assessed strategies for optimal operation in continuous mode using the detailed model for raceways [49], [30]. Two strategies were developed. The first one resides in solving numerically an optimal control problem in which the input flow rate of the raceway is calculated such that the productivity in microalgae biomass is maximized on a finite time horizon. In the second strategy, we aimed at translating the optimization problem into a regulation problem. We proposed a simple operational criterion that when integrated in a strategy of closed-loop control allows to attain biomass productivities very near to the maximal productivities obtained with the optimal control. We demonstrated that the practical advantages for real implementation makes our proposed controller a suitable control strategy for optimizing microalgae production in raceways.
We also propose a nonlinear adaptive controller for light-limited microalgae culture, which regulates the light absorption factor (defined by the ratio between the incident light and the light at the bottom of the reactor). We show by numerical simulation that this adaptive controller can be used to obtain near optimal productivity under day-night cycles [47].

Interactions between species

Large scale culture of microalgae for bioenergy involves a huge biodiversity (different mutants, invasion, growth-promoting bacteria [96]...). Control of such system requires to consider the interactions between the different species.

In the framework of the ANR Facteur 4 project, we propose to drive this competition exploring different strategies in order to select species of interest.

We have proposed an adaptive controller which regulates the light at the bottom of the reactor [48]. When applied for a culture with n species, the control law allows the selection of the strain with the maximum growth rate for a given range of light intensity. This is of particular interest for optimizing biomass production as species adapted to high light levels (with low photoinhibition) can be selected.
Other strategies (e.g. periodic temperature stress) are now under investigation through simulations (in order to design selection experiments that will be performed at LOV) and model analysis.
Finally, in a more theoretical framework, we studied how to select as fast as possible a given species in a chemostat with two species at the initial instant. Using the Pontryagin maximum principle, we have shown that the optimal strategy is to maintain the substrate concentration to the value maximizing the difference between the growth rates of two species [66].

6.2.2. Design of ecologically friendly plant production systems

6.2.2.1. Controlling plant pests

Participants: Frédéric Grognard, Ludovic Mailleret, Suzanne Touzeau, Mickaël Teixeira-Alves, Nicolas Bajeux.

Optimization of biological control agent introductions

The question of how many and how frequently natural enemies should be introduced into crops to most efficiently fight a pest species is an important issue of integrated pest management. The topic of natural enemies introductions optimization has been investigated for several years [9] [105], unveiling the crucial influence of within-predator density dependent processes. Because contrarily to predatory biocontrol agents, parasitoids may be more prone to exhibit positive density dependent dynamics rather than negative ones, the current modeling effort concentrates on studying the impact of positive predator-predator interactions on the optimal introduction strategies [72]. Connected experimental research is also being pursued in the laboratory on *trichogramma spp.* which tends to show positive density dependence because of demographic stochasticity [35], and the PhD thesis of Thibaut Morel Journel (UMR ISA) has just started on this topic. Non-impulsive positive feedback control of predator-prey systems in that framework was also addressed in [45].

Food source diversity and classical biological control efficiency using generalist natural enemies

Because generalist biocontrol agents can feed on different food sources like, e.g. a given pest and pollen, they are capable of surviving pest absence within crops and, when supplied with different food types, generalist biocontrol agents are expected to thrive. However, feeding on different food sources means that a given individual cannot feed on each food source at the same moment, which thus potentially reduces the overall predation pressure imposed by the natural enemy population. We exhibited conditions under which the predator distraction effects can dominate the demographic response of the predator populations, potentially disrupting pest control [12]. Such results were at the center of Mickaël Teixeira Alves's PhD thesis.

Plant compensation, pest control and plant-pest dynamics

Introducing a plant compartment into our models, we first focused on plant-insect interactions and showed how the level and timing of the pest invasion and pests control interventions could have important effects on the plant's growth pattern and its final biomass. We then modelled plant compensation, which is the process by which some plants respond positively to recover from the effects of pest injury. We have shown that depending on plants and pests characteristics, as well as the level of pest attack, plant overcompensation may or may not happen [97].

This work is part of the PhD thesis of Audrey Lebon (Cirad), and done in collaboration with Yves Dumont (Cirad).

6.2.2.2. Controlling plant pathogens

Participants: Frédéric Grognard, Ludovic Mailleret, Suzanne Touzeau, Elsa Rousseau.

Sustainable management of plant resistance

Because in addition to being eaten, plants can also get sick, we studied other forms of biological control dedicated to fight plant pathogens. One such method is the introduction of plant strains that are resistant to one pathogen. This often leads to the appearance of virulent pathogenic strains that are capable of infecting the resistant plants. It is therefore necessary to develop ways of introducing such resistance into crop production without jeopardizing its future efficiency. Considering plant viruses, we computed the proportion of resistant plants that should be cropped together with the non-resistant ones in a seasonal model, in order to optimize the resistance for production or patrimonial objectives [53]. The study of factors influencing resistance breakdown from the within-plant to the landscape level is the topic of Elsa Rousseau's PhD thesis, with emphasis both on experimental and modelling approaches. Experiments have been held in Avignon to determine the respective impacts of selection and genetic drift on resistance breakdown.

This work is done in collaboration with Frédéric Fabre and Benoit Moury (INRA Avignon).

Eco-evolutionary dynamics of plant pathogens in seasonal environments

Understanding better pathogen evolution also requires to understand how closely related plant parasites may coexist. Indeed, such coexistence is widespread and is hardly explained through resource specialization. We showed that, in agricultural systems in temperate environments, the seasonal character of agrosystems can induce complex plant-pathogens dynamics [98] and is an important force promoting evolutionary diversification of plant pathogens [93]. Plant parasites reproduction mode may also strongly interact with seasonality. In this context, we investigated the influence of cyclical parthenogenesis, i.e. the alternation of sexual and asexual reproduction phases, on the eco-evolutionary dynamics of plant parasites [59], [60], [21].

This work is part of the PhD thesis of Magda Castel (Agrocampus Ouest) and is done in collaboration with Frédéric Hamelin (Agrocampus Ouest).

6.2.3. Biological depollution

6.2.3.1. Coupling microalgae to anaerobic digestion

Participants: Olivier Bernard, Antoine Sciandra, Jean-Philippe Steyer, Frédéric Grognard, Philipp Hartmann, Francis Mairet.

The coupling between a microalgal pond and an anaerobic digester is a promising alternative for sustainable energy production and wastewater treatment by transforming carbon dioxide into methane using light energy. The ANR Symbiose project is aiming at evaluating the potential of this process [108], [107].

In a first stage, we developed models for anaerobic digestion of microalgae. Two approaches were used: first, a dynamic model has been developed trying to keep a low level of complexity so that it can be mathematically tractable for optimisation [100] . Considering three main reactions, this model fits adequately the experimental data of an anaerobic digester fed with *Chlorella vulgaris* (data from INRA LBE). On the other hand, we have tested the ability of ADM1 [109] (a reference model which considers 19 biochemical reactions) to represent the same dataset. This model, after modification of the hydrolysis step [102] has then been used to evaluate process performances (methane yield, productivity...) and stability though numerical simulations.

6.2.3.2. Life Cycle Assessment

Participants: Olivier Bernard, Jean-Philippe Steyer.

This work is the result of a collaboration with Laurent Lardon and Arnaud Helias of INRA-LBE through the co-supervision of Pierre Collet's PhD thesis [88].

An analysis of the potential environmental impacts of biodiesel production from microalgae has been carried out using the life cycle assessment (LCA) methodology [95]. This study has allowed to identify the obstacles and limitations which should receive specific research efforts to make this process environmentally sustainable. This study has been updated and the effects of technological improvements (leading to higher productivities) have been compared to the source of electricity. It turns out that the overall environmental balance can much more easily be improved when renewable electricty is produced on the plant [91], [90]. As a consequence, a new paradigm to transform solar energy (in the large) into transportation biofuel is proposed, including a simultaneous energy production stage. This motivated the design of the purple sun ANR-project.

These studies have allowed to identify the obstacles and limitations which should receive specific research efforts to make this process environmentally sustainable [65].

A LCA has been carried out to assess the environmental impact of methane production by coupling microalgae and anaerobic digestion. The study highlights the limitation derived by the low biodegradability of the considered microalgae [89] which induces a large digester design and thus more energy to mix and heat it.

These works have been carried out in collaboration with E. Latrille and B. Sialve (INRA - Laboratory of Environmental Biotechnology, Narbonne).

6.2.4. Models of ecosystems

6.2.4.1. Optimality/games in population dynamics

Participants: Frédéric Grognard, Ludovic Mailleret, Pierre Bernhard.

Adaptive behavior in seasonal consumer-resource dynamics

In this work we studied the evolution of a consumer-resource (or predator-prey) system with seasonal character of the dynamics. We specified two main parts of the process. First, we considered the system during one season with a fixed length: the prey lay eggs continuously and the predators lay eggs or hunt the prey (choose their behavior) according to the solution of an optimal control problem [76]. We then showed that, in most situations, mutants can take advantage of their low frequency and fare better than the residents. Over the course of a large number of seasons, the mutants replace the residents, only to find themselves applying the original resident behavior [61].

Optimal foraging and residence times variations

Charnov's marginal value theorem (MVT) [87] is a central tenet of ecological theory. In fragmented environments, the MVT connects the quality and distribution of patches to the optimal time an individual should spend on any patch, and thus the rate of movement in the habitat. Unfortunately, it does not offer explicit predictions regarding how changing habitat quality would affect residence times. In this work, we answer that question in a very general setting, for habitats with homogeneous or heterogeneous patches and with general fitness functions. We then particularize it to the resource consumption framework and indicate how the residence times variations relate to the curvatures of the functional responses,[20].

This last work is done in collaboration with Vincent Calcagno and Eric Wajnberg (INRA Sophia Antipolis)

The handicap paradox

We have investigated the "handicap paradox" of sexual selection, and more specifically revisited Grafen's mathematical models of Zahavi's "handicap principle". The paradox is that in many species, male secondary sexual characters that clearly attract the females are so developed as to be a handicap to the male's viability, and therefore should be counter-selected by evolution. Zahavi's explanation, made mathematical by Grafen, is that if this secondary sexual character is a signal to the female of the male's quality that she cannot observe otherwise, if this signal were costless, it could be cheated, a low quality male being induced to mimic the signal of a high quality one. We have cast this problem into a signaling game, using the bayesian equilibrium of game theory. This easily shows that indeed, under mild conditions, at equilibrium the signal should be "costly". We have developed several models inspired by Grafen, and to a lesser extent Getty, with explicit solutions, and explained why an undesirable feature appeared in Grafen's model (as well as in one of ours) and proposed a model free of this artifact [19].

6.3. Software design

6.3.1. Odin

Participants: Olivier Bernard, Mélaine Gautier.

Over the years, BIOCORE has been developing a software framework for bioprocess control and supervision called ODIN [80]. This C++ application (working under Windows and Linux) enables researchers and industrials to easily develop and deploy advanced control algorithms through the use of a Scilab interpreter [82], [83]. It also contains a Scilab-based process simulator which can be harnessed for experimentation and training purposes. ODIN is primarily developed in the C++ programming language and uses CORBA to define component interfaces and provide component isolation. ODIN is a distributed platform, enabling remote monitoring of the controlled processes as well as remote data acquisition. Recently, a software development effort has been directed to the graphical user interface, a synoptic view component, new drivers for the experimental hardware and integration of the PlantML data exchange format. ODIN has been tested on four different processes and has been set up with Eric Latrille to supervise the 66m2 high rate pond at the LBE, INRA Narbonne.

6.3.2. In@lgae

Participants: Etienne Delclaux, Francis Mairet, Olivier Bernard.

The simulation platform In@lgae is jointly developed with the Inria Ange team. Its objective is to simulate the productivity of a microalgae production system, taking into account both the process type and its location and time of the year. A first module (Freshkiss) developed by Ange computes the hydrodynamics, and reconstructs the Lagrangian trajectories percept by the cells. Coupled with the Han model, it results in the computation of an overall photosynthesis yield. A second module is coupled with a GIS (geographic information system) to take into account the meteorology of the considered area (any location on earth). The evolution of the temperature in the culture medium together with the solar flux is then computed. Finally, the productivity in terms of biomass, lipids, pigments together with CO_2, nutrients, water consumption, ... are assessed. The productivity map which is produced can then be coupled with a resource map describing the availability in CO_2 nutrients and land.

BIPOP Project-Team

6. New Results

6.1. Multiple impacts modelling

Participants: Bernard Brogliato, Ngoc-Son Nguyen.

The work consists of studying two systems: the rocking block and tapered chains of balls, using the Darboux-Keller model of multiple impacts previously developed. The objectives are threefold: 1) show that the model predicts well the motion by careful comparisons with experimental data found in the literature, 2) study the system's dynamics and extract critical kinetic angles that allow the engineer to predict the system's gross motion, 3) develop numerical code inside the SICONOS platform that incorporates the model of multiple impact. The influence of the kinetic angles in the rocking block motion with friction is analysed as well, numerically. Extensive experimental works have been conducted by our colleague C. Liu at PKU on a disc-ball system. Results are in [30], [24] [64], and in the monograph [15]. Another work is dedicated to analysing the influence of bilateral holonomic constraints on the well-posedness of the complementarity problem obtained from the (frictionless) unilateral constraints. Gauss' principle extension to this case is also analysed [22].

6.2. Discrete-time sliding mode control

Participants: Vincent Acary, Bernard Brogliato, Olivier Huber, Bin Wang.

This topic concerns the study of time-discretized sliding-mode controllers. Inspired by the discretization of nonsmooth mechanical systems, we propose implicit discretizations of discontinuous, set-valued controllers. This is shown to result in preservation of essential properties like simplicity of the parameters tuning, suppression of numerical chattering, reachability of the sliding surface after a finite number of steps, and disturbance attenuation by a factor h or h^2 [36]. This work is part of the ANR project CHASLIM. Within the framework of CHASLIM we have performed many experimental validations on the electropneumatic setup of IRCCyN (Nantes), which nicely confirm our theoretical and numerical predictions: the implicit implementation of sliding mode control, drastically improves the input and output chattering behaviours. In particular the high frequency bang-bang controllers which are observed with explicit discretizations, are completely suppressed.

6.3. Dissipativity preserving methods

Participants: Vincent Acary, Bernard Brogliato.

This work concerns the analysis of so-called theta-methods applied to linear complementarity systems that are dissipative (in the sense of Willems). Necessary and sufficient conditions for dissipativity preservation after the time-discretization are derived (preservation of the storage function, the supply rate and the dissipation function). The possible state jumps are also analyzed [57]. It is shown that excepted when the system is state lossless and theta = 0.5, the conditions for dissipativity preservation are very stringent. In this article we also provide (for the first time, to the best of our knowledge) a rigorous definition of numerical dissipation, which remained until now a vague notion in numerical analysis.

6.4. Lur'e set-valued dynamical systems

Participants: Bernard Brogliato, Aneel Tanwani, Christophe Prieur.

Lur'e systems are quite popular in Automatic Control since the fifties. Set-valued Lur'e systems possess a static feedback nonlinearity that is a multivalued function. This study consists in the mathematical analysis (existence and uniqueness of solutions) and the stability analysis (Lyapunov stability, invariance principle) of classes of set-valued Lur'e systems, with applications in complementarity dynamical systems, relay systems, mechanical systems with dry friction, electrical circuits, etc. Our works in this field started in [51]. The results in [53] extend those in [52] with an accurate characterization of the maximal monotonicity of the central operator of these systems, which consists of a projection-like operator. Concrete and verifiable criteria are provided for the above classes (complementarity, relay systems). Results on state observers for classes of Lur'e systems (namely: Moreau's sweeping process of first and second order, and with prox-regular sets) are proposed in [47], [39]. Therein the convexity is replaced by the far more general notion of prox-regularity, which destroys the monotonicity.

6.5. Analysis of Limit Cycles in Piecewise Linear Systems

Participants: Vincent Acary, Bernard Brogliato, Valentina Sessa.

Autonomous piecewise linear systems in the Lur'e form may exhibit periodic steady-state oscillations. For many practical systems belonging to this class the period and the shape of the oscillation is difficult to be predicted a priori. In this work the complementarity approach is used to tackle the issue. The complementarity formalism is used to represent the closed-loop system and a phase condition acting as an anchor equation for the periodic solution. By discretizing the dynamics a mixed complementarity problem is formulated. The corresponding solution provides an accurate prediction of the steady-state oscillation and its period. Numerical results show the effectiveness of the proposed technique for the computation of stable and sliding periodic solutions. The analysis of the steady-state solution of a Colpitts oscillator is considered as an illustration. This work has been presented at CDC 2013 in [37].

6.6. Simulation and stability of piecewise linear gene networks

Participants: Vincent Acary, Arnaud Tonnelier, Bernard Brogliato.

This work has been done in collaboration with the IBIS project team, it is reported in [45], [19]. Gene regulatory networks control the response of living cells to changes in their environment. A class of piecewise-linear (PWL) models, which capture the switch-like interactions between genes by means of step functions, has been found useful for describing the dynamics of gene regulatory networks. The step functions lead to discontinuities in the right-hand side of the differential equations. This has motivated extensions of the PWL models based on differential inclusions and Filippov solutions, whose analysis requires sophisticated numerical tools. We present a method for the numerical analysis of one proposed extension, called Aizerman-Pyatnitskii (AP)-extension, by reformulating the PWL models as a mixed complementarity system (MCS). This allows the application of powerful methods developed for this class of nonsmooth dynamical systems, in particular those implemented in the Siconos platform. We also show that under a set of reasonable biological assumptions, putting constraints on the right-hand side of the PWL models, AP-extensions and classical Filippov (F)-extensions are equivalent. This means that the proposed numerical method is valid for a range of different solution concepts. We illustrate the practical interest of our approach through the numerical analysis of three well-known networks developed in the field of synthetic biology.

In addition, we have investigated oscillatory regimes in repressilator-type models with piecewise linear dynamics [48]. We derived exact analytical conditions for oscillations and showed that the relative location between the dissociation constants of the Hill functions and the ratio of kinetic parameters determines the possibility of oscillatory activities. We also computed analytically the probability of oscillations. Results suggest that a switch-like coupling behaviour, a time-scale separation and a repressilator-type architecture with an even number of elements facilitate the emergence of sustained oscillations in biological systems.

6.7. Numerical analysis and simulation of mechanical systems with constraints

6.7.1. Event-capturing schemes for nonsmooth mechanical systems

Participant: Vincent Acary.

To perform the numerical time integration of nonsmooth mechanical systems, the family of event-capturing time–stepping schemes are the most robust and efficient tools. Nevertheless, they suffer from several drawbacks : a) a low–order accuracy (at best at order one), b) a drift phenomena when the unilateral constraints are treated at the velocity level and c) a poor "energetic" behavior in terms of stabilizing the high–frequency dynamics. We proposed self-adapting schemes by applying time–discontinuous Galerkin methods to the measure differential equation in [31]. In order to satisfy in discrete time, the impact law and the constraints at the position and the velocity level, an adaptation of the well-known Gear–Gupta–Leimkuhler approach has been developed in [18]. Finally, the energetic behavior of the standard Moreau–Jean scheme has been addressed in [26] by developing a Newmark–type scheme for nonsmooth dynamics.

6.7.2. *Numerical time–integration methods for event-detecting schemes.*

Participants: Vincent Acary, Bernard Brogliato, Mounia Haddouni.

The CIFRE thesis of M. Haddouni concerns the numerical simulation of mechanical systems subject to holonomic bilateral constraints, unilateral constraints and impacts. This work is performed in collaboration with ANSYS and the main goal is to improve the numerical time–integration in the framework of event-detecting schemes. Between nonsmooth events, time integration amounts to numerically solving a differential algebraic equations (DAE) of index 3. We have compared dedicated solvers (Explicit RK schemes, Half-explicit schemes, generalizes α-schemes) that solve reduced index formulations of these systems. Since the drift of the constraints is crucial for the robustness of the simulation through the evaluation of the index sets of active contacts, we have proposed some recommendations on the use of the solvers of dedicated to index-2 DAE. This work has been presented in [35], [40].

6.7.3. *Multibody systems woth contact, friction and clearances*

Participants: Vincent Acary, Bernard Brogliato, Narendra Akadkhar.

The PhD thesis of N. Akadkhar under contract with Schneider Electric concerns the numerical simulation of mechanical systems with unilateral constraints and friction, where the presence of clearances in imperfect joints plays a crucial role. A first work deals with four-bar planar mechanisms with clearances at the joints, which induce unilateral constraints and impacts, rendering the dynamics nonsmooth. The objective is to determine sets of parameters (clearance value, restitution coefficients, friction coefficients) such that the system's trajectories stay in a neighborhood of the ideal mechanism (*i.e.* without clearance) trajectories. The analysis is based on numerical simulations obtained with the projected Moreau-Jean time-stepping scheme. These results have been submitted to the ENOC 2014 conference. It is planned to extend these simulations to frictional cases and to mechanisms of circuit breakers.

6.8. Mechanical rods

6.8.1. *High-order models of mechanical rods*

Participants: Florence Bertails-Descoubes, Romain Casati.

Reduced-coordinate models for rods such as the articulated rigid body model or the super-helix model [50] are able to capture the bending and twisting deformations of thin elastic rods while strictly and robustly avoiding stretching deformations. In this work we are exploring new reduced-coordinate models based on a higher-order geometry. Typically, elements are defined by a polynomial curvature function of the arc length, of degree $d \geq 1$. The main difficulty compared to the super-helix model (where $d = 0$) is that the kinematics has no longer a closed form. Last year, in R. Casati's PhD's thesis, we extended this result to the full 3D case. The key idea was to integrate the rod's kinematics using power series expansion, and to design an accurate and efficient computational algorithm adapted to floating point arithmetics. Our method nicely propagates to the computation of the full dynamic of a linked chain of 3d clothoid. This year we thoroughly compared our methods against other rod models from the literature, in terms of both accuracy and computational efficiency. Our results demonstrate that our model is competitive compared to former models, and yields a better trade-off in the case of highly curly rods. All these results were published and presented this year at SIGGRAPH [25]. The source code is also freely distributed under a GPLv.3 license (see Section 5.3).

6.8.2. Inverse modeling of mechanical rods subject to frictional contact

Participants: Florence Bertails-Descoubes, Alexandre Derouet-Jourdan, Gilles Daviet.

Controlling the input shape of slender structures such as rods is desirable in many design applications (such as hairstyling, reverse engineering, etc.), but solving the corresponding inverse problem is not straightforward. In [54], [55] we noted that reduced-coordinates models such as the super-helix are well-suited for static inversion in presence of gravity.

We are facing two main difficulties: 1/ the geometrical fitting of a piecewise helix to an arbitrary input curve and 2/ the inversion a super-helix subject to gravity *and* contacting forces.

6.8.2.1. Geometrical fitting: from an arbitrary smooth curve to a C^1 piecewise helix

Participants: Florence Bertails-Descoubes, Alexandre Derouet-Jourdan.

In A. Derouet-Jourdan's PhD's thesis (co-supervised by Joëlle Thollot, EPI Maverick), we solved this problem by extending to 3d the floating tangents algorithm introduced in 2d in [54]. In this new method, only tangents are strictly interpolated while points are displaced in an optimal way so as to lie in a feasible configuration, *i.e.*, a configuration that is compatible with the interpolation by a helix. Our approach relies upon the co-helicity condition found by Ghosh [56], which was however only partially proved in [56]. To ensure the existence of the helix and prove its uniqueness in the general case, we complete the proof which serves as the basis for our reconstruction algorithm.

Our method proves to be efficient and robust as it can successfully handle large and complex datasets from real curve aquisitions, such as the capture of hair fibers or the magnetic field of a star. We also compared our method against a standard nonlinear least-squares methods. Unlike the optimization approach which often fail to converge in the case of frizzy input curves, our method remains extremely fast regardless the complexity of the input curves. The set of these results was published this year at Computer-Aided Geometric Design [28]. This work has been transferred to L'Oréal in December 2013. Some source code is also freely released for academics under the GPLv.3 license (see Section 5.3).

6.8.2.2. Inverse modeling of a super-helix assembly subject to frictional contact

Participants: Florence Bertails-Descoubes, Alexandre Derouet-Jourdan, Gilles Daviet.

In A. Derouet-Jourdan's PhD's thesis (co-supervised by Joëlle Thollot, EPI Maverick), we bring a first solution to the challenging problem consisting in identifying the intrinsic geometry of a fiber assembly under gravity and (unknown) frictional external and mutual contacts, from a single configuration geometry (a set of geometric curves). Taking an arbitrary fiber assembly geometry (such as hair) as input together with corresponding interacting meshes (such as the body mesh), we interpret the fiber assembly shape as a static equilibrium configuration of a fiber assembly simulator, in the presence of gravity as well as fiber-mesh and fiber-fiber frictional contacts. Assuming fibers parameters are homogeneous and lie in a plausible range of physical values, we show that this large, underdetermined inverse problem can be formulated as a well-posed constrained optimization problem (second-order cone quadratic program), which can be solved robustly and efficiently by leveraging the frictional contact solver of our direct simulator for fiber assemblies [8]. Our method was successfully applied to the animation of various hair geometries, ranging from synthetic hairstyles manually designed by an artist to the most recent human hair data reconstructed from capture. These results were published this year at SIGGRAPH Asia [27].

6.9. Threshold in neural models

Participant: Arnaud Tonnelier.

We studied the threshold for spike initiation in two-dimensionnal neural models. A threshold criterion that depends on both membrane voltage and recovery (or adaptation) variable is proposed. Our approach provides a simple and unified framework that can account for adapting threshold, threshold variability, dynamic threshold, inhibition-induced spike and postinhibitory facilitation. Implications on neural modeling and on neural dynamics are discussed.

6.10. Nonsmooth modes in chains of impact oscillators

Participants: Vincent Acary, Guillaume James, Franck Pérignon.

Chains of impact oscillators arise for example as finite-element models of thin oscillating mechanical structures (a string under tension or a clamped beam) contacting rigid obstacles. Nonlinear periodic waves are observed in experiments on such systems, but relatively little is known from a theoretical point of view on their existence and stability. In 2008, Gendelman and Manevitch have analyzed the existence and stability of nonlinear localized modes (breathers) for discrete linear chains with a single node undergoing rigid impacts. In this work, we introduce a numerical method allowing to compute branches of time-periodic solutions when an arbitrary number of nodes undergo rigid impacts without energy dissipation. For this purpose, we reformulate the search of periodic solutions as a boundary value problem incorporating unilateral constraints. We illustrate this numerical approach by computing different families of breathers and nonlinear normal modes. Our method is much more effective than a numerical continuation of periodic solutions based on compliant models, which requires to integrate stiff differential equations and lead to costly numerical continuation. These results have been submitted to the ENOC 2014 conference.

6.11. Traveling waves in spatially discrete excitable media

Participants: José Eduardo Morales, Arnaud Tonnelier, Guillaume James.

The propagation of traveling waves in excitable media is a widespread phenomenon, with applications ranging from forest fires to electrical signals propagating along nerve fibers. The case of spatially discrete excitable models is notoriously difficult to analyze. In particular, for the discrete FitzHugh-Nagumo reaction-diffusion system, the existence of pulses for a general class of bistable nonlinearities has been proved only recently (Hupkes and Sandstede, 2010). The existence of pulses under more general types of interactions (e.g. elastic instead of diffusive) remains an open question, as well as traveling wave propagation in higher-dimensional systems. These problems will be tackled in the PhD thesis of J.-E. Morales (advisors A. Tonnelier and G. James), which started on November 2013. J.-E. Morales has started to analyze pulse propagation in the excitable Burridge-Knopoff model, which finds applications in the context of nonlinear friction. This model includes elastic interactions between particles, and an additional difficulty linked with nonsmoothness of the (multivalued) Coulomb friction law.

6.12. Nonlinear waves in granular chains

Participants: Guillaume James, Bernard Brogliato, Ngoc-Son Nguyen.

Granular chains made of aligned beads interacting by contact (e.g. Newton's cradle) are widely studied in the context of impact dynamics and acoustic metamaterials. When a large number of beads are present, their dynamics can be described by infinite-dimensional differential equations, which possess a limited smoothness when unilateral Hertzian contact interactions are considered. In this context, we have developed and analyzed new reduced-order models describing nonlinear wave propagation in such systems. In the work [49] (collaboration with D.Pelinovsky, McMaster Univ.), we analyze small amplitude slowly modulated compression waves in the limit when the exponent of the Hertz force is close to unity. From a multiple scale analysis, we derive a new type of Korteweg-de Vries equation with logarithmic nonlinearity allowing to approximate wave profiles, in particular solitary wave solutions.

In addition the LZB model introduced in [14] has been extensively used to numerically investigate wave phenomena in chains of aligned balls (tapered, monodisperse, anti-tapered, stepped chains). Thorough comparisons with experimental results reported in the Granular Matter literature have been made. The results are reported in the monograph [15].

6.13. Robotics

6.13.1. Lexicographic Least-Squares solver

Participants: Pierre-Brice Wieber, Dimitar Dimitrov.

We have been working on Multi-Objective Least-Squares problems with inequality constraints for the last few years, focusing especially on the Lexicographic case. A previous collaboration with LAAS-CNRS and CEA-LIST led to the development of a software, SOTH , based on Complete Orthogonal Decompositions, which has become a *de facto* reference in robotics when controlling robots (mobile, manipulator or humanoid) through constraints. The focus this year in the Bipop team has been to accelerate computations by reworking the inner matrix decomposition by combining QR and LU decompositions. The resulting solver, called LexLS, is approximately 5 times faster than the previous SOTH solver on most problems. But the main result has been to show both in theory and practice that it is faster to solve a Lexicographic problem than a Weighted problem, on the contrary to popular beliefs both in robotics and optimization theory. That leads to a reversal of popular approaches that prefer to solve weighted problems (thought to be faster to solve) as approximations to lexicographic problems (thought to be slower to solve).

6.13.2. Mobile manipulation by humanoid robots

Participants: Pierre-Brice Wieber, Dimitar Dimitrov, Alexander Sherikov, Jory Lafaye.

The realization of mobile manipulation by humanoid robots requires the handling of two simultaneous problems: taking care of the dynamic balance of the robot, what is usually done with Model Predictive Control (MPC) schemes, and redundant motion and force control of the whole body of the robot, what is usually done with a Quadratic Program, or a more advanced Lexicographic Least-Squares problem (see above). These two problems are usually solved in sequence: an MPC scheme first computes the necessary motion of the feet and Center of Mass (CoM) of the robot, then motion and force redundancy of the whole body of the robot is resolved. We have observed that this sequence corresponds to a lexicographic order between two objectives, feet and CoM motion first, the rest of the body after, which limits the possibility to tackle scenarios where we would like the motion of the CoM of the robot to be driven by the motion of the rest of the body of the robot, for example to catch an object with the hand. We have proposed therefore to reorganize the order between these different objectives, building on the LexLS solver presented above.

6.13.3. Reactive trajectory generation

Participants: Pierre-Brice Wieber, Dimitar Dimitrov, Saed Al Homsy, Matthieu Guilbert.

The goal of the ongoing collaboration with Adept Technologies is to generate near time optimal trajectories in the presence of moving obstacles in real time. Results are not public yet due to industrial constraints.

6.14. Optimization

6.14.1. Semidefinite programming and combinatorial optimization

Participant: Jérôme Malick.

We have worked with Frederic Roupin (Prof. at Paris XIII) and Nathan Krislock (Assistant Prof. at North Illinois University, USA) on the use of semidefinite programming to solve combinatorial optimization problems to optimality.

We proposed a new family of semidefinite bounds for 0-1 quadratic problems with linear or quadratic constraints [61]. We have embedded the new bounds within branch-and-bound algorithms to solve 2 standard combinatorial optimization problems to optimality.

- *Max-cut.* We developed [60] an improved bounding procedure obtained by reducing two key parameters (the target level of accuracy and the stopping tolerance of the inner Quasi-Newton engine) to zero, and iteratively adding triangle inequality cuts. We also precisely analyzed its theoretical convergence properties. We show that our method outperform the state-of-the-art solver ([62]) on the large test-problems.
- *Heaviest k-subgraph problems*. Adapting the techniques we developped for the max-cut problem, we have proposed in [59] an big improvement of the first algorithm (up to 10 times faster). For the first time, we were able to solve exactly k-cluster instances of size 160. In practice, our method works particularly fine on the most difficult instances (with a large number of vertices, small density and small k).

We have also been working on a generic online semidefinite-based solver for binary quadratic problems using the generality of [61]. Finally, a first web interface for our solvers and our data sets are available online at http://lipn.univ-paris13.fr/BiqCrunch/.

6.14.2. On computing marginal prices in electricity production

Participants: Jérôme Malick, Sofia Zaourar.

Unit-commitment optimization problems in electricity production are large-scale, nonconvex and heterogeneous, but they are decomposable by Lagrangian duality. Realistic modeling of technical production constraints makes the dual objective function computed inexactly though. An inexact version of the bundle method has been dedicated to tackle this difficulty [58]. We have worked on two projects related to solving dual unit-commitment problem by inexact bundle methods.

- *Stabilization.* We observed that the computed optimal dual variables show a noisy and unstable behaviour, that could prevent their use as price indicator. We have proposed a simple and controllable way to stabilize the dual optimal solutions, by penalizing the total variation of the prices [63]. Our illustrations on the daily electricity production optimization of EDF show a strinking stabilization at a negligible cost.
- *Acceleration*. We have worked with Welington Oliveira (IMPA, Brazil) on the acceleration of inexact bundle methods by taking advantage of cheap-to-get inexact information on the objective function which comes without any tighness guarantee though. We came up with a new family of bundle methods incorporating this coarse inexact information, to get better iterates. We have studied the convergence of these method and we have conducted numerical experimentation on unit-commitment problems and on two-stage linear problems show a subtantial gain in the overall computing time. This research is about to be released in a preprint in HAL

BONSAI Project-Team

6. New Results

6.1. High-throughtput sequence processing

- Within our collaboration with Montpellier (IRB and LIRMM) we published a paper on CRAC, a software for analysing short RNA sequences and detecting variations among them [5].
- We have been invited to contribute an invited book chapter on metatranscriptomic data analysis (*Methods in Molecular Biology*, in press). This chapter covers the complete bioinformatic analysis from raw reads to taxonomic assignation, and introduces our software SortMeRNA (see Paragraph 5.6). This is a joint work with team LABIS in Genoscope.
- Evguenia Kopylova defended her thesis on December, the 11th (*"New algorithmic and bioinformatic approaches for the analysis of data from high throughput sequencing"*, [1]). The second part of her work deals with a new read mapper for metagenomic sequence data.
- Within our collaboration with the Lille hospital, we developped a seed-based heuristics for the detection of lymphocyte rearrangements from high-throughput data. This method is implemented in the software Vidjil (see Section 5.7). Our results were presented at the Jobim conference [8], and a journal article was submitted.

6.2. RNA algorithms

- We have started a new collabarative project with Bielefeld Universität on an extension of *Algebraic Dynamic Programming*. We introduced a generic specification framework, called *inverted coupled rewrite systems* [9], that can deal with optimization problems on strings, trees, and arc-annotated sequences. It is based on the following ideas: the solutions of combinatorial optimization problems are the inverse image of a term rewrite relation that reduces problem solutions to problem inputs. A tree grammar is used to further refine the search space, and optimization objectives are specified as interpretations of these terms. All these constituents provide a mathematically precise and complete problem specification, leading to concise yet translucent specifications of dynamic programming algorithms.

6.3. Genomic rearrangements

- Within a collaboration with LIAFA (CNRS UMR 7089, and University Paris 7) we published a method for the assembling of ancestral gene orders from contiguous ancestral fragments [4].

6.4. Nonribosomal peptides

- Yoann Dufresne is starting a PhD thesis on computational biology for nonribosomal peptides (NRPs) under the supervision of Maude Pupin and Laurent Noe, after doing his master thesis with them. He already worked on the translation of the chemical structure of the NRPs into their monomeric structure. NRPs can be represented by their chemical structure that is a graph where the atoms are represented by nodes and the chemical bonds by arcs; or by their monomeric structure that is a graph where the monomers are represented by nodes and the chemical bonds between monomers by arcs. We designed a novel algorithm capable of localizing the monomers from a reference list in the chemical structures of peptides [7]. It is based on a heuristic that utilizes chemical information of NRPs. The preliminary results are encouraging, and should lead to further studies.

CAD Team

5. New Results

5.1. Geometry

5.1.1. From CAD to Engineering: Computing FEM on curved surfaces

Participants: Jean-Claude Paul, Kan-Le Shi, Yu-Shen Liu, Jin-San Cheng, Cheng-Lei Yang, Bruno Durand, Jun-Hai Yong.

In cooperation with Bruno Lévy (Inria)

The cooperation with EADS, based on our new B-Spline surface formulation, was very promising, for complex shape modelling. Our surfaces are very efficient in term of precision. Moreover, they avoid the control point explosion of NURBS surfaces. We propose our work in two directions: 1) to Improve the Modelling process for the user (it is a strategic point of the success of our new mathematical surface); 2) to take profit of the control points way of our surface to compute numerical simulation on this surface directly. In industry, Geometry design and Engineering employ a sequence of tools that are generally not well matched to each other. For example, the output of a computer aided geometric design system is typically not suitable as direct input for a finite-element modeler. This is usually addressed through intermediate tools such as mesh generators. Unfortunately, these are notoriously lacking in robustness. Even once a geometric model has been successfully meshed, the output of a finite-element simulation cannot be directly applied to the original geometric model, since there is no straightforward mapping back to the original design degrees of freedom. Additionally there is a need for a trade-off between the speed of analysis and the fidelity of the results. In the early stages of design, quick results are necessary, but approximate results are acceptable. In the later stages, highly precise results are required, and longer computation times are tolerated. Worse, different underlying models are required for each level of refinement. These difficulties make the design process cumbersome and inhibit rapid iteration over design alternatives. We plan to use FEA on Knot vectors surfaces directly (i.e. use the same function basis for the Geometric Modeling and the Numerical Simulation Process. We will apply this approach to fluids analysis: turbulence modeling (fluid-structure interaction). We think that our surface functions exhibiting higher-order continuity are an ideal candidate for approximating such flows. From the practical point of view, the main objectives of the study are to evaluate, in the scope of this application, the efficiency of such approach in term of simulation accuracy, simulation time and computational convergence. We also aim to evaluation how such approach deals with simulation accuracy/convergence according to CAD definition (quality/size of patches used to define the 3D shape).

5.1.2. From CAD to Manufacturing: Robustness tolerance and error control

Participants: Jun-Hai Yong, Yu-Shen Liu, Clara Issandou, Hai-Chuan Song, Lu Yang, Kang-Lai Qian, Jean-Claude Paul.

In cooperation with Dr. Nabil Anwer – ENS Cachan and the Tsinghua PLM Center (supported by Dassault System). Dr. Yi-Jun Yang (Shandong University), Dr. Xiao-Diao Chen (Zhejiang University)

Based on our theoretical contribution in Differential Geometry, especially about our ϵ-Geometry Continuity and our new geometric operators we proposed several elegant solutions to the most important challenges in Computer Aided Design (see Lees A Piegl. "Ten challenges in Computer-Aided-Design". Jal of CAD 2005. 37 (4): 461-470): robustness, tolerances, error control. During CAD processes one uses a myriad of tolerances, many of which are directly related to the actual manufacturing process. Some interesting questions here include: What are the most relevant machining tolerances? How to set the army of computational tolerances, e.g. those of systems of equations, to guarantee machining within the required accuracy? How tolerances in different spaces, e.g. in model space and in parameter space, are related. Numerical instabilities also account for the majority of computational errors in commercial CAD systems. The problems related to robustness haunt every programmer who has ever worked on commercial systems. Fixing numerical bugs can be very frustrating, and often times results in patching up the code simply because no solution exists to remedy the problem. We first plan for assisting the designer when specifying the functional tolerances of a single part included in a mechanism, without any required complex function analysis. The mechanism assembly is first described through a positioning table formalism. In order to create datum reference frames and to respect assembly requirements, an ISO based 3D tolerancing scheme will be proposed, thanks to a set of rules based on geometric patterns and TTRS (Technologically and Topologically Related Surfaces). Since it remains impossible to determine tolerance chains automatically, the designer must impose links between the frames. We want to develop proposes ISO based tolerance specifications to help ensure compliance with the designer's intentions, saving on time and eliminating errors.

5.2. Computer Graphics (2010-2013)

5.2.1. Inverse Procedural Modeling of Facade Layouts

Participants: Weiming Dong, Bin Wang, Dong-Ming Yan, Hua-Liang Xie, Jean-Claude Paul.

We want to address the following open research problem: How can we generate a deterministic shape grammar that explains a given facade layout? An approximate dynamic programming framework will tackle this problem. The proposed solution contributes to the compression of urban models, architectural analysis, and the generation of shape grammars for large-scale urban modeling. As a major contribution of this work we want to formulate the inverse procedural modeling problem for facade layouts as a smallest grammar problem. We also want to propose an automatic algorithm to derive a shape grammar for a given facade layout. In this work, we will assume segmented and labeled facade layouts as input and do not derive the shape grammars directly from photographs. The joint optimization of segmentation and grammar extraction remains an aspirational goal for this work.

5.2.2. Architecture Design

Participants: Jean-Claude Paul, Bin Wang, Weiming Dong, Lin Li, Yan Kong, Yong Zhang, Fan Tang, Fuzhang Wu, Cui-Gong Wang.

In cooperation with UC Berkeley - Department of Architecture

We want to propose a method for automated generation of architectural models for computer graphics applications. Our focus is not only on the building layout: the internal organization of spaces within the building, but also the Architectural composition of volumes, roofs and facades. We focus on the generation of various types of buildings: residences, schools, museums, hospitals, civic enters, office buildings. Our work builds on grammar-based procedural modeling, inverse procedural modeling and composition rules, especially symmetry and scaling, and interactivity. Moreover, we consider the architecture design process as an iterative trial-and-error process that requires significant expertise and learning by doing.

CAGIRE Team

6. New Results

6.1. DNS of a jet in crossflow: generation of a synthetic turbulent signal and coupling with characteristics based boundary conditions

The implementation of the boundary conditions for DNS of the flow configuration that consists of a jet issuing from an inclined cylindrical hole and discharging into a turbulent crossflow is investigated in the framework of our current participation in the Impact-AE EU funded program. First, a method allowing the generation of turbulent inflow that matches targeted statistics (mean velocity and Reynolds stress tensor components measured on the MAVERIC test facility) has been chosen. On the basis of a study of the main classic methods identified in the literature, it has been considered that the Synthetic Eddy Method (SEM) represents the best compromise between effectiveness and cost, from both a computation and a storage point of view. With this approach, eddy structures are created and injected at the inlet plane of the computational domain. These analytically defined structures are chosen in order to reproduce the most relevant ones present in a turbulent channel flow. The SEM implementation has been considered for (1) a basic form of SEM that does not differentiate the vortices in function of their distance to the wall, and (2) a more elaborated version of the method, denoted SEM-WB, where the inlet plane is split into different zones that accommodate different types of coherent structures according to what is observed in a turbulent boundary layer. In order to prescribe realistic turbulence statistics, the targeted mean velocity and Reynolds stress values of the SEM-WB method were obtained by performing dedicated PIV measurements on the MAVERIC test facility (UPPA). The basic form of the method gives quite satisfactory results. The values of some parameters of the SEM-WB method have still to be adjusted in order to achieve a better convergence rate towards the targeted statistics. In november 2013, the deliverable D2.211 (Confidential) documenting in details this methodology and the results obtained with the related module written in C++ has been issued by the team to the IMPACT-AE office. Assuming that the synthetic turbulent signal is generated in a satisfactory way, one is left with the set-up of the procedure necessary to incorporate this signal into a characteristics based method for handling the boundary conditions at the flow inlet(s). We have developed an approach that proved suitable, in a 1-D configuration so far, to accurately superimpose acoustics and turbulence while preserving the non reflective properties at the inlet boundary [5].

6.2. Low Mach number flows simulations issues

Our activity for developing schemes suitable for the simulation of low Mach number flows considers the two main techniques developed initially for dealing with either zero Mach number flows (pressure-velocity coupling) or compressible flows (density based approach). For the methodology adressing the pressure-velocity coupling, we concentrated on the issue of handling in a semi-implicit way the unsteady set of characteristics based equations at both the outlet and the inlet of a subsonic internal flow. The methodology employed to solve the boundary equations has been designed to mimic the pressure-velocity coupling employed in the interior of the computational domain. The numerical experiments carried out with an acoustic CFL number significantly larger than unity show that the expected reflective and non-reflective behavior is preserved at these boundaries [3].

For the density based approach [6], the Euler or Navier-Stokes equations semi-discretised with a Roe-like flux scheme are analysed using an asymptotic development in power of the Mach number. As expected, this development shows that the inaccuracy at low Mach is due to the bad scaling of the pressure gradient in the momentum equation [20]. In addition, the behaviour of any compressible solver based on that scheme proved to be highly dependent on the geometry of the mesh elements [33]. Several cures to this inaccuracy problem exist in the literature for steady flow calculations. But for unsteady low Mach flows simulations, our numerical experiments with high order discontinuous Galerkine discretisation put into evidence the bad stability properties of these modified schemes. In order to adress that second issue, a semi-discrete wave equation for the order one pressure in the system has been derived by including the acoustic time scale in the asymptotic development. An analysis of the dissipative terms of this wave equation has been started in order to determine the possible way of regaining good stability properties while ensuring a good accuracy at low Mach.

CAIRN Project-Team

6. New Results

6.1. Reconfigurable Architecture Design

6.1.1. Arithmetic Operators for Cryptography and Fault-Tolerance

Participants: Arnaud Tisserand, Emmanuel Casseau, Thomas Chabrier, Karim Bigou, Franck Bucheron, Jérémie Métairie, Nicolas Veyrat-Charvillon, Nicolas Estibals.

Arithmetic Operators for Fast and Secure Cryptography. Scalar recoding is popular to speed up ECC (elliptic curve cryptography) scalar multiplication: non-adjacent form, double-base number system, multi-base number system (MBNS). But fast recoding methods require pre-computations: multiples of base point or off-line conversion. In paper [42] presented at ARITH, we presented a multi-base (e.g. (2,3,5,7)) recoding method for ECC scalar multiplication based on i) a greedy algorithm starting least significant terms first, ii) cheap divisibility tests by multi-base elements and iii) fast exact divisions by multi-base elements. Multi-base terms are obtained on-the-fly using a special recoding unit which operates in parallel to curve-level operations and at very high speed. This ensures that all recoding steps are performed fast enough to schedule the next curve-level operations without interruptions. The proposed method can be fully implemented in hardware without pre-computations. We report FPGA implementation details and very good performance compared to state-of-art results. A specific version of our method allows random recodings of the scalar which can be used as a partial counter-measure against side-channel attacks. The PhD thesis defended by Thomas Chabrier [18] deals with MBNS and other types of arithmetic recodings for ECC scalar multiplication (title: "Arithmetic recodings for ECC cryptoprocessors with protections against side-channel attacks").

In the paper [67], presented at ComPAS, we presented efficient arithmetic operators for divisibility tests and modulo operations for large operands (e.g. 160-600 bits like in cryptographic applications) and by a set of small constants such as $(2^a, 3, 5, 7, 9)$ where $1 \le a \le 12$. These operators have been validated and implemented on FPGAs.

In the paper [39] presented at CHES, we described a new RNS modular inversion algorithm based on the extended Euclidean algorithm and the plus-minus trick. In our algorithm, comparisons over large RNS values are replaced by cheap computations modulo 4. Comparisons to an RNS version based on Fermat's little theorem were carried out. Comparisons to a version based on Fermat's little theorem were carried out. The number of elementary modular operations is significantly reduced: a factor 12 to 26 for multiplications and 6 to 21 for additions. Virtex 5 FPGAs implementations show that for a similar area, our plus-minus RNS modular inversion is 6 to 10 times faster. Other implementation results of RNS for ECC cryptosystems have been presented in [75] and [74].

ECC Processor with Protections Against SCA. A dedicated processor for elliptic curve cryptography (ECC) is under development. Functional units for arithmetic operations in GF(2^m) and GF(p) finite fields and 160-600-bit operands have been developed for FPGA implementation. Several protection methods against side channel attacks (SCA) have been studied. The use of some number systems, especially very redundant ones, allows one to change the way some computations are performed and then their effects on side channel traces. This work is done in the PAVOIS project.

Arithmetic Operators for Fault Tolerance. In the ARDyT project, we work on computation algorithms, representations of numbers and hardware implementations of arithmetic operators with integrated fault detection (and/or fault tolerance) capabilities. The target arithmetic operators are: adders, subtracters, multipliers (and variants of multiplications by constants, square, FMA, MAC), division, square-root, approximations of the elementary functions. We study two approaches: residue codes and specific bit-level coding in some redundant number systems for fault detection/tolerance integration at the arithmetic operator/unit level. FPGA prototypes are under development.

6.1.2. Reconfigurable Processor Extensions Generation

Participants: Christophe Wolinski, François Charot.

Most proposed techniques for automatic instruction sets extension usually dissociate pattern selection and instruction scheduling steps. The effects of the selection on the scheduling subsequently produced by the compiler must be predicted. This approach is suitable for specialized instructions having a one-cycle duration because the prediction will be correct in this case. However, for multi-cycle instructions, a selection that does not take scheduling into account is likely to privilege instructions which will be, *a posteriori*, less interesting than others in particular in the case where they can be executed in parallel with the processor core. The originality of our research work is to carry out specialized instructions selection and scheduling in a single optimization step. This complex problem is modeled and solved using constraint programming techniques. This approach allows the features of the extensible processor to be taken into account with a high degree of flexibility. Different architectures models can be envisioned. This can be an extensible processor tightly coupled to a hardware extension having a minimal number of internal registers used to store intermediate results, or a VLIW-oriented extension made up of several processing units working in parallel and controlled by a specialized instruction. These techniques have been implemented in the Gecos source-to-source framework.

Novel techniques addressing the interactions between code transformation (especially loops) and instruction set extension are under study. The idea is to automatically transform the original loop nests of a program (using the polyhedral model) to select specialized and vector instructions. These new instructions may use local memories located in the hardware extension and used to store intermediates data produced at a given loop iteration. Such transformations lead to patterns whose effect is to significantly reduce the pressure on the memory of the processor. An experiment realized on the matrix multiplication (extracted from PolyBench/C, the polyhedral benchmark suite) using an Xtensa extensible and configurable processor from Tensilica shows interesting speedups. Speedup of 4.3 for the transformed code compared to the initial code for matrices of size 512x512 and speedup of 8.75 (respectively 20.15) in case of an extension allowing SIMD vector operations on vector of 4 32-bit words (respectively 16 32-bit words) are observed.

6.1.3. Runtime Mapping of Hardware Accelerators on the FlexTiles 3D Self-Adaptive Heterogeneous Manycore

Participants: Olivier Sentieys, Antoine Courtay, Christophe Huriaux.

FlexTiles is a 3D stacked chip with a manycore layer and a reconfigurable layer. This heterogeneity brings a high level of flexibility in adapting the architecture to the targeted application domain for performance and energy efficiency. A virtualisation layer on top of a kernel hides the heterogeneity and the complexity of the manycore and fine-tunes the mapping of an application at runtime. The virtualisation layer provides self-adaptation capabilities by dynamically relocation of application tasks to software on the manycore or to hardware on the reconfigurable area. This self-adaptation is used to optimize load balancing, power consumption, hot spots and resilience to faulty modules. The reconfigurable technology is based on a Virtual Bit-Stream (VBS) that allows dynamic relocation of accelerators just as software based on virtual binary code allows task relocation.

We have proposed a novel approach to hardware task relocation in an FPGA-based reconfigurable fabric, allowing offline design, routing, and unfinalized placement of hardware IPs and dynamic placement of the corresponding bit-streams at run-time. Our proposal relies on a custom dual-context FPGA configuration memory organization in a shift-register manner and on a dedicated bit-stream insertion controller leading to a break-through in terms of adaptive capabilities of the reconfigurable hardware. We show that using our custom shift-register organization across the configuration memory, and under some weak constraints, can greatly reduce the overhead implied by the 1-D to 2-D mapping of the shift-register onto the logic fabric. The use of partial dynamic reconfiguration in FPGA-based systems has grown in recent years as the spectrum of applications which use this feature has increased. For these systems, it is desirable to create a series of partial bitstreams which represent tasks that can be located in multiple regions in the FPGA substrate. While the transferal of homogeneous collections of lookup-table based logic blocks from region to region has been shown to be relatively straightforward, it is more difficult to transfer partial bitstreams which contain fixed function resources, such as block RAMs and DSP blocks. To do so, we explore adding enhancements to the FPGA architecture which allow for the migration of partial bitstreams including fixed resources from region to region even if these fixed function resources are not located in the same position in the region. Our approach does not require significant, time-consuming place-and-route during the migration process. We quantify the cost of inserting additional routing resources into the FPGA architecture to allow for easy migration of heterogeneous, fixed function resources. Our experiments show that this flexibility can be added for a relatively low overhead and performance penalty. As mentioned above, the Virtual Bit-Stream (VBS) is a concept of an unfinalized, pre-routed bit-stream which could be loaded almost anywhere on a custom FPGA logic fabric. Unlike classical bit-streams, the VBS is not tied to a specific location on the circuit, hence its ”virtual” qualifier. The goal is to generate a single VBS only once for each and every possible location of the logic fabric in the FPGA in a unfinished manner: the time-consuming packing, place and route steps are done offline and only local routing is done at runtime in order to ensure fast decoding time as well as low memory overhead. The VBS concept is pending for a European patent application.

6.1.4. Power Models of Reconfigurable Architectures

Participants: Robin Bonamy, Daniel Chillet, Olivier Sentieys.

Including a reconfigurable area in complex systems-on-chip is considered as an interesting solution to reduce the area of the global system and to support high performance. But the key challenge in the context of embedded systems is currently the power budget and the designer needs some early estimations of the power consumption of its system. Power estimation for reconfigurable systems is a difficult issue since several parameters need to be taken into account to define an accurate model.

One first parameter concerns the choice of tasks to execute and their allocation in the computing resources. Indeed, several hardware implementations of an algorithm can be obtained and exploited by the operating system for a flexible allocation of tasks to optimize energy consumption. These different hardware implementations can be obtained by varying the parallelism level, which has a direct impact on area and execution time and therefore on power and energy consumption. To highlight this point, we made several evaluations of delay, area, power, and energy impacts of loop transformations using High Level Synthesis tools. Real power measurements have been made on an FPGA platform and for different task implementations to build a model of energy consumption versus execution time.

Furthermore, we also considered the opportunity of the dynamic reconfiguration, which makes possible to partially reconfigure a specific part of the circuit while the rest of the system is running. This opportunity has two main effects on power consumption. First, thanks to the area sharing ability, the global size of the device can be reduced and the static (leakage) power consumption can thus be reduced. Secondly, it is possible to delete the configuration of a part of the device which reduces the dynamic power consumption when a task is no longer used.

We analyzed the power consumption during the dynamic reconfiguration on a Virtex 5 board. Three models of the partial and dynamic reconfiguration power consumption with different complexity/accuracy tradeoffs are extracted. These models are used in design space exploration to include impact of reconfiguration on energy consumption of a complete system. We proposed a methodology for power/energy consumption modeling and estimation in the context of heterogeneous (multi)processor(s) and dynamically reconfigurable hardware systems. We developed an algorithm to explore all task mapping possibilities for a complete application (e.g. for H264 video coding) with the aim to extract one of the best solutions with respect to the designer's constraints. This algorithm is a step ahead for defining on-line power management strategies to decide which task instances must be executed to efficiently manage the available power using dynamic partial reconfiguration. All these results are presented in the Robin Bonamy's thesis [17]

6.1.5. Real-time Spatio-Temporal Task Scheduling on 3D Architecture

Participants: Quang-Hai Khuat, Quang-Hoa Le, Emmanuel Casseau, Antoine Courtay, Daniel Chillet.

One of the main advantages offered by a three-dimensional system-on-chip (3D SoC) is the reduction of wire length between different blocks of a system, thus improving circuit performance and alleviating power overheads of on-chip wiring. To fully exploit this advantage, an efficient management referring to allocate temporarily the tasks at different levels of the architecture is greatly important.

In the context of 3D SoC, we have developed several spatio-temporal scheduling algorithms for 3D MultiProcessor Reconfigurable System-on-Chip (3DMPRSoC) architectures composed of a multiprocessor layer and an embedded Field Programmable Gate Array (eFPGA) layer with dynamic reconfiguration. These two layers are interconnected vertically by through-silicon vias (TSVs) ensuring tight coupling between software tasks on processors and associated hardware accelerators on the eFPGA. Our algorithms cope with task dependencies and try to allocate communicating tasks close to each other in order to reduce direct communication cost, thus reducing global communication cost.

In the 3DMPRSoC context, our algorithms favor direct communications including: i) point-to-point communication between hardware accelerators on the eFPGA, ii) communication between software tasks through the Network-on-Chip of the multiprocessor layer, and iii) communication between software task and accelerator through TSV. When a direct communication between two tasks occurs, the data are stored in a shared memory placed onto the multiprocessor layer.

Our work in [68] takes all types of communication into consideration and proposes a scheduling and placement strategy of tasks reducing the global communication cost to 17% compared with our previous algorithm based on Pfair. In this work, the eFPGA layer of the 3DMPRSoC is supposed to contain homogeneous partial reconfiguration regions (PRR) and the size of a hardware accelerator is limited by the size of a PRR. To exceed this limitation, we analyzed the Vertex-List Structure (VLS) method for relocating hardware accelerators of various sizes anywhere onto the eFPGA if resources are available. Then, we proposed VLS-BCF algorithm [49] based on VLS that allows for reducing the overall communication cost significantly – up to 24% – compared to classical methods.

6.1.6. Ultra-Low-Power Reconfigurable Controllers

Participants: Vivek D. Tovinakere, Olivier Sentieys, Steven Derrien.

A key concern in the design of controllers in wireless sensor network (WSN) nodes is the flexibility to execute different control tasks for managing resources, sensing and communications tasks of the node. In this paper, low-power flexible controllers for WSN nodes based on reconfigurable microtasks are presented. A microtask is a digital control unit made up of an FSM and datapath. Scalable architectures for reconfigurable FSMs along with variable precision adders in datapath are proposed for flexible controllers. Power gating as a low power technique is considered for low power operation in reconfigurable microtasks by exploiting coarse grain power gating opportunities in FSMs and adders. Gate-level models are applied to analyze energy savings in logic clusters due to power gating. Power estimation results on typical benchmark microtasks show a $2\times$ to $5\times$ improvement in energy efficiency w.r.t a microcontroller at a cost of $5\times$ when compared with a microtask implemented as an ASIC with higher NRE costs [21].

6.2. Compilation and Synthesis for Reconfigurable Platform

6.2.1. Polyhedral-Based Loop Transformations for High-Level Synthesis

Participants: Steven Derrien, Antoine Morvan, Patrice Quinton, Tomofumi Yuki, Mythri Alle.

After almost two decades of research effort, there now exists a large choice of robust and mature C to hardware tools that are used as production tools by world-class chip vendor companies. Although these tools dramatically slash design time, their ability to generate efficient accelerators is still limited, and they rely on the designer to expose parallelism and to use appropriate data layout in the source program. We believe this can be overcome by tackling the problem directly at the source level, using source-to-source optimizing compilers. More precisely, our aim is to study how polyhedral-based program analysis and transformation can be used to address this problem. In the context of the PhD of Antoine Morvan, we have studied how it was possible to improve the efficiency and applicability of nested loop pipelining (also known as nested software pipelining) in C to hardware tools. Loop pipelining is a key transformation in high-level synthesis tools as it helps maximizing both computational throughput and hardware utilization.

We have first studied how polyhedral based loop transformations (such as coalescing) could be used to improve the efficiency of pipelining small trip-count inner loops [27] and implemented the transformation in the Gecos source to source toolbox. We also have proposed a technique to widen the applicability of loop pipelining to kernels exposing complex dynamic memory access patterns for which compile time dependency analysis techniques cannot be used efficiently. Our approach borrows from the notion of runtime memory disambiguation used in super scalar processors to add a data dependency hazards detection mechanism to the synthesized circuits. The approach has shown promising results and led to a presentation presented at the 50th ACM/IEEE Design Automation Conference [37]. In addition to our work on nested loop pipelining, we also investigated how to extend existing polyhedral code generation techniques to enable the synthesis of fast and area-efficient control-logic. Our approach was implemented in the Gecos framework and presented at the Field Programmable Technology international conference in late 2013 [63].

6.2.2. Compiling for Embedded Reconfigurable Multi-Core Architectures

Participants: Steven Derrien, Olivier Sentieys, Maxime Naullet, Antoine Morvan, Tomofumi Yuki, Ali Hassan El-Moussawi.

Current and future wireless communication and video standards have huge processing power requirements, which cannot be satisfied with current embedded single processor platforms. Most platforms now therefore integrate several processing core within a single chip, leading to what is known as embedded multi-core platforms. This trend will continue, and embedded system design will soon have to implement their systems on platforms comprising tens if not hundred of high performance processing cores. Examples of such architectures are the Xentium processor from by Recore or the Kahrisma processor, a radically new concept of morphable processor from Karlsruhe Institute of Technology (KIT). This evolution will pose significant design challenges, as parallel programming is notoriously difficult, even for domain experts. In the context of the FP7 European Project Alma (Architecture-oriented parallelization for high performance embedded Multi-core systems using scilAb), we are studying how to help designers programming these platforms by allowing them to start from a specification in Matlab and/or Scilab, which are widely used for prototyping image/video and wireless communication applications. Our research work in this field revolves around two topics. The first one aims at exploring how floating-point to fixed-point conversion can be performed jointly with the SIMD instruction selection stage to explore performance/accuracy trade-off in the software final implementation. The second one aims at exploring how program transformation techniques (leveraging the polyhedral model and/or based on the domain specific semantics of scilab built-in functions) can be used to enable an efficient coarse grain parallelization of the target application on such multi-core machines [30].

6.2.3. Numerical Accuracy Analysis and Optimization

Participants: Olivier Sentieys, Steven Derrien, Romuald Rocher, Pascal Scalart, Tomofumi Yuki, Aymen Chakhari, Gaël Deest.

Most of analytical methods for numerical accuracy evaluation use perturbation theory to provide the expression of the quantization noise at the output of a system. Existing analytical methods do not consider correlation between noise sources. This assumption is no longer valid when a unique datum is quantized several times. In [35], an analytical model of the correlation between quantization noises is provided. The different quantization modes are supported and the number of eliminated bits is taken into account. The expression of the power of the output quantization noise is provided when the correlation between the noise sources is considered. The proposed approach allows improving significantly the estimation of the output quantization noise power compared to the classical approach, with a slight increase of the computation time.

Trading off accuracy to the system costs is popularly addressed as the word-length optimization (WLO) problem. Owing to its NP-hard nature, this problem is solved using combinatorial heuristics. In [56], a novel approach is taken by relaxing the integer constraints on the optimization variables and obtain an alternate noise-budgeting problem. This approach uses the quantization noise power introduced into the system due to fixed-point word-lengths as optimization variables instead of using the actual integer valued fixed-point word-lengths. The noise-budgeting problem is proved to be convex in the rounding mode quantization case and can therefore be solved using analytical convex optimization solvers. An algorithm with linear time complexity is provided in order to realize the actual fixed-point word-lengths from the noise budgets obtained by solving the convex noise-budgeting problem.

An analytical approach is studied to determine accuracy of systems including unsmooth operators. An unsmooth operator represents a function which is not derivable in all its definition interval (for example the sign operator). The classical model is no longer valid since these operators introduce errors that do not respect the Widrow assumption (their values are often higher than signal power). So an approach based on the distribution of the signal and the noise was proposed. We focused on recursive structures where an error influences future decision (such as Decision Feedback Equalizer). In that case, numerical analysis method (e.g. Newton Raphson algorithm) can be used. Moreover, an upper bound of the error probability can be analytically determined [43]. We also studied the case of Turbo Coder and Decoder to determine data word-length ensuring sufficient system quality.

One of the limitation of analytical accuracy technique is that they are based on a Signal Flow Graph Representation of the system to be analyzed. This SFG model is currently built-out of a source program by flattening its whole control-flow (including full loop unrolling) which raises significant accuracy analysis issues. In 2013 we have started studying how we could bridge numerical analysis techniques with more compact polyhedral program representations to provide a more general and scalable framework.

6.2.4. Design Tools for Reconfigurable Video Coding

Participants: Emmanuel Casseau, Hervé Yviquel.

In the field of multimedia coding, standardization recommendations are always evolving. To reduce design time taking benefit of available SW and HW designs, Reconfigurable Video Coding (RVC) standard allows defining new codec algorithms. The application is represented by a network of interconnected components (so called actors) defined in a modular library and the behaviour of each actor is described in the specific RVC-CAL language. Dataflow programming, such as RVC applications, express explicit parallelism within an application. However general purpose processors cannot cope with both high performance and low power consumption requirements embedded systems have to face. We have investigated the mapping of RVC applications onto a dedicated multiprocessor platform. Actually, our goal is to propose an automated co-design flow based on the RVC framework. The designer provides the application description in the RVC-CAL language, after which the co-design flow automatically generates a network of processors that can be synthesized on FPGA platforms. The processors are based on a low complexity and configurable TTA processor (Very Long Instruction Word -style processor). The architecture model of the platform is composed of processors with their local memories, an interconnection network and shared memories. Both shared and local memories are used to limit the traditional memory bottleneck. Processors are connected together through the shared memories. The design flow is implemented around two open-source toolsets: Orcc (Open RVC-CAL Compiler: http://orcc.sourceforge.net) and TCE (TTA-based Co-design Environment: http://tce.cs.tut.fi). The inputs of the design flow are the RVC application, the platform configuration (i.e. the configuration of the TTA processors and their number), and the mapping specification (i.e. the mapping of the actors onto the processors). Orcc generates a high-level description of the processors, an intermediate representation of the software code associated to each actor, and the processor interconnection requirements. Then TCE uses these informations to generate a complete multi-processor platform design: the VHDL descriptions of the processors using a pre-existing database of hardware components and the executable binary code that will execute the actors on the processors.

This work is done in collaboration with Mickael Raulet from IETR INSA Rennes and has been implemented in the Orcc open-source compiler and with Jarmo Takala team from Tampere University of Technology (Finland) who is involved in the TCE toolset.

6.3. Interaction between Algorithms and Architectures

6.3.1. Design Methodologies for Software Defined Radios

Participants: Matthieu Gautier, Olivier Sentieys, Emmanuel Casseau, Arnaud Carer, Ganda-Stéphane Ouedraogo, Mai-Thanh Tran, Vaibhav Bhatnagar.

Software Defined Radio (SDR) is a flexible signal processing architecture with reconfiguration capabilities that can adapt itself to various air interfaces. It was first introduced by Joseph Mitola as an underlying structure for Cognitive Radio (CR). The FPGA (Field Programmable Gate Array) technology is expected to play a key role in the development of SDR platforms. FPGA-based SDR is a quite old paradigm and we are fronting this challenge while leveraging the nascent High Level Synthesis tools and languages.
Actually, our goal is to propose methods and tools for rapid implementation of new waveforms in the stringent flexibility paradigm. We proposed a novel design flow for FPGA-based SDR applications [38] [70]. This flow relies upon HLS principles and its entry point is a Domain-Specific Language (DSL) which partly handles the complexity of programming an FPGA and integrates SDR features.

6.3.2. Adaptive Precision under Performance Constraints in OFDM Wireless Receivers

Participants: Olivier Sentieys, Matthieu Gautier, Fernando Cladera [Master's Student].

To cope with rapid variations of channel parameters, wireless receivers are designed with a significant performance margin to reach a given Bit Error Rate (BER), even for worst-case channel conditions. Significant energy savings come from varying at run time processing bit-width, based on estimation of channel conditions, without compromising BER constraints. To validate the energy savings, the energy consumption of basic operators has been obtained from real measurements for different bit-widths on an FPGA and an ARM processor using soft SIMD. Results show that up to 66% of the dynamic energy consumption can be saved using this adaptive technique.

6.3.3. MIMO Systems and Cooperative Strategies for Low-Energy Wireless Networks

Participants: Olivier Berder, Olivier Sentieys, Pascal Scalart, Matthieu Gautier, Le-Quang-Vinh Tran, Duc-Long Nguyen [Master's Student], Ruifeng Zhang, Viet-Hoa Nguyen.

Since a couple of years, the CAIRN team has reached a significant expertise in multi-antenna systems, especially in linear precoding. In order to obtain an efficient, simple and general form of precoders, we considered an SNR-like matrix to approximate the minimum distance. The precoding matrix is first parameterized as the product of a diagonal power allocation matrix and an input-shaping matrix and demonstrated that the minimum diagonal entry of the latter is obtained when the input-shaping matrix is a DFT-matrix. The major advantage of this design is that the solution can be available for all rectangular QAM-modulations and for any number of datastreams [28]. On the other hand the sphere decoder was applied at the receiver side instead of maximum likelihood and the performance complexity trade-off was investigated. Some adjustments of traditional sphere decoding algorithm were mandatory to adapt to the precoded MIMO systems [55].

Another way to exploit the MIMO diversity, especially in WSN where only one antenna can be supported by limited size devices, is to use space-time codes in a distributed manner. In this context, a new protocol, called fully distributed space-time coded (FDSTC) protocol having information exchange between relays, was proposed and compared with the conventional distributed space-time coded (DSTC) protocol using non-regenerative relays (NR-relays) and regenerative relays (R-relays). At the same spectral efficiency, FDSTC has better performance in terms of outage probability in high SNR regions. In terms of energy efficiency, the FDSTC protocol is shown to outperform DSTC for long-range transmissions [32]. As very few dedicated MAC protocols exist, we investigated a novel low-latency MAC protocol (ARQ-CRI) for low-power cooperative wireless sensor networks WSNs, while preserving (in high traffic mode) or even increasing (in low traffic mode) energy-efficiency [54]. An energy efficient opportunistic MAC protocol with the mechanisms of reservation and a relay candidate coordination were also proposed, and the multi-relay transmission probability was analyzed. Simulation and experiment results on a real wireless sensor network platform in different channels demonstrated the proposed scheme greatly reduces the multi-relay transmission probability and achieves about 84% improvement of energy efficiency compared with the traditional opportunistic MAC schemes [66].

6.3.4. Energy Harvesting and Adaptive Wireless Sensor Networks

Participants: Olivier Berder, Olivier Sentieys, Arnaud Carer, Mahtab Alam, Ruifeng Zhang, Trong-Nhan Le.

As tiny sensor nodes are equipped with limited battery, the optimization of the power consumption of these devices is extremely vital. In typical WSN platforms, the radio transceiver consumes major proportion of the energy. Major concerns are therefore to decrease both the transmit power and radio activity. We designed an adaptive transmit power optimization technique that is applied under varying channel to reduce the energy per successful transmitted bit. Each node locally adapts its output power according to the signal-to-noise ratio (SNR) variations (for all the neighbor nodes). It is found that by dynamically adapting the transmit power on average can help to reduce the energy consumption by a factor of two [36].

To further extend the system lifetime of WSN, energy harvesting techniques have been considered as potential solutions for long-term operations. Instead of minimizing the consumed energy as for the case of battery-powered systems, the harvesting node is adapted to Energy Neutral Operation (ENO) to achieve a theoretically infinite lifetime. Several types of energy sources can be used, as light, motion or heat [51]. We even investigated the possibility for a single sediment-microbial fuel cell (MFC) to power a wireless sensor network [31]. Through experiments conducted on the PowWow platform, it was shown that the energy harvesting device adapts to the intermittent power supplied by the MFC, and the radio-transmitter is able to switch from a continuous to degraded mode. Given the harvesting capability, we then tried to design power managers (PM) able to optimize the quality of service of WSN while maintaining ENO. Our PM adapts the duty cycle of the node according to the estimation of harvested energy and the consumed energy provided by a simple energy monitor for a super capacitor based WSN to achieve the ENO [52]. When possible, as is sometimes the case for solar or wind energy, it is also of prime interest to benefit from an accurate energy predictor to estimate the energy that can be harvested in the near future, therefore we proposed a low complexity energy predictor using adaptive filter [53]. Finally, with colleagues from University College of Cork, we recently investigated the possibility to combine energy harvesting platforms with low power wake-up radios. A nano-watt wake-up radio receiver (WUR) was used cooperatively with the main transceiver in order to reduce the wasted energy of idle listening in asynchronous MAC protocols, while still maintaining the same reactivity [50].

6.3.5. Impact of RF Front-End Nonlinearity on WSN Communications.

Participants: Amine Didioui, Olivier Sentieys, Carolynn Bernier [CEA Leti].

In the context of a collaboration with CEA Leti, we studied the impact of RF front-end non-linearity on the performance of wireless sensor networks (WSN). More specifically, we investigated the problem of interference caused by intermodulation between in-band interferers. We analyzed this problem using an enhanced model of signal-to-interference-and-noise ratio (SINR) that includes an interference term due to intermodulation. Using a WSN simulator and the selectivity and the third-order input intercept point (IIP3) specifications of a radio transceiver, we have shown that the new SINR model provides helpful information for the analysis of intermodulation problems caused by in-band signals in IEEE 802.15.4 WSNs. In [45], we presented a reconfigurable receiver model whose purpose is to enable the study of reconfiguration strategies for future energy-aware and adaptive transceivers. This model is based on Figure of Merits of measured circuits. To account for real-life RF interference mechanisms, a link quality estimator is also provided. We show that adapting the receiver performance to the channel conditions can lead to considerable power saving. The models proposed can easily be implemented in a wireless network simulation in order to validate the value of a reconfigurable architecture in real-world deployment scenarios.

6.3.6. HarvWSNet: A Co-Simulation Framework for Energy Harvesting Wireless Sensor Networks.

Participants: Amine Didioui, Olivier Sentieys, Carolynn Bernier [CEA Leti].

Recent advances in energy harvesting (EH) technologies now allow wireless sensor networks (WSNs) to extend their lifetime by scavenging the energy available in their environment. While simulation is the most widely used method to design and evaluate network protocols for WSNs is simulation, existing network simulators are not adapted to the simulation of EH-WSNs and most of them provide only a simple linear battery model. To overcome these issues, we have proposed HarvWSNet, a co-simulation framework based on WSNet and Matlab that provides adequate tools for evaluating EH-WSN lifetime [44]. Indeed, the framework allows for the simulation of multi-node network scenarios while including a detailed description of each node's energy harvesting and management subsystem and its time-varying environmental parameters. A case study based on a temperature monitoring application has demonstrated HarvWSNet's ability to predict network lifetime while minimally penalizing simulation time [40].

6.3.7. Synchronisation Algorithms and Parallel Architecture for Wireless and High-Rate Optical OFDM Systems

Participants: Pramod Udupa, Olivier Sentieys, Arnaud Carer, Pascal Scalart.

Multi-band Coherent Optical OFDM (MB CO-OFDM) is widely predicted to be one of the technologies which will empower 100 Gigabit Ethernet (100GbE) networks. CO-OFDM uses coherent technology and advanced digital signal processing (DSP) to achieve net data rate of 10 Gbps in a single band. This strict throughput requirement puts a constraint on the kind of signal processing algorithms and architectures used for building the system. In [72], a scalable parallel architecture using radix-2^2 for IFFT was proposed. The second proposal consists of a scalable parallel timing synchronization algorithm which can support very high input rates at the receiver. MOPS count as well as area versus throughput for the synchronization algorithm are provided for the OFDM transceiver to show the improvements due to proposed architecture. Architecture exploration was performed using a leading-edge high-level synthesis (HLS) tool.

A novel low complexity parallel algorithm and its associated architecture were proposed for initial synchronization in orthogonal frequency division multiplexing (OFDM) systems. The method is hierarchical and uses auto-correlation for the first step and cross-correlation for the second step [60]. The main advantage of the proposed approach is that it reduces the computational complexity by a factor of five (80%), while achieving similar mean square error (MSE) as cross-correlation based methods. The method uses block-level parallelism for auto-correlation step, which speeds up the computation significantly. After fixed-point analysis, a parallel architecture is proposed to accelerate both coarse and fine synchronization steps. This parallel architecture is scalable and provides speed-up proportional to number of parallel blocks [59].

CALVI Project-Team

6. New Results

6.1. Software development

6.1.1. New methods in Selalib

The Selalib library has seen important developments during the year 2013 as we move towards a release in 2014. Several existing modules were improved in terms of their interfaces or implementations, while many other modules were added. Notably, we have improved our interaction with external software (Pigasus, developed by Dr. Ahmed Ratnani) capable of producing NURBS-based coordinate transformations and introduced a general elliptic PDE solver based on finite elements and arbitrary degree splines that can be used as a field solver in domains deformed by an arbitrary coordinate transformation. Preliminary results of these developments have been published [2]. In addition, we have included new abstractions to facilitate the development of parallel codes using domain decomposition methods. Modules like these have been already used in some of the multiple pre-packaged simulations also included during this period. For instance, it allows us to implement a new Vlasov-Poisson solver by the Eulerian reduced approach, with applications to four-dimensional Landau-Damping. The latest simulations also use newly developed interfaces related with the semi-lagrangian methodology, such as generic interfaces for advections and calculation of characteristics. At the end of 2013 virtually all conceivable abstractions related with the semi-lagrangian methodology have a natural place to live within the library.

Many new and classical methods and models have been cleanly incorporated into our software Selalib:

- Vlasov-Poisson solver by the Eulerian reduced approach. Application to 4D Landau-Damping.
- cartesian semi-Lagrangian 2D guiding center sequential simulation tested on periodic Kelvin Helmholtz instability
- polar semi-Lagrangian 2D guiding center sequential simulation tested on diocotron instability
- general curvilinear semi-Lagrangian 2D guiding center sequential simulation; first results, still in progress
- cartesian semi-Lagrangian 2D Vlasov-Poisson parallel simulation with high order splitting tested on Landau damping, bump on tail, two stream instability and beam
- cartesian semi-Lagrangian 2D Vlasov-Poisson sequential simulation without splitting tested on beam
- cartesian semi-Lagrangian 4D Vlasov-Poisson parallel simulation on cartesian grid with high order splitting tested on Landau-Damping
- polar semi-Lagrangian 4D drift kinetic parallel simulation tested on a simple ITG instability
- general curvilinear semi-Lagrangian 4D drift kinetic parallel simulation (in development)

6.1.2. New developments in CLAC

CLAC is a generic DG solver for hyperbolic conservation laws. It is optimized for running efficiently on GPU clusters. We have reorganized the software conception in order to accelerate the computations. A first point is to group the finite-elements into uniform zones in order to get optimized kernels for SIMD architectures. A second point is to manage efficiently the data transfers between the zone. An important last point is to consider a non blocking parallel task management. This is achieved through a coupling between the event mechanisms of OpenCL and MPI. Some ideas and results are presented in [44]. In addition to these developments, we have started to test some parallel programming approaches in order to achieve good efficiency on multicore processors. These ideas have been tested on fluid models [27] and the MHD model [47]. They are very efficient and will be incorparated into CLAC later on.

[2] A. Back, E. Chacon-Golcher, V. Grandgirard, A. Ratnani, E. Sonnendrücker, A 4D semi-Lagrangian Vlasov solver based on an arbitrary curvilinear grid in physical space, poster at Vlasovia, 25-28 November 2013, Nancy

6.2. Mathematical analysis of kinetic models

Participants: N. Besse, M. Bostan.

Contribution [13] concerns a one-dimensional version of the Vlasov equation dubbed the Vlasov-Dirac-Benney equation (in short V-D-B) where the self interacting potential is replaced by a Dirac mass. Emphasis is put on the relations between the linearized version, the full nonlinear problem and equations of fluids. In particular the connection with the so-called Benney equation leads to new stability results. Eventually the V-D-B appears to be at the cross road of several problems of mathematical physics which have as far as stability is concerned very similar properties.

The subject matter of paper [3] concerns anisotropic diffusion equations: we consider heat equations whose diffusion matrices have disparate eigenvalues. We determine first and second order approximations, we study their well-posedness and then, we establish convergence results. The analysis relies on averaging techniques, which have been used previously for studying transport equations whose advection fields have disparate components.

In [4] we perform an asymptotic analysis of general particle systems arising in collective behavior in the limit of large self-propulsion and friction forces. These asymptotics impose a fixed speed in the limit, and thus a reduction of the dynamics to a sphere in the velocity variables. The limit models are obtained by averaging with respect to the fast dynamics. We can include all typical effects in the applications: short-range repulsion, long-range attraction, and alignment. For instance, we can rigorously show that the Cucker-Smale model is reduced to the Vicsek model without noise in this asymptotic limit. Finally, a formal expansion based on the reduced dynamics allows us to treat the case of diffusion. This technique follows closely the gyroaverage method used when studying the magnetic confinement of charged particles. The main new mathematical difficulty is to deal with measure solutions in this expansion procedure.

6.2.1. Gyrokinetic approximation

Participants: E. Frénod, M. Lutz.

Considering a Hamiltonian Dynamical System describing the motion of charged particle in a Tokamak or a Stellarator, we build in [42] a change of coordinates to reduce its dimension. This change of coordinates is in fact an intricate succession of mappings that are built using Hyperbolic Partial Differential Equations, Differential Geometry, Hamiltonian Dynamical System Theory and Symplectic Geometry, Lie Transforms and a new tool which is here introduced : Partial Lie Sums.

6.3. Development of semi-Lagrangian methods

Participants: N. Crouseilles, P. Glanc, A. Hamiaz, S. Hirstoaga, M. Mehrenberger, J. Petri, E. Sonnendrücker, C. Steiner.

The development of numerical methods - here semi-Lagrangian schemes for plasma physic applications- is continued and strengthened in the context of the on-going library Selalib. We intend to improve the robustness of the numerical tools in order to be prepared for future more realistic test problems.

6.3.1. Vlasov-Poisson simulations on cartesian grids

We have developed a 1D x 1D Vlasov-Poisson solver on GPU using optimized FFT of CUDA and applied it on KEEN waves test case, which needs a fine resolution in velocity [46]. An efficiency of 100 Gflops on 4096x4096 grid is obtained while using single precision, and about 30 GFlops on a 2048x3048 grid using double precision. The approach is valid: implementation effort is reduced, because we rely on external optimizations and the speed-up is quite impressive (only 1 or 2 GFlops were obtained using CPU). We emphasize that FFT is used for the implementation but not (necessarily) for the numerical method. Classical methods like splines or arbitrary high order odd Lagrange interpolation are used, as they can be fitted in

[3] M. Bostan, Strongly anisotropic diffusion problems; asymptotic analysis, in *J. Differential Equations*, vol. 256 pp. 1043-1092 (2013)

[4] M. Bostan, J.-A. Carrillo, Asymptotic fixed speed reduced dynamics for kinetic equations in swarming, *Math. Models Methods Appl. Sci.*, Vol.23, No. 13 pp. 2353-2393 (2013)

this framework. In order to reduce mass conservation issues while using single precision, a delta-f method is validated. The limitation is here the grid size; we were not able to run the code for bigger sizes. We then developed other strategies based on non uniform grids in velocity with cubic splines and two grid strategies [5] and with the semi-Lagrangian discontinuous Galerkin (SLDG) scheme [6]. Integration of the code in Selalib with upgraded interface to deal with non uniform grids has been thought but remains to be done. Thanks to the MPI Parallelization of the Selalib code, we should be able to run the code for more interesting physical parameters, in particular, when the drive amplitude goes to zero, which leads to even more localized delta-f function in velocity.

Considering the SLDG scheme, we were able to prove a super convergence property in the case of constant linear advection [48].

6.3.2. Guiding-center based simulations on polar grids

We continue our work on polar grids, which are intermediate, between cartesian and general curvilinear grids. We have revisited a diocotron simulation previously done with the PIC method [75] by using a (classical) semi-Lagrangian approach. A detailed study of boundary conditions, energy and mass conservation as well as linear growth rates is performed and validated with teh code [33]. We then have extended the code to a first drift kinetic simulation [7] using at first the classical cubic splines method and then a new 2D conservative method, called CSL2D (conservative semi-Lagrangian 2D), based on mesh intersections and displacement of volumes [11]. For the latter method to work, we had to take care of the Jacobian and we used a delta-f method, in order to treat more easily non zero boundary conditions. We have benefitted from previous experience on the FSL2D (forward semi-Lagrangian) method. Again, the integration in Selalib is under development. Linear growth rate is here validated, by solving numerically the dispersion relation using recent results of [21].

6.3.3. Guiding-center simulations on general curvilinear grids

In order to deal with more complex geometries or to consider field aligned coordinates, we work on generalizing existing methods for curvilinear grids. Guiding center simulations have been successfully performed with the classical cubic splines method and a finite element solver for the Poisson equation developed by A. Back [32]. Further works concern integration in Selalib, in a more modular way. This should help the comparison with other methods as for example the recent CSL2D method [11] but also the previous CSL1D method [5].

6.4. Development of reduced Eulerian methods

Participants: E. Chacon Golcher, P. Helluy, L. Navoret, N. Pham.

6.4.1. Eulerian methods in the physical phase-space

Kinetic plasmas computer simulations are very intensive, because of the gyrokinetic turbulence. In some situations, it is possible to make assumptions on the shape of the distribution function that simplify the model. We obtain in this way a family of fluid or reduced models.If the distribution function has a Maxwellian shape (strong collisions), we obtain the MagnetoHydroDynamic (MHD) model. Even without collisions, the plasma may still relax to an equilibrium state over sufficiently long time scales (Landau damping effect). This indicates that the approximation of the distribution function could require fewer data while still achieving a good representation, even in the collisionless regime. In what follows we call this the “reduced model” approach. A reduced model is a model where the explicit dependence on the velocity variable is suppressed. In a more mathematical way, we consider that in some regions of the plasma, it is possible to exhibit a (preferably small) set of parameters α that allows us to describe the main properties of the plasma with a generalized

[5] M. Mehrenberger, N. Crouseilles, E. Sonnendrücker, B. Afeyan High-Order Numerical Methods for KEEN Wave Vlasov-Poisson Simulations, Poster at PPPS, 16-21 June 2013, San Francisco

[6] C. Steiner, M. Mehrenberger, A semi-Lagrangian discontinuous Galerkin scheme for Vlasov-Poisson equation, poster at Vlasovia, 25-28 November 2013, Nancy

[7] N. Crouseilles, P. Glanc, S. Hirstoaga, E. Madaule, M. Mehrenberger, J. Pétri, Semi-Lagrangian simulations on polar grids: from diocotron instability to ITG turbulence, poster at Vlasovia, 25-28 November 2013, Nancy

"Maxwellian" M . Then $f(x,v,t) = M(\alpha(x,t),v)$. In this case it is sufficient to solve for $\alpha(x,t)$. Generally, the vector α is solution of a first order hyperbolic system.

Several approaches are possible that we have started to study theoretically and numerically: waterbag approximations, velocity space transforms, etc.

It is also possible to construct in this way intermediate models between the kinetic and the fluid models by truncating the velocity expansion. The unknowns α of the problem become the coefficients of the expansion, which depend only on space and time. They obey a first order hyperbolic PDE system. And then it is possible to capitalize on the large theoretical and numerical machinery developed for such PDEs.

A first step is to develop the one-dimensional models in order to test several numerical methods. The chosen approach is the high order Discontinuous Galerkin (DG) family of methods for solving the hyperbolic system. We compare the reduced Eulerian model with semi-Lagrangian or PIC methods on classical test cases: Landau damping, two-stream instability [28].

6.4.2. Eulerian method in the Fourier transformed phase-space

An experiment made in the 60's [8] exhibits in a spectacular way the reversible nature of the Vlasov equations. When two perturbations are applied to a plasma at different times, at first the plasma seems to damp and reach an equilibrium. But the information of the perturbations is still here and "hidden" in the high frequency microscopic oscillations of the distribution function. At a later time a resonance occurs and the plasma produces an echo. The time at which the echo occurs can be computed (see Villani[9], page 74). The fine mathematical study of this phenomenon allowed C. Villani and C. Mouhot to prove their famous result on the rigorous nonlinear Landau damping[10].

More practically, this experiment and its theoretical framework show that it is interesting to represent the distribution function by an truncated expansion on an orthonormal basis of oscillating functions in the velocity variables. This representation allows a better control of the energy transfer between the low frequencies and the high frequencies in the velocity direction, and thus provides more relevant numerical methods. This kind of approach is studied for instance by Eliasson [11].

We have started to study such kind of approaches in [43]. An interesting point is that the truncated reduced model is also an hyperbolic system in the space direction only. This allows the classical methods for hyperbolic systems to be reused.

6.5. Two-Scale numerical methods

Participant: E. Frénod.

In note [39] a classification of Homogenization-Based Numerical Methods and (in particular) of Numerical Methods that are based on the Two-Scale Convergence is done. In this classification stand: Direct Homogenization-Based Numerical Methods, H-Measure-Based Numerical Methods, Two-Scale Numerical Methods and TSAPS (Two-Scale Asymptotic Preserving Schemes).

In [34] we develop and we explain the two-scale convergence in the covariant formalism, i.e. using differential forms on a Riemannian manifold. For that purpose, we consider two manifolds M and Y, the first one contains the positions and the second one the oscillations. We establish some convergence results working on geodesics on a manifold. Then, we apply this framework on examples.

6.6. Spline Discrete Differential Forms and applications

Participant: E. Sonnendrücker.

[8] Malmberg, J. and Wharton, C. Collisionless damping of electrostatic plasma waves. Phys. Rev. Lett. 13, 6 (1964), 184–186.
[9] Villani, C. Landau damping. CEMRACS 2010 lectures.
[10] Mouhot, C. ; Villani, C. On Landau damping, Acta Mathematica 207 (September 2011), 29-201.
[11] Eliasson, B. Outflow boundary conditions for the Fourier transformed one-dimensional Vlasov-Poisson system. J. Sci. Comput. 16 (2001), no. 1, 1–28.

In [36] we construct a new set of discrete differential forms based on B-splines of arbitrary degree as well as an associated Hodge operator. The theory is first developed in 1D and then extended to multi-dimension using tensor products. We link our discrete differential forms with the theory of chains and cochains. The spline discrete differential forms are then applied to the numerical solution of Maxwell's equations.

The notion of B-spline based discrete differential forms is recalled and along with a Finite Element Hodge operator, it is used in [35] to design new numerical methods for solving the Vlasov-Poisson equations.

6.7. Simulations of highly oscillatory Vlasov-Poisson system

Participants: E. Frénod, S. Hirstoaga, M. Lutz, E. Sonnendrücker.

In paper [45] a Lie Transform method is applied for a charged beam under the action of a radial external electric field. The aim of the Lie transform method that is used here is to construct a change of variable which transforms the 2D kinetic problem into a 1D problem. This reduces the dimensionality of the problem and make it easier to solve numerically.

In paper [41], in the framework of a Particle-In-Cell scheme for some 1D Vlasov-Poisson system depending on a small parameter, we propose a time-stepping method which is numerically uniformly accurate when the parameter goes to zero. Based on an exponential time differencing approach, the scheme is able to use large time steps with respect to the typical size of the fast oscillations of the solution.

6.8. Waterbag models: analysis and simulations

Participant: N. Besse.

Ion temperature gradient instabilities play a major role in the understanding of anomalous transport in core fusion plasmas. In the considered cylindrical geometry, ion dynamics is described in [20] using a drift-kinetic multi-water-bag model for the parallel velocity dependency of the ion distribution function. In a first stage, global linear stability analysis is performed. From the obtained normal modes, parametric dependencies of the main spectral characteristics of the instability are then examined. Comparison of the multi-water-bag results with a reference continuous Maxwellian case allows us to evaluate the effects of discrete parallel velocity sampling induced by the Multi-Water-Bag model. Differences between the global model and local models considered in previous works are discussed. Using results from linear, quasilinear, and nonlinear numerical simulations, an analysis of the first stage saturation dynamics of the instability is proposed, where the divergence between the three models is examined.

In paper [21] we present two new codes devoted to the study of ion temperature gradient (ITG) driven plasma turbulence in cylindrical geometry using a drift-kinetic multi-water- bag model for ion dynamics. Both codes were developed to complement the Runge–Kutta semi-lagrangian multi-water-bag code GMWB3D-SLC described in [1]. The CYLGYR code is an eigenvalue solver performing linear stability analysis from given mean radial profiles. It features three resolution schemes and three parallel velocity response models (fluid, multi-water-bag, continuous Maxwellian). The QUALIMUWABA quasilinear code is an initial value code allowing the study of zonal flow influence on drift-waves dynamics. Cross-validation test performed between the three codes show good agreement on both temporal and spatial characteristics of unstable modes in the linear growth phase.

6.9. Full wave modeling of lower hybrid current drive in tokamaks

Participants: Takashi Hattori, Simon Labrunie, Jean R. Roche.

This work is performed in collaboration with Yves Peysson (DRFC, CEA Cadarrache). Since September 2012 this work is included in the ANR CHROME.

The aim of this project is to develop a finite element numerical method for the full-wave simulation of electromagnetic wave propagation in plasma. Full-wave calculations of the LH wave propagation is a challenging issue because of the short wave length with respect to the machine size. In the continuation of the works led in cylindrical geometry , a full toroidal description for an arbitrary poloidal cross-section of the plasma has been developed.

Since its wavelength λ at the LH frequency is very small as compared to the machine size R, a conventional full wave description represents a considerable numerical effort.Therefore, the problem is addressed by an appropriate mathematical finite element technique, which incorporates naturally parallel processing capabilities. It is based on a mixed augmented variational (weak) formulation taking account of the divergence constraint and essential boundary conditions, which provides an original and efficient scheme to describe in a global manner both propagation and absorption of electromagnetic waves in plasmas.

With such a description, usual limitations of the conventional ray tracing related to the approximation $\lambda \ll \phi_B \ll R$, where ϕ_B is the size of the beam transverse to the rf power flow direction, may be overcome. Since conditions are corresponding to $\lambda \ll \phi_B \sim R$, the code under development may be considered as a WKB full wave, dielectric properties being local.

This formulation provides a natural implementation for parallel processing, a particularly important aspect when simulations for plasmas of large size must be considered.

The domain considered is as near as possible of the cavity fill by a tokomak plasma. Toroidal coordinates are introduced. In our approach we consider Fourier decomposition in the angular coordinate to obtain stationary Maxwell equations in a cross-section of the tokamak cavity.

A finite element method is proposed for the simulation of time-harmonic electromagnetic waves in a plasma, which is an anisotropic medium. The approach chosen here is sometimes referred to as *full-wave modeling* in the literature: the original Maxwell's equations are used to obtain a second order equation for the time-harmonic electric field. These are written in a weak form using a augmented variational formulation (AVF), which takes into account the divergence. The variational formulation is then discretized using modified Taylor-Hood (nodal) elements.

The analyze of the model considered, existence and unicity of solution, equivalence of the formulation for the domain decomposition formulation was completed in the frame of Takashi Hattori Phd thesis.

During 2013 we continue to develop the domain decomposition method introduced in 2012 and a new preconditioned system was considered in the code "FullWaveFEM", [31].

6.10. Eulerian simulations of parallel transport in the SOL

Participants: S. Hirstoaga, G. Manfredi.

During the year 2013, we have progressed in the implementation of an asymptotic preserving (AP) Eulerian Vlasov code (`VESPA: Vlasov Eulerian Simulator of PArallel transport`) for the study of parallel transport in the scrape-off layer of tokamaks. An AP Vlasov-Poisson code had already been partially developed for the quasi-neutral regime. In this case the small parameter is the Debye length λ (normalized to the parallel connection length). The Poisson equation becomes singular when $\lambda \to 0$: the AP solution consists in reformulating Poisson's equation in a way that is no longer singular in this limit. In theory, any value of λ can be used, including $\lambda = 0$, without numerical instability and without any constraint on the grid spacing and time-step. In practice, we have observed a CFL stability condition (although not very restrictive) and a limit on the smallness of λ. During the past year, we have performed systematic tests on the code, which is now capable of attaining very small values of λ, down to 10^{-4} or even lower. Meaningful results can be obtained with just 1-2 hours of computation on a standard desktop computer (see for example [29]).

The next upgrade of the VESPA code concerns the modelling of collisions, which have been implemented through a relaxation (BGK) term that also retains the possibility to include ionization and recombination in the model. The BGK term has been tested and validated against analytical results. In particular, the AP scheme had to be modified in order to correctly treat the BGK term. These upgrades are now fully integrated into the VESPA code.

Using the VESPA code, we have studied the dynamical response of a stationary sheath-presheath system to an external perturbation, which takes the form of a small density disturbance in the central region of the plasma, far from the sheaths. The numerical results suggest that, for most regimes of physical interest, the perturbation is damped away before it reaches the wall and does not have a significant impact on the structure of the sheath. This scenario has been studied for different temperatures and density profiles of the disturbance.

We have started to look at the impact of secondary electrons (SE) on the structure and the formation of the sheath. SEs were neglected in previous versions of the code but can play a significant role on the wall potential. In the VESPA code, they are now modelled as a Maxwellian electron source located near the wall. First results indicate that a large yield rate of SEs reduces the potential drop between the plasma bulk and the wall.

6.11. Other application domains

6.11.1. Applications of Two-Scale numerical methods

Participant: E. Frénod.

In paper [37] we consider a model for short term dynamics of dunes in tidal area. We construct a Two-Scale Numerical Method based on the fact that the solution of the equation which has oscillations Two-Scale converges to the solution of a well-posed problem. This numerical method uses on Fourier series.

In [18] we present Chapman–Enskog and Hilbert expansions applied to the $O(v/c)$ Boltzmann equation for the radiative transfer of neutrinos in core-collapse supernovae. Based on the Legendre expansion of the scattering kernel for the collision integral truncated after the second term, we derive the diffusion limit for the Boltzmann equation by truncation of Chapman–Enskog or Hilbert expansions with reaction and collision scaling. We also give asymptotically sharp results obtained by the use of an additional time scaling. The diffusion limit determines the diffusion source in the *Isotropic Diffusion Source Approximation (IDSA)* of Boltzmann's equation for which the free streaming limit and the reaction limit serve as limiters. Here, we derive the reaction limit as well as the free streaming limit by truncation of Chapman–Enskog or Hilbert expansions using reaction and collision scaling as well as time scaling, respectively. Finally, we motivate why limiters are a good choice for the definition of the source term in the IDSA.

6.11.2. Inverse problem governed by Maxwell equations

Participant: Jean R. Roche.

This work is performed in collaboration with José Herskovits Norman of UFRJ, Rio de Janeiro, Antonio André Novotny from the LNCC, Petropolis, both from Brazil and Alfredo Canelas from the University of the Republic, Montevideo, Uruguay.

The industrial technique of electromagnetic casting allows for contactless heating, shaping and controlling of chemical aggressive, hot melts. The main advantage over the conventional crucible shape forming is that the liquid metal does not come into contact with the crucible wall, so there is no danger of contamination. This is very important in the preparation of very pure specimens in metallurgical experiments, as even small traces of impurities, such as carbon and sulphur, can affect the physical properties of the sample. Industrial applications are, for example, electromagnetic shaping of aluminum ingots using soft-contact confinement of the liquid metal, electromagnetic shaping of components of aeronautical engines made of superalloy materials (Ni,Ti, etc.), control of the structure solidification.

The electromagnetic casting is based on the repulsive forces that an electromagnetic field produces on the surface of a mass of liquid metal. In the presence of an induced electromagnetic field, the liquid metal changes its shape until an equilibrium relation between the electromagnetic pressure and the surface tension is satisfied. The direct problem in electromagnetic casting consists in determining the equilibrium shape of the liquid metal. In general, this problem can be solved either directly studying the equilibrium equation defined on the surface of the liquid metal, or minimizing an appropriate energy functional. The main advantage of this last method is that the resulting shapes are mechanically stable.

The inverse problem consists in determining the electric currents and the induced exterior field for which the liquid metal takes on a given desired shape. This is a very important problem that one needs to solve in order to define a process of electromagnetic liquid metal forming.

In a previous work we studied the inverse electromagnetic casting problem considering the case where the inductors are made of single solid-core wires with a negligible area of the cross-section. In a second paper we considered the more realistic case where each inductor is a set of bundled insulated strands. In both cases the number of inductors was fixed in advance, see [61]. In order to look for configurations of inductors considering different topologies we introduce a new formulation for the inverse problem using a shape functional based on the Kohn-Vogelius criterion. A topology optimization procedure is defined by means of topological derivatives, a new method that simplifies computation issues was considered, see [60] and [49].

During 2013 we rewrite the inverse electromagnetic casting model in order to have a quadratic programming problem, this simplified the numerical solution and simulation [19].

CAMUS Team

6. New Results

6.1. VMAD and APOLLO

The goal of the APOLLO project is to provide a set of annotations (pragmas) that the user can insert in the source code to perform advanced analyses and optimizations, for example dynamic speculative parallelization. It is based on the prototype VMAD developed previously by the team between 2009 and 2012.

APOLLO includes a modified LLVM compiler and a runtime system. The program binary files are first generated by our compiler to include necessary data, instrumentation instructions, parallel code skeletons, and callbacks to the runtime system which is implemented as a dynamic library. External modules associated to specific analyses and transformations are dynamically loaded when required at runtime.

APOLLO uses sampling and multi-versioning to limit the runtime overhead (profiling, analysis, and code generation). At runtime, targeted codes are launched by successive chunks that can be either original, instrumented or optimized/parallelized versions. After each chunk execution, decisions can be taken relatively to the current optimization strategy. APOLLO is handling advanced memory access profiling [26], [17] through linear interpolation of the addresses, dynamic dependence analysis [18], version selection [26] and speculative polyhedral parallelization [22], [17].

Alexandra Jimborean defended her PhD thesis on this topic in 2012 [25].

In 2012, Aravind Sukumaran-Rajam started his PhD in our team to extend this work in order to handle more general programs which do not exhibit a pure polyhedral memory behavior. The investigated approach will explore approximative modelling of dependences still allowing advanced optimizing transformations of loop nests. A main issue concerns speculation verification when using approximative modelling.

Juan Manuel Martinez started his PhD in our team in 2013, with the goal of improving the flexibility of the parallel code generation phase inside Apollo. Indeed, although code skeletons are a good solution to fast dynamic parallel code generation, their shapes limit the kind of optimizing transformations that may be applied at runtime. Juan Manuel's work consists in defining elementary code skeletons that may be assembled at runtime to form a large panel of possible codes. These elementary skeletons will be defined as the objects forming the Apollo specific intermediate representation. Juan Manuel Martinez is a former master student of the University of Buenos Aires, Argentina (associate team EA-Ancome), and has already been working on VMAD to make the code generation support tiling. He defended his master thesis on this subject in October 2013 at the University of Buenos Aires.

Jean-Fran cois Dollinger will extend the framework to handle heterogeneous architectures (GPGPUs) in 2014. Willy Wolff, a master student from the University of Strasbourg, joined the APOLLO group in September 2013. His work is to implement just-in-time compilation in the APOLLO framework.

6.2. The Multifor programming construct

We have proposed a new programming control structure called "multifor", allowing to take advantage of optimization and parallelization opportunities that are not easily attainable using the standard programming structures.

In a multifor-loop, several loops whose bodies are run in interleaved fashion can be defined. Respective iteration domains are mapped onto each other according to a run frequency – the grain – and a relative position – the offset. Imen Fassi developped a source-to-source compiler called IBB (Iterate-But-Better) which is automatically translating any C source code containing multifor-loops into an equivalent source code where multifor-loops have been transformed into equivalent for-loops. Traditional polyhedral software tools, and particularly CLooG [21], are used to generate the corresponding code. Additionally, a promising perspective related to non-linear mapping of iteration spaces has also been developed, yielding to run a loop nest inside any other one by solving the problem of inverting "ranking Ehrhart polynomials".

This work is the PhD work of Imen Fassi, who started her work in 2013 and who is co-advised by Yosr Slama, Assistant Professor at the University El Manar in Tunis, Tunisia, and Philippe Clauss. A first paper [15] on this topic has been published at the IMPACT workshop that was held in conjunction with the HIPEAC conference in Berlin, Germany, in January 2013. Another paper describing the IBB compiler and showing the efficiency of multifor codes has been submitted to an international conference.

Obviously, reasoning on such a syntactic sugar suppose an associated precise and unambiguous meaning. Therefore a denotational semantics has been defined that resolves all such semantic issues and that is well-suited to prove code transformations. It has been presented to the French community of Compilation during the sixth meeting in Annecy [8].

6.3. CPU+GPU adaptive computation

In this work, we aim to automatically use CPU and GPU to jointly execute a parallel code. To ensure load balance between different PUs, thus to preserve performance, it is necessary to consider the underlying hardware and the program parameters. Compiler optimizations, execution context, hardware availability and specification make it difficult to determine execution times statically. To overcome this hurdle we rely on a portable and automatic method for predicting execution times of statically generated codes on multicore CPUs and on CUDA GPUs. This approach relies on three stages: automatic code generation, offline profiling and online prediction.

This is the latest result of PhD student Jean-François Dollinger, advised by Vincent Loechner since 2011. Preliminary results, a "fastest-wins" algorithm between a multicore CPU and the best predicted GPU code version, was published in 2013 in ICPP [14]. We are currently writing a conference paper presenting the latest advances, and preparing a journal paper to be submitted in 2014, before Jean-François Dollinger's PhD defense by the end of the year.

6.4. Minimizing the synchronization overhead of X10 programs

The CAMUS team has for long focused on compiling, optimizing, and parallelizing *sequential* programs. The project described in this section is somewhat unusual in this context, in that it targets programs written in an explicitly parallel language, and applies polyhedral modeling techniques to reschedule computations, effectively introducing parallel-to-parallel program transformations. This work has been done in collaboration with the Inria COMPSYS team at ENS Lyon, and first results will be presented at the *Compiler Construction* conference (CC'14) in April 2014.

The need to leverage the computing power of multi-core processors (and distributed computers) has lead to the design of explicitly parallel programming languages. Such languages often employ a fork/join model, and include syntax to launch and synchronize tasks (also called activities) with well-defined semantics. This brings parallel constructions under the control of the compiler, and introduces new optimization opportunities. Our work has focused on the various synchronization primitives available to the programmer, and more specifically on how one type of synchronization can be replaced with another for specific classes of programs, the goal being to minimize the synchronization overhead. We have demonstrated significant speedups on programs written using the X10 programming language, and have obtained similar results on equivalent Habanero-Java programs.

More specifically, our proposed optimization works by eliminating the use of clocks in X10 programs whose activities can be characterized with a polyhedral time-domain. The X10 language basically has two activity synchronization primitives: one is the explicit use of “clocks” (synchronization barriers) during activity execution, the other is the implicit use of activity containers that synchronize only on the end of activities. Under reasonable conditions on the patterns of activity creation and control, we have shown that long-running activities using clocks can be replaced by short-lived activities synchronized only on the end of their containers, and that this transformation provides a significant gain at run time. This work has two main contributions. First, it extends a known transformation framework to the case where the original program is already parallel.

[8] http://compilation.gforge.inria.fr/2013_04_Annecy

Second, it shows that the polyhedral model has applications far beyond its current use in data dependence and memory locality analyzes. This work also opens up new research directions. First, it turns out that our transformation is far more general than the use we currently make of it, and therefore that it provides a solid basis for other optimizations of parallel programs. Second, the polyhedral model we have developed provides an immediate cost model for synchronization primitives, which is not used in our current work, but may provide sound heuristics to adapt the optimization phase to the characteristics of specific run time components. We plan to explore these aspects in the near future.

This work has been done in collaboration with Paul Feautrier, member of the COMPSYS Inria team, in ENS Lyon. The CAMUS team has invited Paul Featrier for one week in June 2013 in Strasbourg. We are currently seeking funding to organize more frequent stays at either Lyon or Strasbourg.

This work has been invited for presentation at the LCPC workshop held in Lyon in July 2013 (http://labexcompilation.ens-lyon.fr/cpc2013). An extended version of this work has been accepted for publication at the *Compiler Construction* conference, to be held in April 2014.

6.5. Switcheable scheduling

Parallel applications used to be executed alone until their termination on partitions of supercomputers. The recent shift to multicore architectures for desktop and embedded systems is raising the problem of the coexistence of several parallel programs. Operating systems already take into account the *affinity* mechanism to ensure a thread will run only onto a subset of available processors (e.g., to reuse data remaining in the cache since its previous execution). But this is not enough, as demonstrated by the large performance gaps between executions of a given parallel program on desktop computers running several processes. To support many parallel applications, advances must be made on the system side (scheduling policies, runtimes, memory management...). However, automatic optimization and parallelization can play a significant role by generating programs with dynamic-auto-tuning capabilities to adapt themselves to the complete execution context, including the system load.

Our approach is to design at compile-time programs that can adapt at run-time to the execution context. The originality of our solution is to rely on *switcheable scheduling*, a selected set of program restructuring which allows to swap between program versions at some meeting points without backtracking. A first step selects pertinent versions according to their performance behavior on some execution contexts. The second step builds the auto-adaptive program with the various versions. Then at runtime the program selects the best version by a low overhead sampling and profiling of the versions, ensuring every computation is useful.

This work is an addition to the research directions of CAMUS related to dynamic optimization. It has been started at Paris-Sud University by Cédric Bastoul before he joined CAMUS during this year. This is an ongoing work with the PhD student Lénaïc Bagnères (GRAND-LARGE Team at Inria Saclay-Île-de-France, co-advised by Christine Eisenbeis and Cédric Bastoul). The first results have been presented in 2013 at the HiPEAC Computing System Week [9] and at the Rencontres Françaises de Compilation [10].

6.6. Interactive Code Restructuring

This work falls within the exploration and development of semi-automatic programs optimization techniques. It consists in designing and evaluating new visualization and interaction techniques for code restructuring, by defining and taking advantage of visual representations of the underlying mathematical model. The main goal is to assist programmers during program optimization tasks in a safe and efficient way, even if they neither have expertise into code restructuring nor knowledge of the underlying theories. This project is an important step for the efficient use and wider acceptance of semi-automatic optimization techniques, which are still tedious to use and incomprehensible for most programmers. More generally, this research is also investigating new presentation and manipulation techniques for code, algorithms and programs, which could lead to many practical applications: collaboration, tracking and verification of changes, visual search in large amount of code, teaching, etc.

[9] http://www.hipeac.net/thematic-session/let-us-push-thread-level-speculation
[10] http://compil13.cri.mines-paristech.fr

This is a rather new research direction which strengthen CAMUS's static parallelization and optimization issue. It has been initiated at Paris-Sud University as a collaboration between Compilation, represented by Cédric Bastoul before he joined CAMUS during this year, and Human-Machine Interaction, represented by Stéphane Huot from the IN-SITU Team at Inria Saclay-Île-de-France. This work is essentially the PhD topic of Alexander Zinenko (IN-SITU Team at Inria Saclay-Île-de-France, co-advised by Stéphane Huot and Cédric Bastoul, CORDI Grant) which started in 2013.

CARAMEL Project-Team

6. New Results

6.1. Computation of Discrete Logarithms in $\mathrm{GF}(2^{809})$

Participants: Razvan Barbulescu, Cyril Bouvier, Jérémie Detrey, Pierrick Gaudry, Hamza Jeljeli, Emmanuel Thomé [contact], Marion Videau, Paul Zimmermann.

In the context of the CATREL ANR project, most team members contributed to the achievement of a new record computation for discrete logarithms in $\mathrm{GF}(2^{809})$, with the Function Field Sieve (FFS) algorithm. This is, to date, the largest computation in a binary field of prime extension degree. Beyond the experimental data and the improvements related to "what it takes" to beat such a record, this work provides very useful basis information towards the assessment of the cut-off with the novel quasi-polynomial algorithm discussed below.

This work has been reported in the article [15], accepted for publication in the conference PKC 2014 (Public Key Cryptography). It was the occasion to illustrate several contributions of members of the teams to various phases of the algorithm: Răzvan Bărbulescu [21] analyzed the polynomial selection step for FFS; Jérémie Detrey, Pierrick Gaudry and Marion Videau [17] improved the practical implementation of the relation collection; Cyril Bouvier [23] studied the filtering step; and Hamza Jeljeli [28] proposed to use the Residue Number System representation for the linear algebra step on GPU and CPU.

6.2. A Quasi-polynomial Algorithm for the Computation of Discrete Logarithms in Finite Fields of Small Characteristic

Participants: Razvan Barbulescu, Pierrick Gaudry, Emmanuel Thomé [contact].

In collaboration with Antoine Joux (Université Pierre et Marie Curie), Răzvan Bărbulescu, Pierrick Gaudry, and Emmanuel Thomé designed a new algorithm of quasi-polynomial complexity for computing discrete logarithms in finite fields $\mathrm{GF}(p^n)$, under the constraint that the characteristic p is small: it must not grow faster than a polynomial in the input size $n \log p$. This constraint accomodates for instance the cryptographically relevant case of finite fields of fixed characteristic $\mathrm{GF}(2^n)$ and $\mathrm{GF}(3^n)$.

This new algorithm dramatically changes the complexity landscape of the computation of discrete logarithms in finite fields. This has in particular an immense impact on the small characteristic pairing-based cryptography proposals. As it turns out, the field of definition of the Weil pairing for curves over small characteristic fields lends itself incredibly well to the new algorithm, to the point that the key sizes which are necessary to claim a sufficient security suddenly become unacceptably large. The newly proposed algorithm practically kills such cryptosystems.

This work has been published in preprint form in June 2013 [22] and was immediately acclaimed as a breakthrough, receiving also some external publicity. Pending the submission outcome, a first publication is expected in 2014.

6.3. Computation of CM Class Polynomials for Genus 2 Jacobians

Participant: Emmanuel Thomé [contact].

In collaboration with Andreas Enge, Emmanuel Thomé has developed software for computing class polynomials, in the context of complex multiplication theory in genus 2. The current computations set new records which are well above the previous state of the art, as Igusa class polynomials for class number above 20,000 have been computed in december 2013 using this software. An article describing this work has been accepted for publication in *Experimental Mathematics* [11].

Using similar underlying tools and theory, and based on work by Sorina Ionica [13], Sorina Ionica and Emmanuel Thomé have worked on the analysis of isogeny graphs in genus 2, when certain properties of the endomorphism ring are satisfied. A publication is being worked on, and is expected to be submitted in early 2014.

6.4. Binary to Decimal Conversion

Participants: Cyril Bouvier, Paul Zimmermann.

Cyril Bouvier and Paul Zimmermann designed a new algorithm to convert a large binary integer to decimal (or more generally any non-power-of-two radix). Compared to the reference implementation in GNU MP, this algorithm replaces divisions by multiplications, and exhibits a speedup of up to a factor of two (or more) in some cases [24].

6.5. Fast Change of Ordering for Gröbner Bases

Participant: Pierrick Gaudry.

When solving polynomial systems, the usual approach is to compute a Gröbner basis for a monomial order that is compatible with the degree with the F4 or F5 algorithm, and then compute a Gröbner basis for the lexicographical order using the FGLM algorithm. In collaboration with Jean-Charles Faugère, Louise Huot and Guénaël Renault, Pierrick Gaudry designed another approach [27] for this second step, leading to a better asymptotic complexity: the cubic complexity is replaced by the complexity of the linear algebra where the exponent can theoretically be as small as 2.37.

CARMEN Team

6. New Results

6.1. Mathematical models

- *Mathematical derivation of a bilayer surface model of the atria using asymptotic analysis methods [28], [16]*

  We derived rigorously, by using asymptotic analysis tools, a bilayer model of atrial electrophysiology. Starting with a 3D model of atrial tissue that includes two layers with distinct electrophysiological characteristics and with an aspect ratio of ϵ, we obtained an asymptotic equivalent model when $\epsilon \longrightarrow 0$ made up of two surface models coupled by a coupling term. The bilayer model discribes the evolution of the mean in the thickness of the 3D potential in each layer. This approach is an improvement of the classical surface model of cardiac electrophysiology, because it guaranties a higher convergence speed, and allows to take into account transmural heterogeneities. We numerically implemented the 3D and bilayer models and compared it to the classical surface model. We observed a second order accuracy of the bilayer model and drastically reduced computational times respectively to the 3D model.

- *Formal derivation of a macroscopic model of propagation that includes the non linear behavior of gap junctions [16]*

  A macroscopic model of electrical propagation that take into account the non linear conductivity of the gap junction is obtained by a formal homogenization method. We derived a one dimensional macroscopic model which diffusive tensor varies in time. We compared this macroscopic model with a cell-to-cell propagation. This is an very important improvement of existing models that only consider a linear cell-to-cell coupling. The introduction of this non linear phenomenon in homogenized models gives a simulation tool to investigate the impact of the microscopic nonlinear mechanisms on the macroscopic propagation.

- *Influence of periodic diffusive inclusions on the bidomain model [29]*

  We present a new mathematical model of the electric activity of the heart. In the standard bidomain model we can distinguish the intra- and the extracellular space with different conductivities for excitable cells and the fibrotic tissue around them. The main drawback is that it assumes the existence of excitable cells everywhere in the heart, while it is known that there exist non small regions where fibroblasts take place. The fibroblasts are equally distributed and since they are non excitable cells, they can be considered as a diffusive part. Hence we extend the standard bidomain model as follows: we assume that we have periodic alternation of the healthy tissue (linear bidomain model) and fibrotic extracellular space (diffusive part). We use homogenisation techniques to derive our macroscopic partial differential equations. Interestingly, we obtain again a bidomain type model with modified conductivities that involve the volume fraction of the diffusive domain. Preliminary numerical experiments will conclude on the influence of these diffusive inclusions.

6.2. Construction of numerical models

- *Implementation of an accurate bilayer model of human atria, including realistic geometry and qualitative fibre direction [19], [21], [16]*

  We introduce a bilayer model of the human atria. We set a specific mathematical model based on two surface monodomain problems coupled by a coupling term. We recalled convergence results of the bilayer model towards a 3D model for thin tissues, we formalized an optimization method to set the coupling coefficient and we present two different asymptotically equivalent numerical implementation of the model. We then present a geometrically and electrophysiologically accurate model of the atrial heterogeneities, including two layers of fibre directions and ionic function heterogeneities based on histological and modelling works. We assess the physiological relevance of the model during a sinus wave and we check the occurrence of three-dimensional electrical behaviour such as slight electrical dissociation. This bilayer model is able to take into account transmural heterogeneities only accessible since then with full 3D models, while keeping the low computational load associated with surface models. It is then a light and relevant tool for long-lasting simulations designed to investigate atrial arrhythmia.

- *Personalization method of the bilayer model to registrate the geometry to a patient dependant geometry [16]*

  If the generic atrial bilayer model developed in [21] allows to conduct general experiments, greater customization of the model is necessary to carry out more specific studies on a given patient. We present a methodology to obtain a patient-dependant model containing the geometry of the patient, a generic fibrous organization and an image of the patient's fibrosis obtained by late-enhancement MRI. This is a common work with the clinical team of the CHU du Haut-Lévêque (H. Cochet and P. Jaïs) and the Asclepios Inria team.

  The methodology is based on a registration method developed by Durrleman et al. [34] that allows to register surfaces : the generic model is registrated towards a patient-specific geometry (work by M. Sermseant and R. Cabrera-Lozoya, Asclepios Team). The fibre organisation is transported by the same linear local transformations. The late-enhancement is projected on the model to take into account the complex patient-specific fibrotic repartition. A similar methodology was presented by McDowell et al. [37]. However, the authors took as a starting point a three-dimensional geometry and a different methodology to registrate the geometry. The work presented here is therefore innovative.

- *Faster solvers for cardiac electro-physiology problems [27]*

  There are many applications in cardiac electro-physiology where computational time is the main requirement to fulfill, even by sacrificing accuracy. Some techniques were investigated in this direction, in order to obtain a break-even point between accuracy and speed. The complete problem involves solving some ODEs on each mesh node and inverting large sparse matrices, often ill-conditionned.

  We first designed a method based on the Proper Orthogonal Decomposition (POD) technique: we project the linear system onto a well-chosen orthogonal basis of smaller dimension while still solving the ODES. We tried the method on both the bidomain and monodomain equations, and extended the tests on an HPC machine, in order to observe scalability performances. There is no improvment for the monodomain equations because its linear systems are well-conditionned. For the bidomain equations, the CPU time decreases by a factor of 10 between the full and reduced models, and better scalability performances.

  We secondly developped an eikonal model, in view of serious games applications for the Medic Activ project. The Dijkstra algorithm is used to solve the eikonal equation and the transmembrane potential is determined by the solution of a Mitchell-Schaeffer model on each mesh node. Some modifications where introduced to take into account re-excitability and allox re-entrant waves. Compared by the algorithm proposed by [38], the transmembrane potential comes from the solution of an underlying model, not through an approximation. This represent an innovation, to our knowledge not present in literature.

6.3. Medical applications of numerical models

- *Influence of Transmural Slow-Conduction Zones on the Long-Time Behaviour Of Atrial Arrhythmia. A Numerical Study with a Human Bilayer Atrial Model. [20], [31], [16]*

  Atrial fibrosis is known to be a factor in the perpetuation of atrial arrhythmia. Despite the thinness of atrial tissue, the fibrosis distribution may not be homogeneous through the entire thickness of the atria. The aim of this study is twofold. 1) We want to elucidate the respective influences of a transmural and a non-transmural distribution of fibrosis, described as a slow conduction zone, on the perpetuation of a rotor-like arrhythmic episode, compared to a control situation. 2) We aim to assess which is the more efficient ablation protocol between a) a lesion-box ablation, b) an ablation line connecting the fibrotic zone to the closest anatomical obstacle, c) ablation spots.

We used a bilayer monodomain representation of the atria that included transmural heterogeneities of fibre organisation, and an arrhythmic scenario composed of a rotor initiated near the pulmonary veins. This model allowed long simulations for a sustainable computational load. We observed that when the fibrosis was transmural, the centre of the rotor was anchored in the slow conduction zone and was stable during a 10 seconds simulation, whereas the other simulations showed meandering rotors that disappeared after a few seconds. In our model framework, only a transmural fibrosis distribution had a stabilizing effect on reentrant circuits. Furthermore, the lesion-box ablation and the line ablation were able to stop the arrhythmia, unlike the spot lesions. The bilayer model proved to be a good trade-off between accuracy and speed for observing the influence of transmural heterogeneities on atrial arrhythmia over long periods.

- *Effects of L-type Calcium channel and hERG blockers on the electrical activity of the human heart: A simulation study.*

  Class III and IV drugs affect cardiac hERG (IKr) and L-type calcium (ICaL) channels, resulting in complex alterations in repolarization with both anti and pro-arrhythmic consequences. Interpretation of their effects on cellular and ECG-based biomarkers for risk stratification is challenging. As pharmaceutical compounds often exhibit multiple ion channel effects, our goal is to investigate the simultaneous effect of ICaL and IKr block on human ventricular electrophysiology from ionic to ECG level. ECG simulations show that ICaL block results in shortening of the QT interval, ST elevation and reduced T wave amplitude, caused by reduction in APD and AP amplitude during the plateau phase, and in repolarization times. In contrast, IKr block results in QT prolongation and reduced T wave amplitude. Combined ICaL and IKr block are combined, the degree of ICaL block strongly determines QT interval whereas the effect of IKr block is more pronounced on the T wave amplitude.

6.4. Inverse problems

- *A Steklov-Poincaré approch to solve the inverse problem in electrocardiography [23]*

  In the cardiac electrophysiology imaging commu- nity the most widely used approach to solve the inverse prob- lem is the least square formulation with different Thikhonov regularizations. Clinicians are not yet fully satisfied by the technology that solves the inverse problem. Reformulating the inverse problem could bring new techniques to solve it. In this paper we use the Steklov-Poincare ´ formulation of the Cauchy problem in order to solve the inverse problem in electrocardiography imaging. We present in this work the technique and an algorithm of gradient descent. We also show numerical results based on simulated synthetical data.

- *A machine learning regularization of the inverse problem in electrocardiography imaging [22]*

  Radio-frequency ablation is one of the most ef- ficient treatments of atrial fibrillation. The idea behind it is to stop the propagation of ectopic beats coming from the pulmonary vein and the abnormal conduction pathways. Medical doctors need to use invasive catheters to localize the position of the triggers and they have to decide where to ablate during the intervention. ElectroCardioGraphy Imaging (ECGI) provides the opportunity to reconstruct the electrical potential and activation maps on the heart surface and analyze data prior to the intervention. The mathematical problem behind the reconstruction of heart potential is known to be ill posed. In this study we propose to regularize the inverse problem with a statistically reconstructed heart potential, and we test the method on synthetically data produced using an ECG simulator.

- *Inverse problem in electrocardiography via factorization method of boundary value problems : How to reconstruct epicardial potential maps from measurements on the torso ? [26]*

We are working on a new approach for solving the inverse problem of electrocardiography. This approach is based on an invariant embedding method: the factorization method of boundary values problems [35]. The idea is to embed the initial problem into a family of similar problems on subdomains bounded by a moving boundary from the torso skin to the epicardium surface. For the direct problem this method provides an equivalent formulation with two Cauchy problems evolving on this moving boundary and which have to be solved successively in opposite directions. This method calculates Neumman-Dirichlet and Dirichlet-Neumann operators on this moving boundary that satisfy Riccati equations. Regarding the inverse problem, mathematical analysis allows to write an optimal estimation of the epicardial potential based on a quadratic criterion. Then, the ill-posed behaviour of the inverse problem can be analyzed and a better regularization and discretization of the problem can be proposed. One of the advantages of this method is the computation of the potential at different times during cardial cycle: it is not necessary to repeat the resolution of all the equations at every time. In a first time the simplar case of a cylinder is considered. In a second time the method is applied to the 3D model of concentric spheres. The next step will be to use 3D deformed surfaces.

- *Reconstruction of 3D depolarization wavefronts from surface optical mapping images [33]*

  Starting from the diffusion-absorption equation of light in a tissue we solved the forward problem for excitation light using the `FreeFem++` software (www.freefem.org/ff++). We first considered a spherical wave front expanding in time: the tissue is depolarized inside the sphere. This choice allowed us to locate the position of the excitation. Using this representation of the wavefront, we obtained in silico data. We defined a functional to minimize and implemented the BFGS method to solve the inverse problem. We tested our method on in silico data and obtained good results. We next compared our results with an approach developed by Khait [36] and found that our method is more accurate and that we have less restrictions for the convergence of the method. We modified the wave front into ellipsoid in order to start working on experimental data.

CARTE Project-Team

6. New Results

6.1. Computation and Dynamical Systems

In [12], we analyzed the power of dynamical system that are robust to infinitesimal perturbations. While previous works on this question were limited to very specific kinds of systems such as piecewise constant derivative systems, we obtained results for a quite general class of systems: the main hypothesis being smoothness (which is already a prerequisite in systems that perform analog computation). We show that if a system is robust, then the language it recognizes is computable, and the converse: all computable languages can be recognized by a robust smooth system. Those results are true for discrete-time as well as continuous-time dynamical systems on bounded or unbounded domains.

We investigated in [23], [15], [33] the isomorphism (conjugacy) problem for dynamical systems. While the decidability in the one-dimensional case is a long-standing open problem, we characterize its exact complexity [23] in higher dimensions. Our result suggest that the isomorphism problem is easier than the factoring and embedding problem (decide if one dynamical system is a subsystem of another). A traditional approach to prove two dynamical systems are not isomorphic is to prove that they have different dynamical invariants. We characterised in terms of complexity and computability classes different well known dynamic invariants (periodic points, Turing degrees) in [23], [33].

While Turing machines are usually used for computing, it is an interesting model of dynamical systems, which looks very much like two-dimensional piecewise-affine maps. We investigated dynmacial invariants (entropy and Lyapunov exponents) for Turing machines, and proved quite surprisingly that they are computable. Essentially this means that Turing machines that do interesting computations must do it so slowly that this cannot be seen in their dynamics. This work will be presented in STACS 2014

6.2. Computability, Complexity and Topology

6.2.1. Complexity of real functionals

Computability and topology are closely related as computability assumptions impose topological restrictions: on a topological space, computable functions are continuous and continuous functions are computable relative to some oracle. In the same way, complexity assumptions as bounds on the computation time impose analytical restrictions, but in a way that is not understood yet. For functions from the real numbers to the real numbers, it is known that polynomial-time computable functions correspond to functions with a polynomial modulus of continuity. However for functions on other spaces no such correspondence is known. We investigate the particular case of norms on the space of continuous real functions defined on the unit interval. We introduce analytical characteristics of a norm, namely its dependency on points and the concept of *relevant points*, and use them to characterize the polynomial-time computable norms. This work was presented at LICS 2013 [19]. A full version including other results on non-deterministic complexity classes is currently submitted [28].

6.2.2. Higher-order complexity

While computability theory is well-developed and understood on large classes of topological spaces, complexity theory in analysis is still in its infancy. We argue that the usual way of representing mathematical objects by functions from finite strings to finite strings (order 1 functions) is not appropriate for general spaces. We show that as soon as the space becomes large in a topological sense, it cannot be represented by order 1 functions in a way that respects complexity notions, so we propose to represent objects using higher order functions over finite strings. However higher order complexity theory is not well-understood. The only known class to date is BFF, the class of Basic Feasible Functionals, which does not enjoy nice properties: some intuitively feasible functionals do not belong to the class. We develop a new way of carrying out complexity theory at higher order types, using an adaptation of game semantics. A preliminary version of this work was presented at CCA 2013 [26].

6.2.3. Irreversible computable functions

As mentioned before, computable functions must be continuous. It gives a simple way of proving that some operator is not computable by showing that it is discontinuous. We recall that a function f is computable if there is a *single* oracle Turing machine M that on each x given by an oracle, computes $f(x)$. The following weaker notion is also interesting: a function f*preserves computability* if for each computable x, $f(x)$ is computable. Preservation of computability no more implies continuity, so there is no topological argument to show that some operator does not preserve computability. We develop a strong notion of discontinuity and prove a general result stating that this notion of discontinuity prevents preservation of computability. We apply this result to solve an open problem about the non-computability of the ergodic decomposition. We show that many classical constructions in computability theory are instances of our result. Hence we exhibit deeper connections between computability and topology. The work has been accepted at STACS 2014 [22]. A partial result was published in [13].

6.3. Implicit Computational Complexity

In the setting of non-interference and implicit computational complexity, Emmanuel Hainry, Jean-Yves Marion, and Romain Péchoux presented a characterization of FPSPACE in a language with a fork/wait mechanism [20]. The language used in this work is a classical imperative language with while loops complemented with a mechanism to launch new processes through forks. The fork instruction is heavily inspired by C's fork/wait construction for Unix operating systems, which anchors this work in a down-to-earth setting. Using a type system that enforces a data-ramification on variables, they show that all programs that can be typed and are terminating compute an FPSPACE function, that with a natural evaluation strategy, they indeed use only polynomial space, and conversely that this type system is complete as all FPSPACE functions can be implemented in this language in a typable way.

Emmanuel Hainry and Romain Péchoux also used data-ramification combined with non-interference principles to effectively bound the memory used by object oriented languages in [21]. This work introduces a type system for an object oriented language (derived from java). This type system allows to compute polynomial bounds on the heap and stack used by a typable program, ensuring that if the program halts, it will only use memory under this explicit bound. As the typing procedure is doable in time polynomial in the size of the program, those bounds are easy to obtain, though not tight. Interesting features of this work include inheritance (with overloading and overriding) and, the ability to analyze programs with flow statements controled by objects, contrary to most other works in implicit computational complexity. In [24], Romain Péchoux has shown that the notion of (polynomial) interpretation over term rewrite systems can be adapated on a process language, a variant of the pi-calculus with process recursive definitions. This work shows that the order induced by simulation can be used wrt a given process semantics to infer time and space upper bounds on process resource usage (reduction length, size of sent values, ...).

6.4. Computer Virology

The study on behavioural malware detection has been continued. Guillaume Bonfante, Isabelle Gnaedig and Jean-Yves Marion have been developing an approach detecting suspicious schemes on an abstract representation of the behavior of a program, by abstracting program traces, rewriting given subtraces into abstract symbols representing their functionality. Considering abstract behaviors allows us to be implementation-independent and robust to variants and mutations of malware. Suspicious behaviors are then detected by comparing trace abstractions to reference malicious behaviors.

Model checking is a strong point of our approach: the predefined behavior patterns, used to abstract program traces, are defined by first order temporal logic formulas, as well as the reference suspicious behaviors, given in a signature. The infection problem can then be seen as the satisfaction problem of the formula of the signature by an abstracted trace of the program, which can be checked using existing model checking techniques

The previous work by the team involved abstracting trace automata by rewriting them with respect to a set of predefined behavior patterns defined as a regular language described by a string rewriting system [37], and then, by a term rewriting system [38], which allows to detect information leak.

This work has been finished this year by designing a probabilistic generalization of our approach. Introducing probabilities in our technique allows to express a pertinence degree of detection when analysis of the program results in an incomplete or uncertain program dataflow, or when abstraction cannot be performed reliably. Proposing malware detection with a probabilistic rate is finer and more realistic in practice than giving the binary answer of whether a program is infected or not.

Using a tropical semiring over the reals, they have presented a formalism relying on a weighted term rewriting mechanism, where a weight w, naturally associated to a probability p by the formula: $w = -log(p)$, represents the probability that the realized abstraction be right.

Detection of an abstract behavior has then be defined with respect to a threshold, and a program P exhibits an abstract behavior M if and only if one of its traces admits an abstract form realizing M with a weight not exceeding this threshold.

The weighted abstraction formalism has the advantage of providing a detection algorithm with the same complexity as in the unweighted case, that is linear in the size of the trace automaton [27].

6.5. Graph rewriting

Guillaume Bonfante and Bruno Guillaume provide a new graph rewriting framework adapted to Natural Language Processing. It involves a new form of edge transformation. A new termination technique is also described. The extended paper [17] is accepted for publication in Mathematical Structure in Computer Science.

CASCADE Project-Team (section vide)

CASSIS Project-Team

6. New Results

6.1. Automated Deduction

We develop general techniques which allow us to re-use available tools in order to build a new generation of solvers offering a good trade-off between expressiveness, flexibility, and scalability. We focus on the careful integration of combination techniques and rewriting techniques to design decision procedures for a wide range of verification problems.

6.1.1. Building and verifying decision procedures

Participants: Alain Giorgetti, Olga Kouchnarenko, Christophe Ringeissen, Elena Tushkanova.

We have developed a methodology to build decision procedures by using superposition calculi which are at the core of equational theorem provers. In [14], we have developed automated deduction techniques to prove properties about these superposition-based decision procedures. To this aim, we have further investigated the use of schematic superposition, to check the termination and the combinability of superposition-based procedures. We have worked on the development of a framework for specifying and verifying superposition-based procedures. We have designed an implementation in Maude of the schematic superposition calculus. Thanks to this implementation we automatically derive termination of superposition for a couple of theories of interest in verification.

Until now, schematic superposition was only studied for standard superposition. In [53], [55], we introduce a schematic superposition calculus modulo a fragment of arithmetics, namely the theory of Integer Offsets. This new schematic calculus is used to prove the decidability of the satisfiability problem for some theories extending Integer Offsets. We illustrate our theoretical contribution on theories representing extensions of classical data structures, e.g., lists and records. Our Maude-based implementation has been extended to incorporate this new schematic superposition calculus modulo Integer Offsets. It enables automatic decidability proofs for theories of practical use.

6.1.2. Hierarchical combination of unification procedures

Participant: Christophe Ringeissen.

In [45], [54], a novel approach is described for the combination of unification algorithms for two equational theories which share function symbols. We are able to identify a set of restrictions and a combination method such that if the restrictions are satisfied the method produces a unification algorithm for the union of non-disjoint equational theories. Furthermore, we identify a class of theories satisfying the restrictions. The critical characteristics of the class is the hierarchical organization and the shared symbols being restricted to "inner constructors". Our approach can be applied to theories used for the analysis of protocols. The property of having an inner constructor in one side of an equality is common in the use of exponentiation in Diffie-Hellman inspired key agreement protocols. We are working on considering additional hierarchical theories. A possible candidate theory is a partial theory of Cipher Block Chaining.

6.1.3. Unification modulo equational theories of cryptographic primitives

Participant: Michaël Rusinowitch.

Some attacks exploit in a clever way the interaction between protocol rules and algebraic properties of cryptographic operators. In [74], we provide a list of such properties and attacks as well as existing formal approaches for analyzing cryptographic protocols under algebraic properties.

We have further investigated unification problems related to the Cipher Block Chaining (CBC) mode of encryption. We first model chaining in terms of a simple, convergent, rewrite system over a signature with two disjoint sorts: list and element. The 2-sorted convergent rewrite system is then extended into one that captures a block chaining encryption-decryption mode at an abstract level, (using no AC-symbols); unification modulo this extended system is shown to be decidable [15].

6.2. Security Protocol Verification

The design of cryptographic protocols is error-prone. Without a careful analysis, subtle flaws may be discovered several years after the publication of a protocol, yielding potential harmful attacks. In this context, formal methods have proved their interest for obtaining good security guarantees. Many analysis techniques have been proposed in the literature [70]. We have edited a book [62] where each chapter presents an important and now standard analysis technique. We develop new techniques for richer primitives, wider classes of protocols and higher security guarantees. In Section 6.4.3 we consider derived testing techniques for verifying protocol implementations.

6.2.1. Voting protocols

Participants: Véronique Cortier, David Galindo-Chacon, Stéphane Glondu, Malika Izabachene, Steve Kremer, Cyrille Wiedling.

Voting is a cornerstone of democracy and many voting systems have been proposed so far, from old paper ballot systems to purely electronic voting schemes. Although many works have been dedicated to standard protocols, very few address the challenging class of voting protocols. We have studied several protocols that are currently in use:

- Helios is an open-source web-based end-to-end verifiable electronic voting system, used e.g. by UCL and the IACR association in real elections. One main advantage of Helios is its verifiability, up-to the ballot box (a dishonest ballot box may add ballots). We have defined a variant of Helios, named Belenios, that prevents from ballot stuffing, even against a dishonest ballot box. Our approach consists in introducing an additional authorities that provides credentials that the ballot box can verify but not forge. This new version has been implemented by Stéphane Glondu and has been tested in a mock election in the teams Cassis and Caramel.
  We have proved computational security for both ballot secrecy and full verifiability (due to our credentials). Helios, as well as Belenios, makes use of threshold decryption, to ensure that decryption keys are distributed among several authorities, yet allowing decryption even some of the authorities are missing. We have provided a fully distributed (with no dealer) threshold cryptosystem suitable for the Helios voting system (in particular, suitable to partial decryption), and prove it secure under the Decisional Diffie-Hellman assumption [40]. Ballot privacy of Belenios then follows from ballot privacy of Helios. For full verifiability, we had first to adapt existing definitions of verifiability in the case of a corrupted ballot box and then prove verifiability of Helios [60].
- The Section 07 of CNRS (now split into Section 06 and Section 07) has proposed a voting protocol for Face-to-Face meetings to enhanced the verifiability of an election run through electronic devices. We have formally modeled this protocol and proved both ballot secrecy and verifiability [32].

Security based on cryptography relies on the fact that certain operations (such as decrypting) are computationally infeasible. However, e-voting protocols should also guarantee privacy in the future, when computers will have an increased computational power and will be able e.g. to break nowadays keys. Such privacy in the future is called *everlasting privacy* and we have proposed a definition of *practical everlasting privacy* [31]. As an illustration, we show that several variants of Helios (including Helios with Pedersen commitments) and a protocol by Moran and Naor achieve practical everlasting privacy, using the ProVerif and the AKiSs tools, which we had to adapt to cope with everlasting privacy.

We have written a popularization science paper on e-voting in Interstices[4].

[4] https://interstices.info/jcms/int_68258/vote-par-internet

6.2.2. Other families of protocols

Participants: Véronique Cortier, Steve Kremer, Robert Künnemann, Cyrille Wiedling.

Securing routing Protocols. The goal of routing protocols is to construct valid routes between distant nodes in the network. If no security is used, it is possible for an attacker to disorganize the network by maliciously interacting with the routing protocols, yielding invalid routes to be built. We have proposed a new model and an associated decision procedure to check whether a routing protocol can ensure that honest nodes only accept valid routes, even if one of the nodes of the network is compromised. This result has been obtained for a bounded number of sessions, adapting constraint solving techniques to node topologies as well as some families of recursive tests, used in routing protocols [16].

Security APIs. In some systems, it is not possible to trust the host machine on which sensitive codes are executed. In that case, security-critical fragments of a program should be executed on some tamper resistant device (TRD), such as a smartcard, USB security token or hardware security module (HSM). The exchanges between the trusted and the untrusted infrastructures are ensured by special kind of API (Application Programming Interface), that are called *security APIs*. We have designed a generic API for key-management based on key hierarchy [23], that can self-recover from corruption of arbitrary keys, provided the number of corrupted, active keys is smaller than some threshold. In [50], we propose a universally composable key management functionality and show how to achieve a secure, distributed implementation on TRDs. We are currently also working on automated verification of security APIs (and more generally protocols that require global mutable state). A tool implementation using the tamarin prover as a backend is currently in progress.

6.2.3. Automated verification of indistinguishability properties.

Participants: Rémy Chrétien, Véronique Cortier, Stéphane Glondu, Steve Kremer.

New emerging classes of protocols such as voting protocols often require to model less classical security properties, such as anonymity properties, strong versions of confidentiality and resistance to offline guessing attacks. Many of these properties can be modelled using the notion of indistinguishability by an adversary, which can be conveniently modeled using process equivalences.

Static case. The YAPA tool [17] can check static equivalence for convergent equational theories. It is proved to terminate for a wide class of equational theories that includes subterm convergent theories (e.g. encryption, signatures, pairing and hash) and layered convergent theories (e.g. blind signatures). The procedure is generic in the sense that it remains sound and complete (but may not terminate) for any convergent theory. It has been implemented in the YAPA tool[5].

Active case. We have shown that, for arbitrary equational theories, verifying indistinguishability properties such as trace equivalence in security protocols amounts to deciding the equivalence of constraint systems, i.e., checking whether they have the same set of solutions [20]. When considering the equational theory corresponding to the standard primitives, Vincent Cheval has proposed a decision procedure for checking equivalence of set constraints, which yields a procedure for checking trace equivalence [73]. We have extended this decision procedure to the case where the attacker can observe the *length* of messages [37]. This yields the discovery of a new attack on the biometric passport. This attack has been implemented and successfully tested on a small set of passports. This attack is explained in details in a webpage[6] and has obtained some press coverage.

Active case, unbounded number of sessions. Rémy Chrétien has started a PhD on deciding trace equivalence for an unbounded number of sessions. He has shown that for some classes of protocols, decidability of trace equivalence can be reduced to equivalence of deterministic pushdown automata [38]. Equivalence of deterministic pushdown automata is decidable [81] and the corresponding decision procedure is currently implemented by Géraud Senizergues. Based on his tool, we are developing a tool for automatically checking equivalence, for an unbounded number of sessions.

[5] http://www.lsv.ens-cachan.fr/~baudet/yapa/
[6] http://www.loria.fr/ glondu/epassport/attack-lengths.html

6.2.4. Securely Composing Protocols

Participants: Véronique Cortier, Steve Kremer, Éric Le Morvan.

Protocols are often built in a modular way. For example, authentication protocols may assume pre-distributed keys or may assume secure channel. However, when an authentication protocol has been proved secure assuming pre-distributed keys, there is absolutely no guarantee that it remains secure when executing a real protocol for distributing the keys. How the security of these protocols can be combined is an important issue that is studied in the PhD thesis recently started by Éric Le Morvan.

A related problem arises when several protocols use the same secrets, e.g. the same keys. While each protocol may be proved secure in isolation, the protocols may become insecure when executed in parallel. In [21] we study whether password protocols can be safely composed, even when a same password is reused. It seems indeed unrealistic to suppose that users do not re-use the same password for different applications. More precisely, we present a transformation which maps a password protocol that is secure for a single protocol session (a decidable problem) to a protocol that is secure for an unbounded number of sessions. Our result provides an effective strategy to design secure password protocols: (i) design a protocol intended to be secure for one protocol session; (ii) apply our transformation and obtain a protocol which is secure for an unbounded number of sessions. Our technique also applies to compose different password protocols allowing us to obtain both inter-protocol and inter-session composition.

6.2.5. Soundness of the Dolev-Yao Model

Participants: Véronique Cortier, Guillaume Scerri.

All the previous results rely on symbolic models of protocol executions in which cryptographic primitives are abstracted by symbolic expressions. This approach enables significantly simple and often automated proofs. However, the guarantees that it offers have been quite unclear compared to cryptographic models that consider issues of complexity and probability. A somewhat recent line of research consists in identifying cases where it is possible to obtain the best of both cryptographic and formal worlds: fully automated proofs and strong, clear security guarantees.

A first approach consists in proving that symbolic models (as the ones studied on the previous sections) are actually *sound* w.r.t. cryptographic models, provided the primitives satisfy some (strong) security properties. Soundness result are usually established for some set of cryptographic primitives and extending the result to encompass new primitives typically requires redoing most of the work. In [35], we propose a notion of computational soundness, amenable to modular extensions. Specifically, we prove that a deduction sound implementation of some arbitrary primitives can be extended to include all standard primitives (asymmetric ans symmetric encryption, public data-structures - e.g. pairings or list, signatures, MACs, and hashes) without repeating the original proof effort. Furthermore, our notion of soundness concerns cryptographic primitives in a way that is independent of any protocol specification language.

Such soundness results require however strong hypotheses on the implementation. For example, primitives must be tagged to avoid confusion between e.g. pairs and encryption. Gergei Bana and Hubert Comon have proposed a new framework [67] where the symbolic model now specifies what an attacker *cannot* do instead of specifying what it can do. Checking protocols security can then be reduced to checking inconsistency of some set of first order formula. During his PhD, Guillaume Scerri studies how to develop a (polynomial) decision procedure for deciding consistency of sets of formulas, for some class of formulas corresponding to security protocols [39].

6.3. Model-based Verification

We have investigated extensions of regular model-checking to new classes of rewrite relations on trees. We have studied specification and proof of modular imperative programs.

6.3.1. Verification of Linear Temporal Patterns over Finite and Infinite Traces

Participants: Pierre-Cyrille Héam, Vincent Hugot, Olga Kouchnarenko.

In the regular model-checking framework, reachability analysis can be guided by temporal logic properties, for instance to achieve the counter example guided abstraction refinement (CEGAR) objectives. A way to perform this analysis is to translate a temporal logic formula expressed on maximal rewriting words into a "rewrite proposition" – a propositional formula whose atoms are language comparisons, and then to generate semi-decision procedures based on (approximations of) the rewrite proposition. In [13] we have investigated suitable semantics for LTL on maximal rewriting words and their influence on the feasibility of a translation, and we have proposed a general scheme providing exact results for a fragment of LTL corresponding mainly to safety formulæ, and approximations for a larger fragment.

6.3.2. Approximations Techniques for Regular Model-Checking

Participants: Aloïs Dreyfus, Pierre-Cyrille Héam, Olga Kouchnarenko.

We address the following general problem of regular model-checking: decide whether $R^*(L) \cap L_p = \varnothing$ where R^* is the reflexive and transitive closure of a successor relation R, and L and L_p are both regular tree languages. Considering a relation R on finite words and a regular language L encoding the initial configurations of a system, the set $R^*(L)$ of accessible words is not necessarily regular. Therefore, a way to verify safety properties is to over-approximate the set of reachable words by a regular language. In [42], we develop new efficient approximation techniques based on syntactic criteria. When these syntactic over-approximations are too coarse, we propose CEGAR-like techniques to refine them using counter-examples. The approach has been successfully applied to verify mutual exclusion protocols.

6.4. Model-based Testing

Our research in Model-Based Testing (MBT) aims to extend the coverage of tests. The coverage refers to several artefacts: model, test scenario/property, and code of the program under test. The test generation uses various underlying techniques such as symbolic animation of models [75], or symbolic execution of programs by means of dedicated constraints, SMT solvers, or model-checkers.

6.4.1. Automated Test Generation from Behavioral Models

Participants: Fabrice Bouquet, Kalou Cabrera, Jérome Cantenot, Frédéric Dadeau, Jean-Marie Gauthier, Julien Lorrain.

We have developed an original model-based testing approach that takes a behavioural view (modelled in UML) of the system under test and automatically generates test cases and executable test scripts according to model coverage criteria. We continue to extend this result to SysML specifications for validating embedded systems [30]. To allow the test generation from SysML models, in [47] we study the transformation into a low level language suitable for hardware specification.

In the context of software evolution, we have worked on exploiting the evolution of requirements in order to classify test sequences, and precisely target the parts of the system impacted by this evolution. We have proposed to define the life cycle of a test via three test classes: (i) Regression, used to validate that unimpacted parts of the system did not change, (ii) Evolution, used to validate that impacted parts of the system correctly evolved, and (iii) Stagnation, used to validate that impacted parts of the system did actually evolve. The associated algorithms are under implementation in a dedicated prototype already used in the SecureChange European project.

6.4.2. Scenario-Based Verification and Validation

Participants: Fabrice Bouquet, Kalou Cabrera, Frédéric Dadeau.

Test scenarios represent an abstract test case specification that aims at guiding the model animation in order to produce relevant test cases. Contrary to the previous section, this technique is not fully automated since it requires the user to design the scenario, in addition to the model.

We have designed a scenario based testing language for UML/OCL that can be either connected to a model animation engine or to a symbolic animation engine, based on a set-theoretical constraint solver [75]. In the context of the ANR TASCCC project, we investigated the automation of test generation from Security Functional Requirements (SFR), as defined in the Common Criteria terminology. To achieve that, we worked on the definition of description patterns for security properties, to which a given set of SFRs can be related. These properties are used to automatically generate test scenarios that produce model based test cases. The traceability, ensured all along the testing process, makes it possible to provide evidences of the coverage of the SFR by the tests, required by the Common Criteria to reach the highest Evaluation Assurance Levels.

We have proposed a dedicated formalism to express test properties. translated into a finite state automaton which describes a monitor of its behaviors [36]. We have proposed dedicated property coverage criteria that can be used either to measure the property coverage of a given test suite, or to generate test cases, exercising nominal or robustness aspects of the property. This process has been fully tool-supported into an integrated software prototype[7] [41].

In the context of the SecureChange project, we have also investigatde the evolution of test scenarios. As the system evolves, the model evolves, and the associated test scenarios may also evolve. We are currently extending the test generation and management of system evolutions to ensure the preservation of the security.

6.4.3. Mutation-based Testing of Security Protocols

Participants: Frédéric Dadeau, Pierre-Cyrille Héam, Ghazi Maatoug, Michaël Rusinowitch.

Verification of security protocols models is an important issue. Nevertheless, the verification reasons on a model of the protocol, and does not consider its concrete implementation. While representing a safe model, the protocol may be incorrectly implemented, leading to security flaws when it is deployed. We have proposed a model-based penetration testing approach for security protocols [9]. This technique relies on the use of mutations of an original protocol, proved to be correct, for injecting realistic errors that may occur during the protocol implementation (e.g. re-use of existing keys, partial checking of received messages, incorrect formatting of sent messages, use of exponential/xor encryption, etc.). Mutations that lead to security flaws are used to build test cases, which are defined as a sequence of messages representing the behavior of the intruder. secret. We have applied our technique on protocols designed in HLPSL, and implemented a protocol mutation tool that performs the mutations. The mutants are then analyzed by the CL-Atse [82] front-end of the AVISPA toolset [64]. We have experimented our approach on a set of protocols, and we have shown the relevance of the proposed mutation operators and the efficiency of the CL-Atse tool to conclude on the vulnerability of a protocol and produce an attack trace that can be used as a test case for implementations. We applied our approach on the Paypal Express protocol, and we were able to retrieve an existing attack trace on this protocol[8]. We are now investigating the transformation of an attack trace into executable tests scripts. To achieve that, we propose to automatically generate skeletons of Java test programs that the validation engineer only has to fill in order to concretize the steps of the test. A first experience in this direction has been described in [48].

6.4.4. Rewriting-based Mathematical Model Transformations

Participants: Walid Belkhir, Alain Giorgetti.

Since 2011 we collaborate with the Department “Temps-Fréquence” of the FEMTO-ST institute (Franche-Comté Electronique Mécanique Thermique et Optique - Sciences et Technologies, CNRS UMR 6174) on the formalization of asymptotic methods (based on two-scale convergence).The goal is to design a software, called *MEMSALab*, for the automatic derivation of multiscale models of arrays of micro- and nanosystems. In this domain a model is a partial differential equation. Multiscale methods approximate it by another partial differential equation which can be numerically simulated in a reasonable time. The challenge consists in taking into account a wide range of geometries combining thin and periodic structures with the possibility of multiple nested scales. We have designed a transformation language facilitating the design of MEMSALab [18]. It

[7] A video of the prototype is available at: http://vimeo.com/53210102
[8] http://www.nbs-system.com/blog/faille-securite-magento-paypal.html

is proposed as a Maple™ package for rule-based programming, rewriting strategies and their combination with standard Maple™ code. We illustrate the practical interest of this language by using it to encode two examples of multiscale derivations, namely the two-scale limit of the derivative operator and the two-scale model of the stationary heat equation. A more general framework for the derivation of the multi-scale models was established in [26].

6.4.5. Code-related Test Generation and Static Analysis

Participants: Fabrice Bouquet, Frédéric Dadeau, Ivan Enderlin, Alain Giorgetti.

We have designed a new annotation language for PHP, named PRASPEL (for *PHP Realistic Annotation SPEcification Language*). This language relies on *realistic domains* which serve two purposes. First, they assign to a data a domain that is supposed to be specific w.r.t. a context in which it is employed. Second, they provide two features that are used for test generation: (i) *samplability* makes it possible to automatically generate a value that belongs to the realistic domain so as to generate test data, (ii) *predicability* makes it possible to check if the value belongs to a realistic domain. This approach is tool-supported in a dedicated framework for PHP which makes it possible to produce unit test cases using random data generators, execute the test cases on an instrumented implementation, and decide the conformance of the code w.r.t. the annotations by runtime assertion checking. This principle has been extended to generate grammar-based textual data based on various strategies, namely uniform random generation, bounded exhaustive generation and rule-coverage-based test generation. In a recent work, we have proposed a dedicated constraint solver for PHP arrays [44] aiming to avoid rejection during the generation of array structures.

6.4.6. Random Testing

Participants: Aloïs Dreyfus, Pierre-Cyrille Héam, Olga Kouchnarenko.

The random testing paradigm represents a quite simple and tractable software assessment method for various testing approaches. When performing random testing, the random sampler is supposed to be independent of tester choices or convictions: a solution is to exploit uniform random generators.

In [78] a method is proposed for drawing paths in finite graphs uniformly, and it is explained how to use these techniques for testing C programs within a control flow graph based approach. Nevertheless, as finite graphs often provide strong abstractions of the systems under test, many abstract tests generated by the approach cannot be played on the implementation. In [79], we have proposed a new approach, extending [78], to manage stack-call during the random test generation while preserving uniformity. In [61], we go further by investigating a way to biase the random testing, in order to optimize the probability to fulfil a coverage criterion. The new approaches have been implemented in a prototype and experimented on several examples. A similar approach for grammar based testing is developped in [43]: we show how to hedge the random generation of execution trees to optimize the probability of covering either all rules or all non terminal symbols.

6.5. Verification of Collaborative Systems

We investigate security problems occurring in decentralized systems. We develop general techniques to enforce read and update policies for controlling access to XML documents based on recursive DTDs (Document Type Definition). Moreover, we provide a necessary and sufficient condition for undoing safely replicated objects in order to enforce access control policies in an optimistic way.

6.5.1. Automatic Analysis of Web Services Security

Participants: Walid Belkhir, Michaël Rusinowitch, Mathieu Turuani, Laurent Vigneron.

Automatic composition of web services is a challenging task. Many works have considered simplified automata models that abstract away from the structure of messages exchanged by the services. For the domain of secured services (using e.g. digital signing or timestamping) we propose a novel approach to automated orchestration of services under security constraints. Given a community of services and a goal service, we reduce the problem of generating a mediator between a client and a service community to a security problem where an intruder should intercept and redirect messages from the service community and a client service till reaching a satisfying state. This orchestration specification is expressed in ASLan language, a formal language designed for modeling Web Services tied with security policies that was developed in AVANTSSAR project. The AVANTSSAR Orchestrator (presented in [28]) generates an attack trace describing the execution of a the mediator and translates it into ASLan. Then we can check with automatic tools that this ASLan specification verifies required security properties such as secrecy and authentication. If no flaw is found, we can compile the ASLan specification into a Java servlet that can be used to execute the orchestration.

In [34] we introduce an alternative approach based on *fresh-variable automata*, a natural extension of finite-state automata over infinite alphabet. In this model the transitions are labeled with constants or variables that can be refreshed in some specified states. We prove several closure properties for this class of automata and study their decision problems. We show the applicability of our model to Web services handling data from an infinite domain. We introduce a notion of simulation that enables us to reduce the Web service composition problem to the construction of a simulation of a target service by the asynchronous product of existing services, and prove that this construction is computable. We now work on synthesizing composed services that satisfy required security policies.

6.5.2. Secure Querying and Updating of XML Data

Participants: Abdessamad Imine, Houari Mahfoud, Michaël Rusinowitch.

It is increasingly common to find XML views used to enforce access control as found in many applications and commercial database systems. To overcome the overhead of view materialization and maintenance, XML views are necessarily virtual. With this comes the need for answering XML queries posed over virtual views, by rewriting them into equivalent queries on the underlying documents. A major concern here is that query rewriting for recursive XML views is still an open problem, and proposed approaches deal only with non-recursive XML views. Moreover, a small number of works have studied the access rights for updates. In [51], we present SVMAX (Secure and Valid MAnipulation of XML), the first system that supports specification and enforcement of both read and update access policies over arbitrary XML views (recursive or non). SVMAX defines general and expressive models for controlling access to XML data using significant class of XPath queries and in the presence of the update primitives of W3C XQuery Update Facility. Furthermore, SVMAX features an additional module enabling efficient validation of XML documents after primitive updates of XQuery. The wide use of W3C standards makes of SVMAX a useful system that can be easily integrated within commercial database systems as we will show. We give extensive experimental results, based on real-life DTDs, that show the efficiency and scalability of our system.

We introduce in [49] an extension of hedge automata called bidimensional context-free hedge automata, proposing a new uniform representation of vertical and horizontal computation steps in unranked ordered trees. We also extend the parameterized rewriting rules used for modeling the W3C XQuery Update Facility in previous works, by the possibility to insert a new parent node above a given node. Since the rewrite closure of hedge automata languages with these extended rewriting systems is a computable context-free hedge language we can perform some static typechecking on these XML transformations.

6.5.3. On Adding Friends Problem in Social Networks

Participants: Bao Thien Hoang, Abdessamad Imine.

Online social networks are currently experiencing a peak and they resemble real platforms of social conversion and content delivery. Indeed, they are exploited in many ways: from conducting public opinion polls about any political issue to planning big social events for a large public. To securely perform these large-scale computations, current protocols use a simple secret sharing scheme which enables users to obfuscate their inputs. However, these protocols require a minimum number of friends, i.e. the minimum degree of the social graph should be not smaller than a given threshold. Often this condition is not satisfied by all social graphs. Yet we can reuse these graphs after some structural modifications consisting in adding new friendship relations. In this paper, we provide the first definition and theoretical analysis of the "*adding friends*" problem. We formally describe this problem that, given a graph G and parameter c, asks for the graph satisfying the threshold c that results from G with the minimum of edge-addition operations. We present algorithms for solving this problem in centralized social networks [33]. An experimental evaluation on real-world social graphs demonstrates that our protocols are accurate and inside the theoretical bounds.

6.5.4. Access Control Models for Collaborative Applications

Participants: Fabrice Bouquet, Abdessamad Imine, Michaël Rusinowitch.

The importance of collaborative systems in real-world applications has grown significantly over the recent years. The most of new applications are designed in a distributed fashion to meet collaborative work requirements. Among these applications, we focus on Real-Time Collaborative Editors (RCE) that provide computer support for modifying simultaneously shared documents, such as articles, wiki pages and programming source code by dispersed users. Although such applications are more and more used into many fields, the lack of an adequate access control concept is still limiting their full potential. In fact, controlling access in a decentralized fashion in such systems is a challenging problem, as they need dynamic access changes and low latency access to shared documents. In [19], we propose a generic access control model based on replicating the shared document and its authorization policy at the local memory of each user. We consider the propagation of authorizations and their interactions. We propose an optimistic approach to enforce access control in existing collaborative editing solutions in the sense that the access control policy can be temporarily violated. To enforce the policy, we resort to the selective undo approach in order to eliminate the effect of illegal document updates. Since, the safe undo is an open issue in collaborative applications. We investigate a theoretical study of the undo problem and propose a generic solution for selectively undoing operations. Finally, we apply our framework on a collaboration prototype and measure its performance in the distributed grid GRID'5000 to highlight the scalability of our solution.

However, verifying whether the combination of access control and coordination protocols preserves the data consistency is a hard task since it requires examining a large number of situations. In [52], we specify this access control protocol in the first-order relational logic with Alloy, and we verify that it preserves the correctness of the system on which it is deployed, namely that the access control policy is enforced identically at all participating user sites and, accordingly, the data consistency remains still maintained.

CASTOR Team

6. New Results

6.1. High order approximation of the two fluid Braginskii model

Participants: Sebastian Minjeaud, Richard Pasquetti.

We work on a two fluid physical model developed in close connection with Ph. Ghendrih (IRFM). It is based on the electrostatic assumption, i.e. the magnetic field is given (the magnetic field induced by the plasma itself is negligible), on the hypothesis of electroneutrality (the density of ions and electrons are proportional) and on the Braginskii closure. On the basis of the conservation equations of density, electron and ion velocities, electron and ion temperatures and electrical charges, a set of 10 non-linear and strongly anisotropic coupled partial differential equations (PDE) can be set up. A high order Fourier-SEM (Spectral Element Method) code is currently developed to address this problem. This Fourier-SEM code is close to be operational for the full set of PDEs in a 3D toroidal geometry. The torus section is discretized with quadrangular elements and Fourier expansions are used in the toroidal direction. In time one uses an RK3 (third order Runge-Kutta) IMEX (Implicit-Explicit), so that the Lorentz terms are handled implicitly. The capability of this code to handle a strongly anisotropic diffusion in a 3D toroidal geometry has already been tested. The Braginskii closure has been implemented. The Bohm boundary conditions at the plates are also considered. In 2013, we worked on a parallel version of the code and on the robustness of our algorithms, to be able to make long time computations, e.g. a few hundreds of thermal times.

6.1.1. Parallelization of the full Braginskii (FBGKI) code.

A first parallel version of the FBGKI code is now operational. Tests were made on the Computional center of the University of Nice-Sophia Antipolis. Tests on a large number of processors have however not yet been carried out, since presently our goal is to improve the robustness of our algorithms. Our parallelization strategy is based on a domain decomposition technique in the poloidal plane, where the spectral element approximation is local. On the contrary, in toroidal direction the approximation is global since based on Fourier expansions.

6.1.2. Numerical stabilization.

A spectral vanishing viscosity (SVV) technique was implemented in the sequential version of the code. Such a stabilization technique relies on the idea of introducing viscosity in the high frequency range of the spectral element / Fourier approximation. Such an approach was first proposed for hyperbolic problems, typically the Burgers equation (E. Tadmor, 1989). Later on, it was used for the large-eddy simulation of turbulent flows. Thus, we have a large experience of the SVV stabilization for the computation of turbulent wake flows.

6.1.3. Projection techniques.

A projection technique is used to enforce the divergence free constraint of the current. Projection techniques have been developed for a long time, in the frame of the Navier-Stokes equations to provide efficient algorithms when incompressible flows are concerned. For the Braginskii system, it appears natural to make use of such techniques for the current. Different projection techniques have been implemented in the FBGKI code, from the most classical one (Chorin-Temam, 1969) to the most recent. It turns out however that using projection techniques is less straightforward for Braginskii than for Navier-Stokes. We actively work on this point in order to cure some not yet understood failures of convergence with the time-step.

In the frame of the Eurofusion program, it is planned to check this version of the code on a simple configuration proposed by the EPFL (Paolo Ricci) where experimental as well as numerical results are available.

6.2. Equilibrium reconstruction and current density profile identification

Participants: Jacques Blum, Cédric Boulbe, Blaise Faugeras.

EQUINOX is a real-time equilibrium reconstruction code. It solves the equation satisfied by the poloidal flux in a computation domain, which can be the vacuum vessel for example, using a P1 finite element method and solves the inverse problem of the identification of the current density profile by minimizing a least square cost-function. It uses as minimal input the knowledge of the flux and its normal derivative on the boundary of the computation domain. It can also use supplementary constraints to solve the inverse problem: interferometric, polarimetric and MSE measurements. Part of the work reported here has been done in the frame of a RTM-JET contract [2]

6.2.1. Direct use of the magnetic measurements

The code EQUINOX was not originally designed to take as magnetic inputs directly the magnetic measurements, as it should be the case in the ITM (Integrated Tokamak Modeling European platform), but some outputs from the real-time codes APOLO at ToreSupra and XLOC at JET. These codes provide EQUINOX with the values of the flux and its normal derivative on a closed contour defining the boundary of the computation domain (this contour can be the limiter for example). As a consequence the main difficulty arising in the objective of integrating the code EQUINOX in the ITM structure is to interpolate between the magnetic measurements (flux loops and poloidal B-probes) with a machine independent method. A solution to do this is to use toroidal harmonics functions as a basis for the decomposition of the poloidal flux in the vacuum region in complement to the contribution of the PF coils. The first version of the algorithm implementing this method for EQUINOX-ITM developed in 2012 has been updated and tested during 2013:

- WEST and JET: This method can provide an alternative tool, comparable to APOLO (for Tore Supra) and FELIX (for JET), to compute the plasma boundary in real time from the magnetic measurements. Some twin experiments for WEST have been successfully conducted. In a first step the equivalents of magnetic measurements were generated using the free boundary equilibrium code CEDRES++. In a second step these measurements were used by the toroidal harmonics algorithm to reconstruct the plasma boundary. Additional calculations aiming at validating the design of the WEST magnetic diagnostics have been performed. They consisted in checking the equilibrium reconstruction accuracy with respect to: (i) a reduced number of magnetic sensors; (ii) noise on magnetic sensor and/or current measurements. Then, experiments on the possibility to reconstruct not only the plasma boundary but also the current density have been conducted. A paper on this subject is accepted for publication [13]. The same algorithm has been tested using real JET measurements in order to provide an equilibrium reconstruction code that directly uses the magnetic measurements instead of using FELIX as an intermediate preprocessing of the measurements.
- EFDA-ITM (Task WP13-ITM-IMP12-ACT3): EQUINOX-ITM has been upgraded and tested on the new gateway machine of the ITM. The Kepler actor was tested and used on 3 different tokamaks (JET, Tore Supra and WEST) (with F. Imbeaux, T. Aniel, P. Moreau, E. Nardon (CEA)). A benchmark work is on going between the codes Equal, Efit and Equinox on JET shot 74221 (with Dimitriy Yadykin and Wolfgang Zwingmann).

6.3. Evolutive equilibrium and transport coupling and optimization of scenarii

Participants: Jacques Blum, Cédric Boulbe, Blaise Faugeras.

6.3.1. New developments in the direct evolutive version of CEDRES++

6.3.1.1. External circuits and saddle currents in the blankets

In the previous version of the free boundary equilibrium code CEDRES++, each PF-coil (Poloidal Field Coil) was considered separately. In the evolutive version, a voltage was applied to each coil. In the machine, PF-coils can be connected in series or in parallel with one or several power supplies. In order to consider more realistic configurations of the PF system, the model used in CEDRES++ has been generalized to circuits involving several coils and supplies connected in series or in parallel. This model has been implemented in CEDRES++ and has been tested on simple configurations with circuits composed of one supply and several coils connected in series. More complicated configurations like circuits with several supplies and coils in series and in parallel in the same circuit have to be tested. This will be done on a WEST test case. A model for saddle currents in the blankets has also been implemented. This model is actually under validation on DEMO geometry.

6.3.1.2. Coupling CEDRES++ with a feedback controller (task ITM-WP13-ITM-IMP12-ACT1-T3)

Cedres++ has been successfully coupled with a controller generated from the true TCV hybrid Simulink controller in an ITM (Integrated Tokamak Modeling) workflow. The "yoyo" discharge (shot 40475) has been reproduced. In Figure 1, the vertical position of the magnetic axis simulated matches the experimental one.

../../../../projets/castor/IMG/controlcedres.png

Figure 1. Simulation of Yoyo discharge on TCV - Comparison between Z axis simulated and Z axis obtained from experiments

6.3.1.3. Cedres++ - transport coupling

Last year, different algorithms coupling free boundary equilibrium solvers and transport solvers (CEDRES++-diff, CEDRES++-CRONOS, CEDRES++-ETS, FREEBIE-CRONOS) have been developed. The ETS-C coupling between CEDRES++ and the transport solver under the ITM environment has been finalized this year. A simulation of a VDE test case has been performed (task WP13-ITM-IMP12-ACT1-T2). A benchmark between the different strategies has been performed in order to validate the numerical methods required to ensure stability of the coupling system and to compare the physical simulations of each model. A benchmark between CEDRES++-diff solving the resistive diffusion equation coupled to CEDRES and the CEDRES-CRONOS coupling has been performed on an ITER test case. Some divergences between the two codes appear and are not fully understood despite long investigations. This difficulty has led us to delay the introduction of non inductive terms in the resistive diffusion equation implemented in CEDRES++-diff. These developments will be realized when the different coupling strategies will be fully validated.

6.3.2. Reasearch of optimal trajectories for the preparation of Tokamak discharges

A new approach has been developed for the optimization of dynamic plasma scenarios in tokamaks. This task has been formulated as an optimal control problem, using numerical solution methods for optimization problems with PDE constraints. Due to free boundary setting, a new linearization of the non-linear equations has been introduced, which is consistent with the numerical discretization. It is this consistency that guarantees that the method converges to the optimum.

6.4. Parallel Kelvin-Helhmohltz-like instability in edge plasma

Participants: Hervé Guillard, Boniface Nkonga, Marco Bilanceri, Giorgio Giorgiani.

A large part of this year activities have been devoted to the investigation of the Kelvin-Helmholtz-like instabilities that can be triggered at the core/SOL transition, the sheath acceleration at the limiter or divertor plates leading to a radial shear for the velocity. Linear stability analysis developed in Schwander et al [2] actually reveals that unstable modes at the edge-core transition can develop in the presence of core rotation. This study was also an opportunity for a benchmarking comparison with the TOKAM3X code developed in IRFM and at Marseille. Both codes have confirmed the linear stability analysis and have shown that large fluctuations grow in the shear layer downstream of the limiter on the LFS (Low Field Side), the growth of these fluctuations being accompanied by a radial drift away from the limiter. Fig. 2 displays a 3D representation of these unstable

[2] Schwander et al., J. Nucl. Materials, doi:10.1016/j.jnucmat.2010.10.073 2011

density fluctuations. Apart from the physical results, this study has also shown the large sensitivity of the solutions to the discretization and to the implementation of the Bohm's boundary conditions. It has also shown that the study of these parallel KH like instabilities is a very demanding benchmark : these simulations require large meshes and since the growth rate of these instabilities is very weak, this results in an extremely long simulation time involving an extremely large number of time steps. In the future, it is planned to investigate the saturation of the instability as well as its possible presence in diverted plasmas.

../../../../projets/castor/IMG/parallelKHinstabilities.png

Figure 2. 3D representation of unstable density fluctuations in edge plasma tokamak with limiter. The core plasma rotates in the anti-clockwise direction M//central=0.75, safety factor q=6

6.5. Development of a two temperature model

Participants: Hervé Guillard, Afeintou Sangam, Elise Estibals.

A two temperature (ions - electrons) version of the code is in development. At present an approximate Riemann solver using the total energy equation and the electron entropy as main variables has been designed. This Riemann solver has been validated against standard shock tube problems and incorporated in the PlaTo platform. The implementation of this solver in toroidal geometry is in progress.

6.6. Implementation of a Taylor-Galerkin stabilized Finite Element

Participants: José Costa, Marie Martin, Boniface Nkonga.

The theoretical part of Taylor-Galerkin (TG) stabilized strategy applied to MHD and reduced MHD modeling has been realized. The final method amounts to add in the formulation, a self-adjoint operator associated to the most critical hyperbolic component of the system to be solved. The design of the critical contours and the identification of associated waves to be stabilized is problem dependent and related to the Jacobian matrix. A simplified version has been developed for reduced MHD and takes into account the high anisotropy of strongly magnetized plasma under consideration here. This first implementation of the TG stabilization in Jorek, has made possible efficient and robust simulations of Edge Localized Modes (Elms) and their mitigation by Resonant Magnetic Perturbations (RMP) and pellets injections. Work under progress will use more elaborated TG formulations that will be applied to reduced and full MHD models.

6.7. Development of a full MHD Modeling

Participants: José Costa, Jeaniffer Vides, Boniface Nkonga.

The single fluid full MHD numerical model has been developed. The divergence free constraint on the magnetic field is achieved by introduction of a potential vector. The use of the potential vector has the additional advantage that the toroidal component is the magnetic flux of the Grad-Shafranov equilibrium. Therefore, using the same finite element for the computation of initial equilibrium and the evolution of perturbed system, the numerical scheme is well balanced when the projection of the momentum equation use a component parallel to the magnetic field. Indeed, at the discrete level the projection is exactly orthogonal to equilibrium sub-space. Using the potential vector as variable introduces third order derivatives in the system and classical C0 finite elements cannot be directly applied. This is why our finite element strategy uses shape/test functions whose derivatives have global continuity in space. Finite element method is designed for poloidal plane discretization using quadrangles or triangles. Validations have been performed for internal kink and tearing modes instabilities in tokamak with a circular plasma. For this configuration, all magnetic surfaces are closed and simple boundary conditions are used. Future work will address X-point configurations with Bohm boundary conditions.

6.8. Environmental and Astrophysical flows

Participants: Hervé Guillard, Boniface Nkonga, Marco Bilanceri, Maria-Vittoria Salvetti [University of Pisa, Italy], Karim Elhakim [University Ains Shams, Egypt].

The numerical approximation of a model coupling the shallow-water equations with a sediment transport equation for the morphodynamics has been studied. In shallow-water problems, time advancing can be carried out by explicit schemes. However, if the interaction with the mobile bed is weak, the characteristic time scales of the flow and of the sediment transport can be very different introducing time stiffness in the global problem. For this case, it is of great interest to use implicit schemes. The time integration stategy that we have devised is based on a defect-correction approach and on a time linearization, in which the flux Jacobians are computed through automatic differentiation. This work has been published in this reference [3]. The aim of the present work is to investigate the behavior of this time scheme for different Riemann solvers, sediment transport models and situations related to environmental flows [12]. This activity takes place in the framework of the Euromediterranée 3+3 MedLagoon program and a PHC Imhotep program. A preliminary work has begun to apply this strategy to the study of the Burulus lake in Egypt.

6.9. Ionospheric plasma

Participants: Didier Auroux, Sebastian Minjeaud.

In order to guarantee the integrity of the european positionning system Galileo, it is fundamental to identify all the potential sources of system unavailability. One of the main sources that has been identified is the phenomenon of ionospheric scintillations which causes radio frequency signal amplitude and phase variations when satellite signals pass through the ionosphere. Scintillations appear as the turbulent aspect of a larger disturbance of the ionospheric plasma density, which has the shape of a plasma bubble. In this context and in the framework of the ANR IODISSEE, a model hierarchy aimed at representing the evolution of the ionospheric plasma was proposed (Besse and al., 2004). It is based on an asymptotic analysis of the Euler-Maxwell system thanks to typical scales of the physical parameters involved in this framework. Among these models the simplest, referred to as the Striation model, describes the evolution of the quasineutral plasma in a plane perpendicular to the earth magnetic field. The magnetic field is assumed constant, and both electron and ions inertia are neglected. In this model the mobility of charge particles is assumed infinite along the magnetic field lines so that they constitute equipotential for the electric field. This property allows the computation of the electric field by means of a two dimensional elliptic equation with coefficients integrated along the magnetic field lines. This equation is coupled to either a two or a three dimensional transport equation for the evolution of the plasma density.

[3] Linearized implicit time advancing and defect correction applied to sediment transport simulations Marco Bilanceri; François Beux; Imad Elmahi; Hervé Guillard; Maria Vittoria Salvetti Computers and Fluids, elsevier, 2012, 63, pp. 82-104

6.9.1. Data assimilation

We worked on data-models coupling method to identify the parameters of the Striation model (especially, the initial data for the electronic density and the ion/neutral collision frequency). Some measurements acquired during the mission of satellite DEMETER will constitute the set of observed data. We consider this problem from an optimal control point of view. We define a cost function measuring the misfit between the observed data and the corresponding model states. This function can be seen as a function of only the input model parameters. The previous inverse problem is then equivalent to minimizing this cost function. Of course, this problem can be ill posed (over or under-determined, non convex...) and we need to add some regularization terms (mainly Tikhonov terms), using some a priori estimation of the model parameters. The effective minimization of the cost function (in order to estimate the best possible set of model parameters) relies on the computation of the adjoint state. Thanks to the help of L. Hascoet, the adjoint of the Striation code was derived and validated with the automatic differentiation tool TAPENADE. We obtained the first results of coefficients identification in very simple situations.

6.9.2. Wave Propagation

The electronic density fluctuations of the ionospheric plasma have been identified as the main causes of the scintillation phenomena since they induce some variations of the amplitudes and phases of the signals passing through the ionosphere. These fluctuations have indeed a direct influence on the refractive index of the medium. A code was developed to simulate the propagation of wave signals in the perturbed ionosphere (whose representation is obtained thanks to the Striation code). We chose to use the method of the Rytov approximations which allow to obtain sufficiently accurate results (since, e.g., it takes into account the diffraction due to the small structures) whithin a reasonable computational time (contrary to the resolution of the whole Maxwell system). This work was carried out during the Master 2 internship of Gonzalo José Carracedo Carballal (advised with P. Lafitte, Ecole Centrale Paris).

6.10. Mesh adaptative MG Methods

Participants: Gautier Brethes [Projet Ecuador], Alain Dervieux, Olivier Allain [Lemma].

Anisotropic tetrahedrization, Continuous metric

This activity concerns the use of mesh adaptation and multigrid for simplified plasma models in the context of ANEMOS ANR project.

6.11. Turbulence models

Participants: Alain Dervieux, Bruno Koobus [University of Montpellier 2], Carine Moussaed [University of Montpellier 2], Maria-Vittoria Salvetti [University of Pisa], Stephen Wornom [Lemma], Marianna Braza [IMF-Toulouse].

Large Eddy Simulation, Variational Multi-scale, hybrid models, unstructured meshes, vortex shedding

The purpose of our works in hybrid RANS/LES is to develop new approaches for industrial applications of LES-based analyses. This year, a lot of experiments have validated the dynamic version of our VMS-LES. The quality of simulations is either comparable to non-dynamic, or better. In the foreseen applications (aeronautics, hydraulics), the Reynolds number can be as high as several tenth millions, a far too large number for pure LES models. However, certain regions in the flow can be much better predicted with LES than with usual statistical RANS (Reynolds averaged Navier-Stokes) models. These are mainly vortical separated regions as assumed in one of the most popular hybrid model, the hybrid Detached Eddy Simulation model. Here, “hybrid” means that a blending is applied between LES and RANS. The french-italian team is working on a novel type of hybrid model between the VMS-LES model and a $k - \varepsilon$ one. The team has this year concentrated on the shift between the RANS boundary region and the VLES boundary one. This is also the problematic on the IDDES studies. We are working on propositions relying not only on the value of the RANS viscosity but also on its gradient. A paper is in preparation on this subject.

CELTIQUE Project-Team

5. New Results

5.1. Information Flow Tracking

Participants: Frédéric Besson, Nataliia Bielova, Delphine Demange, Thomas Jensen, David Pichardie.

We investigate different approaches for dynamically tracking information flows.

The first track of work is motivated by web-browser security. In a survey [15], we have classified JavaScript security policies and their enforcement mechanisms in a web-browser. We have identified the problem of stateless web tracking (fingerprinting) and have proposent a novel approach to hybrid information flow monitoring by tracking the knowledge about secret variables using logical formulae. A logic formula quantifies the amount of knowledge stored in a variable. This knowledge representation helps to compare and improve precision of hybrid information flow monitors. We define a generic hybrid monitor parametrised by a static analysis and derive sufficient conditions on the static analysis for soundness and relative precision of hybrid monitors. We instantiate the generic monitor with a combined static constant and dependency analysis. Several other hybrid monitors including those based on well-known hybrid techniques for information flow control are formalised as instances of our generic hybrid mon- itor. These monitors are organised into a hierarchy that establishes their relative precision. The whole framework is accompanied by a formalisation of the theory in the Coq proof assistant [19].

Our second activity is related to SAFE, a clean-slate design for a highly secure computer system, with pervasive mechanisms for tracking and limiting information flows. At the lowest level, the SAFE hardware supports fine-grained programmable tags, with efficient and flexible propagation and combination of tags as instructions are executed. The operating system virtualizes these generic facilities to present an information-flow abstract machine that allows user programs to label sensitive data with rich confidentiality policies. We present a formal, machine-checked model of the key hardware and software mechanisms used to control information flow in SAFE and an end-to-end proof of noninterference for this model in the Coq proof assistant [17].

5.2. Towards efficient abstract domains for regular language based static analysis

Participants: Thomas Genet, Valérie Murat, Yann Salmon.

We develop a specific theory and the related tools for analyzing programs whose semantics is defined using term rewriting systems. The analysis principle is based on regular approximations of infinite sets of terms reachable by rewriting. The tools we develop use, so-called, Tree Automata Completion to compute a tree automaton recognizing a superset of all reachable terms. This over-approximation is then used to prove safety properties on the program by showing that some "bad" terms, encoding dangerous or problematic configurations, are not in the superset and thus not reachable. This is a specific form of, so-called, Regular Tree Model Checking. However, when dealing with infinite-state systems, Regular Tree Model Checking approaches may have some difficulties to represent infinite sets of data. We proposed Lattice Tree Automata, an extended version of tree automata to represent complex data domains and their related operations in an efficient manner. Moreover, we introduce a new completion-based algorithm for computing the possibly infinite set of reachable states in a finite amount of time. This algorithm is independent of the lattice making it possible to seamlessly plug abstract domains into a Regular Tree Model Checking algorithm[27]. As a first instance, we implemented in Timbuk a completion with an interval abstract domain. We shown that this implementation permits to scale up regular tree model-checking of Java programs dealing with integer arithmetics. Now, we aim at applying this technique to the static analysis of programming languages whose semantics is based on terms, like functional programming languages [38].

5.3. Result Certification of Static Program Analysers with Automated Theorem Provers

Participants: Frédéric Besson, Pierre-Emmanuel Cornilleau, Thomas Jensen.

The automation of the deductive approach to program verification crucially depends on the ability to efficiently infer and discharge program invariants. In an ideal world, user-provided invariants would be strengthened by incorporating the result of static analysers as untrusted annotations and discharged by automated theorem provers. However, the results of object-oriented analyses are heavily quantified and cannot be discharged, within reasonable time limits, by state-of-the-art auto- mated theorem provers. In the present work, we investigate an original approach for verifying automatically and efficiently the result of certain classes of object-oriented static analyses using off-the-shelf automated theorem provers. We propose to generate verification conditions that are generic enough to capture, not a single, but a family of analyses which encompasses Java bytecode verification and Fähndrich and Leino type-system for checking null pointers. For those analyses, we show how to generate tractable verification conditions that are still quantified but fall in a decidable logic fragment that is reducible to the Effectively Propositional logic. Our experiments confirm that such verification conditions are efficiently discharged by off-the-shelf automated theorem provers [20].

5.4. Formal Semantics for Multi-threaded Java

Participants: Delphine Demange, Vincent Laporte, David Pichardie.

Recent advances in verification have made it possible to envision trusted implementations of real-world languages. Java with its type-safety and fully specified semantics would appear to be an ideal candidate; yet, the complexity of the translation steps used in production virtual machines have made it a challenging target for verifying compiler technology. One of Java's key innovations, its memory model, poses significant obstacles to such an endeavor. The Java Memory Model is an ambitious attempt at specifying the behavior of multithreaded programs in a portable, hardware agnostic, way. While experts have an intuitive grasp of the properties that the model should enjoy, the specification is complex and not well-suited for integration within a verifying compiler infrastructure. Moreover, the specification is given in an axiomatic style that is distant from the intuitive reordering-based reasonings traditionally used to justify or rule out behaviors, and ill suited to the kind of operational reasoning one would expect to employ in a compiler. We take a step back, and introduces a *Buffered Memory Model* (BMM) for Java [26]. We choose a pragmatic point in the design space sacrificing generality in favor of a model that is fully characterized in terms of the reorderings it allows, amenable to formal reasoning, and which can be efficiently applied to a specific hardware family, namely x86 multiprocessors. Although the BMM restricts the reorderings compilers are allowed to perform, it serves as the key enabling device to achieving a verification pathway from bytecode to machine instructions. Despite its restrictions, we show that it is backwards compatible with the Java Memory Model and that it does not cripple performance on TSO architectures.

5.5. Formal Verification of Static Analysis

Participants: Sandrine Blazy, Martin Bodin, Thomas Jensen, Vincent Laporte, André Oliveira Maroneze, David Pichardie, Alan Schmitt.

Static analyzers based on abstract interpretation are complex pieces of software implementing delicate algorithms. Even if static analysis techniques are well understood, their implementation on real languages is still error-prone.

Using the Coq proof assistant, we formalized of a value analysis (based on abstract interpretation), and a soundness proof of the value analysis. The formalization relies on generic interfaces. The mechanized proof is facilitated by a translation validation of a Bourdoncle fixpoint iterator. The work has been integrated into the CompCert verified C-compiler. Our verified analysis directly operates over an intermediate language of the compiler having the same expressiveness as C. The automatic extraction of our value analysis into OCaml yields a program with competitive results, obtained from experiments on a number of benchmarks and comparisons with the Frama-C tool [21]. The value analysis was applied to a loop bound estimation tool for WCET analysis [22] relying also on program slicing and loop bound calculation.

Moreover, we formalized static analyses for logic programming, relying on results about the relative correctness of semantics in different styles; forward and backward, top-down and bottom-up. The results chosen are paradigmatic of the kind of correctness theorems that semantic analyses rely on and are therefore well-suited to explore the possibilities afforded by the application of interactive theorem provers to this task, as well as the difficulties likely to be encountered in the endeavour [29].

We also study the development of certified information flow analyses based on a formal semantics of JavaScript. We have in particular presented a technique for deriving semantic program analyses from a natural semantics specification of the programming language. The technique is based on the pretty-big-step semantics approach applied to a language with simple objects called O'While. We have specified a series of instrumentations of the semantics that makes explicit the flows of values in a program. This leads to a semantics-based dependency analysis, at the core, e.g., of tainting or information flow analyses in software security [32].

5.6. Certified JavaScript Semantics

Participants: Martin Bodin, Alan Schmitt.

JavaScript is the most widely used web language for client-side applications. Whilst the development of JavaScript was initially just led by implementation, there is now increasing momentum behind the ECMA standardisation process. The time is ripe for a formal, mechanised specification of JavaScript, to clarify ambiguities in the ECMA standards, to serve as a trusted reference for high-level language compilation and JavaScript implementations, and to provide a platform for high-assurance proofs of language properties. We present JScert, a formalisation of the current ECMA standard in the Coq proof assistant, and JSref, a reference interpreter for JavaScript extracted from Coq to OCaml. We give a Coq proof that JSref is correct with respect to JScert and assess JSref using test262, the ECMA conformance test suite. Our methodology ensures that JScert is a comparatively accurate formulation of the English standard, which will only improve as time goes on. We have demonstrated that modern techniques of mechanised specification can handle the complexity of JavaScript [25], [24].

5.7. Concurrent Reversibility

Participant: Alan Schmitt.

Concurrent reversibility has been studied in different areas, such as biological or dependable distributed systems. However, only "rigid" reversibility has been considered, allowing to go back to a past state and restart the exact same computation, possibly leading to divergence. We present a concurrent calculus featuring *flexible reversibility*, allowing the specification of alternatives to a computation to be used upon rollback. Alternatives in processes of this calculus are attached to messages. We show the robustness of this mechanism by encoding more complex idioms for specifying flexible reversibility, and we illustrate the benefits of our approach by encoding a calculus of communicating transactions [30].

5.8. Non linear analysis: fast inference of polynomial invariants

Participants: Thomas Jensen, David Cachera, Arnaud Jobin.

We have proposed an abstract interpretation based method for inferring polynomial invariants.Our analysis uses a form of weakest precondition calculus which was already observed to be well adapted to polynomial disequality guards, and which we extend to equality guards by using parameterized polynomial division. We have shown that the choice of a suitable division operation is crucial at each iteration step in order to compute the invariant. Based on this analysis, we have designed a constraint-based algorithm for inferring polynomial invariants. We have identified heuristics to solve equality constraints between ideals, and implemented the whole analysis algorithm in Maple.A salient feature of this analysis, which distinguishes it from the approaches proposed so far in the literature, is that it does not require the use of Gröbner base computations, which are known to be costly on parameterized polynomials. Our benchmarks show that our analyzer can successfully infer invariants on a sizeable set of examples, while performing two orders of magnitude faster than other existing implementations [16].

CEPAGE Project-Team

6. New Results

6.1. Resource allocation and Scheduling

6.1.1. Broadcasting on Large Scale Heterogeneous Platforms under the Bounded Multi-Port Model

Participants: Olivier Beaumont, Nicolas Bonichon, Lionel Eyraud-Dubois, Przemyslaw Uznanski.

In [17], we consider the problem of broadcasting a large message in a large scale distributed network under the multi-port communication model. We are interested in building an overlay network, with the aim of maximizing the throughput and minimizing the degree of the participating nodes. We consider a classification of participating nodes into two parts: open nodes that stay in the open-Internet and "guarded" nodes that lie behind firewalls or NATs, with the constraint that two guarded nodes cannot communicate directly. Without guarded nodes, we prove that it is possible to reach the optimal throughput with a quasi-optimal (up to a small additive increase) degree of the participating nodes. In presence of guarded nodes, we provide a closed form formula for the optimal cyclic throughput and we observe that the optimal solution may require arbitrarily large degrees. In the acyclic case, we propose an algorithm that reaches the optimal acyclic throughput with low degree. Then, we prove a worst case 5/7 ratio between the optimal acyclic and cyclic throughput and show through simulations that this ratio is on average very close to 1, what makes acyclic solutions efficient both in terms of throughput maximization and degree minimization.

6.1.2. Non Linear Divisible Load Scheduling

Participants: Olivier Beaumont, Hubert Larchevêque.

Divisible Load Theory (DLT) has received a lot of attention in the past decade. A divisible load is a perfect parallel task, that can be split arbitrarily and executed in parallel on a set of possibly heterogeneous resources. The success of DLT is strongly related to the existence of many optimal resource allocation and scheduling algorithms, what strongly differs from general scheduling theory. Moreover, recently, close relationships have been underlined between DLT, that provides a fruitful theoretical framework for scheduling jobs on heterogeneous platforms, and MapReduce, that provides a simple and efficient programming framework to deploy applications on large scale distributed platforms. The success of both have suggested to extend their framework to non-linear complexity tasks. In [32], we show that both DLT and MapReduce are better suited to workloads with linear complexity. In particular, we prove that divisible load theory cannot directly be applied to quadratic workloads, such as it has been proposed recently. We precisely state the limits for classical DLT studies and we review and propose solutions based on a careful preparation of the dataset and clever data partitioning algorithms. In particular, through simulations, we show the possible impact of this approach on the volume of communications generated by MapReduce, in the context of Matrix Multiplication and Outer Product algorithms. (Joint work with Loris Marchal from ENS Lyon)

6.1.3. Reliable Service Allocation in Clouds

Participants: Olivier Beaumont, Lionel Eyraud-Dubois, Hubert Larchevêque, Paul Renaud-Goud, Philippe Duchon.

In [30], we consider several reliability problems that arise when allocating applications to processing resources in a Cloud computing platform. More specifically, we assume on the one hand that each computing resource is associated to a capacity constraint and to a probability of failure. On the other hand, we assume that each service runs as a set of independent instances of identical Virtual Machines, and that the Service Level Agreement between the Cloud provider and the client states that a minimal number of instances of the service should run with a given probability. In this context, given the capacity and failure probabilities of the machines, and the capacity and reliability demands of the services, the question for the cloud provider is to find an allocation of the instances of the services (possibly using replication) onto machines satisfying all types of constraints during a given time period. The goal of this work is to assess the impact of the reliability constraint on the complexity of resource allocation problems. We consider several variants of this problem, depending on the number of services and whether their reliability demand is individual or global. We prove several fundamental complexity results (#P' and NP-completeness results) and we provide several optimal and approximation algorithms. In particular, we prove that a basic randomized allocation algorithm, that is easy to implement, provides optimal or quasi-optimal results in several contexts, and we show through simulations that it also achieves very good results in more general settings.

In [29], we extend this work to an energy minimisation framework, by considering two energy consumption models based on DVFS techniques, where the clock frequency of physical resources can be changed with a Dynamic Voltage and Frequency Scaling (DVFS) method. For each allocation problem and each energy model, we prove deterministic approximation ratios on the consumed energy for algorithms that provide guaranteed probability failures, as well as an efficient heuristic, whose energy ratio is not guaranteed.

In [37], we study the robustness of an allocation of Virtual Machines (VM) on a set of Physical Machines (PM) when the resource demand of the VMs can change over time. This may imply sometimes expensive "SLA violations", corresponding to some VM's consumption not satisfied because of overloaded PMs. Thus, while optimizing the global resource utilization of the PMs, it is necessary to ensure that at any moment a VM's need evolves, a few number of migrations (moving a VM from PM to PM) is sufficient to find a new configuration in which all the VMs' consumptions are satisfied. We modelize this problem using a fully dynamic bin packing approach and we present an algorithm ensuring a global utilization of the resources of 66%. Moreover, each time a PM is overloaded at most one migration is necessary to fall back in a configuration with no overloaded PM, and only 3 different PMs are concerned by required migrations that may occur to keep the global resource utilization correct. This allows the platform to be highly resilient to a great number of changes.

6.1.4. Splittable Single Source-Sink Routing on CMP Grids: A Sublinear Number of Paths Suffice

Participants: Adrian Kosowski, Przemyslaw Uznanski.

In [44], we study single chip multiprocessors (CMP) with grid topologies, where a significant part of power consumption is attributed to communications between the cores of the grid. We investigate the problem of routing communications between CMP cores using shortest paths, in a model in which the power cost associated with activating a communication link at a transmission speed of f bytes/second is proportional to f^α, for some constant exponent $\alpha > 2$. Our main result is a trade-off showing how the power required for communication in CMP grids depends on the ability to split communication requests between a given pair of node, routing each such request along multiple paths. For a pair of cores in a $m \times n$ grid, the number of available communication paths between them grows exponentially with n, m. By contrast, we show that optimal power consumption (up to constant factors) can be achieved by splitting each communication request into k paths, starting from a threshold value of $k = \Theta(n^{1/(\alpha-1)})$. This threshold is much smaller than n for typical values of $\alpha \approx 3$, and may be considered practically feasible for use in routing schemes on the grid. More generally, we provide efficient algorithms for routing multiple k-splittable communication requests between two cores in the grid, providing solutions within a constant approximation of the optimum cost. We support our results with algorithm simulations, showing that for practical instances, our approach using k-splittable requests leads to a power cost close to that of the optimal solution with arbitrarily splittable requests, starting from the stated threshold value of k.

6.1.5. Maximum matching in multi-interface networks

Participants: Adrian Kosowski, Dominik Pajak.

In [26], we consider the standard matching problem in the context of multi-interface wireless networks. In heterogeneous networks, devices can communicate by means of multiple wireless interfaces. By choosing which interfaces to switch on at each device, several connections might be established. That is, the devices at the endpoints of each connection share at least one active interface. In the studied problem, the aim is to maximize the number of parallel connections without incurring in interferences. Given a network $G = (V, E)$, nodes V represent the devices, edges E represent the connections that can be established. If node x participates in the communication with one of its neighbors by means of interface i, then another neighboring node of x can establish a connection (but not with x) only if it makes use of interface $j \neq i$. The size of a solution for an instance of the outcoming matching problem, that we call *Maximum Matching in Multi-Interface* networks (MMMI for short), is always in between the sizes of the solutions for the same instance with respect to the standard matching and its induced version problems. However, we prove that MMMI is NP-hard even for proper interval graphs and for bipartite graphs of maximum degree $\Delta \geq 3$. We also show polynomially solvable cases of MMMI with respect to different assumptions.

6.1.6. Parallel scheduling of task trees with limited memory

Participant: Lionel Eyraud-Dubois.

In a paper submitted to ACM TOPC, we have investigated the execution of tree-shaped task graphs using multiple processors. Each edge of such a tree represents some large data. A task can only be executed if all input and output data fit into memory, and a data can only be removed from memory after the completion of the task that uses it as an input data. Such trees arise, for instance, in the multifrontal method of sparse matrix factorization. The peak memory needed for the processing of the entire tree depends on the execution order of the tasks. With one processor the objective of the tree traversal is to minimize the required memory. This problem was well studied and optimal polynomial algorithms were proposed. We have extended the problem by considering multiple processors, which is of obvious interest in the application area of matrix factorization. With multiple processors comes the additional objective to minimize the time needed to traverse the tree, i.e., to minimize the makespan. Not surprisingly, this problem proves to be much harder than the sequential one. We study the computational complexity of this problem and provide inapproximability results even for unit weight trees. We design a series of practical heuristics achieving different trade-offs between the minimization of peak memory usage and makespan. Some of these heuristics are able to process a tree while keeping the memory usage under a given memory limit. The different heuristics are evaluated in an extensive experimental evaluation using realistic trees.

6.1.7. Point-to-point and congestion bandwidth estimation: experimental evaluation on PlanetLab

Participants: Lionel Eyraud-Dubois, Przemyslaw Uznanski.

In large scale Internet platforms, measuring the available bandwidth between nodes of the platform is difficult and costly. However, having access to this information allows to design clever algorithms to optimize resource usage for some collective communications, like broadcasting a message or organizing master/slave computations. In [54], we analyze the feasibility to provide estimations, based on a limited number of measurements, for the point-to-point available bandwidth values, and for the congestion which happens when several communications take place at the same time. We present a dataset obtained with both types of measurements performed on a set of nodes from the PlanetLab platform. We show that matrix factorization techniques are quite efficient at predicting point-to-point available bandwidth, but are not adapted for congestion analysis. However, a LastMile modeling of the platform allows to perform congestion predictions with a reasonable level of accuracy, even with a small amount of information, despite the variability of the measured platform.

6.1.8. Parallel Mining of Functional Dependencies

Participants: Sofian Maabout, Nicolas Hanusse.

The problem of extracting functional dependencies (FDs) from databases has a long history dating back to the 90's. Still, efficient solutions taking into account both material evolution, namely the advent of multicore machines, and the amount of data that are to be mined, are still needed. In [46] we propose a parallel algorithm which, upon small modifications, extracts (i) the minimal keys, (ii) the minimal exact FDs, (iii) the minimal approximate FDs and (iv) the Conditional functional dependencies (CFDs) holding in a table. Under some natural conditions, we prove a theoretical speed up of our solution with respect to a baseline algorithm which follows a depth first search strategy. Since mining most of these dependencies require a procedure for computing the number of distinct values (NDV) which is a space consuming operation, we show how sketching techniques for estimating the exact value of NDV can be used for reducing both memory consumption as well as communications overhead when considering distributed data while guaranteeing a certain quality of the result. Our solution is implemented in both shared, using C++ and OpenMP, and distributed memory, using Hadoop implementation of Map-Reduce. The experimental results showthe efficiency and scalability of our proposal. Most notably, the theoretical speed ups are confirmed by the experiments.

6.1.9. Fast Skyline Query Evaluation with Skycuboids Materialization based on Functional Dependencies

Participants: Sofian Maabout, Nicolas Hanusse.

Ranking multidimensional data via different Skyline queries gives rise to the so called skycube structure. Most of previous work on optimizing the subspaces skyline queries have concentrated on full materialization of the skycube. Due to the exponential number of skylines one must pre-compute, the full materialization is unfeasible in practice. However, due to the non monotonic nature of skylines, there is no immediate inclusion relationship between the skycuboids when we have an inclusion of the dimensions. This makes the partial materialization harder. In this paper, we identify sufficient conditions for establishing inclusions between skycuboids thanks to the functional dependencies that hold in the underlying data. This leads to the characterization of a *minimal* set of skycuboids to be materialized in order to answer all the possible skyline queries without resorting to the underlying data. We conduct an extensive set of experiments showing that with the help of a small fraction of the skycube, we can efficiently answer all the possible skyline queries. In addition, our proposal turns to be helpful even in the full materialization setting. Indeed, thanks to the inclusions we identify, we devise a full materialization algorithm which outperforms state of the art skycube computation algorithms especially when data and dimensions get large. The results are reported in the technical report submitted to SIGMOD'14.

6.2. Compact Routing

6.2.1. On the Communication Complexity of Distributed Name-Independent Routing Schemes

Participants: Cyril Gavoille, Nicolas Hanusse, David Ilcinkas.

In [38], we present a distributed asynchronous algorithm that, for every undirected weighted n-node graph G, constructs name-independent routing tables for G. The size of each table is $\widetilde{O}(\sqrt{n}\,)$, whereas the length of any route is stretched by a factor of at most 7 w.r.t. the shortest path. At any step, the memory space of each node is $\widetilde{O}(\sqrt{n}\,)$. The algorithm terminates in time $O(D)$, where D is the hop-diameter of G. In synchronous scenarios and with uniform weights, it consumes $\widetilde{O}(m\sqrt{n} + n^{3/2}\min D, \sqrt{n}\,)$ messages, where m is the number of edges of G.

In the realistic case of sparse networks of poly-logarithmic diameter, the communication complexity of our scheme, that is $\widetilde{O}(n^{3/2})$, improves by a factor of $\sqrt{n}$ the communication complexity of *any* shortest-path routing scheme on the same family of networks. This factor is provable thanks to a new lower bound of independent interest.

6.2.2. There are Plane Spanners of Maximum Degree 4

Participant: Nicolas Bonichon.

Let $\mathcal{E}$ be the complete Euclidean graph on a set of points embedded in the plane. Given a fixed constant $t \geq 1$, a spanning subgraph G of $\mathcal{E}$ is said to be a *t-spanner* of $\mathcal{E}$ if for any pair of vertices u, v in $\mathcal{E}$ the distance between u and v in G is at most t times their distance in $\mathcal{E}$. A spanner is *plane* if its edges do not cross.

We consider the question: "What is the smallest *maximum degree* that can be achieved for a *plane* spanner of $\mathcal{E}$?" Without the planarity constraint, it is known that the answer is 3 which is thus the best known lower bound on the degree of any plane spanner. With the planarity requirement, the best known upper bound on the maximum degree is 6, the last in a long sequence of results improving the upper bound. In this work we show that there is a constant $t \geq 1$ such that the complete Euclidean graph always contains a plane t-spanner of maximum degree 4 and make a big step toward closing the question. Our construction leads to an efficient algorithm for obtaining the spanner from Chew's L_1-Delaunay triangulation.

6.3. Mobile Agents

6.3.1. Collision-Free Network Exploration

Participants: Ralf Klasing, Adrian Kosowski, Dominik Pajak.

A set of mobile agents is placed at different nodes of a n-node network. The agents synchronously move along the network edges in a *collision-free* way, i.e., in no round may two agents occupy the same node. In each round, an agent may choose to stay at its currently occupied node or to move to one of its neighbors. An agent has no knowledge of the number and initial positions of other agents. We are looking for the shortest possible time required to complete the collision-free *network exploration*, i.e., to reach a configuration in which each agent is guaranteed to have visited all network nodes and has returned to its starting location. In [34], we first consider the scenario when each mobile agent knows the map of the network, as well as its own initial position. We establish a connection between the number of rounds required for collision-free exploration and the degree of the minimum-degree spanning tree of the graph. We provide tight (up to a constant factor) lower and upper bounds on the collision-free exploration time in general graphs, and the exact value of this parameter for trees. For our second scenario, in which the network is unknown to the agents, we propose collision-free exploration strategies running in $O(n^2)$ rounds for tree networks and in $O(n^5 \log n)$ rounds for general networks.

6.3.2. Deterministic Rendezvous of Asynchronous Bounded-Memory Agents in Polygonal Terrains

Participant: Adrian Kosowski.

In [22], we deal with a more geometric variant of the rendezvous problem. Two mobile agents, modeled as points starting at different locations of an unknown terrain, have to meet. The terrain is a polygon with polygonal holes. We consider two versions of this rendezvous problem: exact RV, when the points representing the agents have to coincide at some time, and ϵ-RV, when these points have to get at distance less than ϵ in the terrain. In any terrain, each agent chooses its trajectory, but the movements of the agent on this trajectory are controlled by an adversary that may, e.g., speed up or slow down the agent. Agents have bounded memory: their computational power is that of finite state machines. Our aim is to compare the feasibility of exact and of ϵ-RV when agents are anonymous vs. when they are labeled. We show classes of polygonal terrains which distinguish all the studied scenarios from the point of view of feasibility of rendezvous. The features which influence the feasibility of rendezvous include symmetries present in the terrains, boundedness of their diameter, and the number of vertices of polygons in the terrains.

6.3.3. Optimal Patrolling of Fragmented Boundaries

Participant: Adrian Kosowski.

Mobile agents in geometric scenarios are also studied in [33], where a set of mobile robots is deployed on a simple curve of finite length, composed of a finite set of vital segments separated by neutral segments. The robots have to patrol the vital segments by perpetually moving on the curve, without exceeding their maximum speed. The quality of patrolling is measured by the idleness, i.e., the longest time period during which any vital point on the curve is not visited by any robot. Given a configuration of vital segments, our goal is to provide algorithms describing the movement of the robots along the curve so as to minimize the idleness. Our main contribution is a proof that the optimal solution to the patrolling problem is attained either by the cyclic strategy, in which all the robots move in one direction around the curve, or by the partition strategy, in which the curve is partitioned into sections which are patrolled separately by individual robots. These two fundamental types of strategies were studied in the past in the robotics community in different theoretical and experimental settings. However, to our knowledge, this is the first theoretical analysis proving optimality in such a general scenario.

6.3.4. Fast Collaborative Graph Exploration

Participants: Adrian Kosowski, Dominik Pajak, Przemyslaw Uznanski.

In [35], we study the following scenario of online graph exploration. A team of k agents is initially located at a distinguished vertex r of an undirected graph. At every time step, each agent can traverse an edge of the graph. All vertices have unique identifiers, and upon entering a vertex, an agent obtains the list of identifiers of all its neighbors. We ask how many time steps are required to complete exploration, i.e., to make sure that every vertex has been visited by some agent. We consider two communication models: one in which all agents have global knowledge of the state of the exploration, and one in which agents may only exchange information when simultaneously located at the same vertex. As our main result, we provide the first strategy which performs exploration of a graph with n vertices at a distance of at most D from r in time $O(D)$, using a team of agents of polynomial size $k = Dn^{1+\epsilon} < n^{2+\epsilon}$, for any $\epsilon > 0$. Our strategy works in the local communication model, without knowledge of global parameters such as n or D. We also obtain almost-tight bounds on the asymptotic relation between exploration time and team size, for large k. For any constant $c > 1$, we show that in the global communication model, a team of $k = Dn^c$ agents can always complete exploration in $D(1 + \frac{1}{c-1} + o(1))$ time steps, whereas at least $D(1 + \frac{1}{c} - o(1))$ steps are sometimes required. In the local communication model, $D(1 + \frac{2}{c-1} + o(1))$ steps always suffice to complete exploration, and at least $D(1 + \frac{2}{c} - o(1))$ steps are sometimes required. This shows a clear separation between the global and local communication models.

6.3.5. A $\widetilde{O}(n2)$ Time-Space Trade-off for Undirected s-t Connectivity

Participant: Adrian Kosowski.

The work [43] makes use of the Metropolis-type walks due to Nonaka et al. (2010) to provide a faster solution to the S-T-connectivity problem in undirected graphs (USTCON). As the main result of this research, we propose a family of randomized algorithms for USTCON which achieves a time-space product of $S \cdot T = \widetilde{O}(n^2)$ in graphs with n nodes and m edges (where the $\widetilde{O}$-notation disregards poly-logarithmic terms). This improves the previously best trade-off of $\widetilde{O}(nm)$, due to Feige (1995). Our algorithm consists in deploying several short Metropolis-type walks, starting from landmark nodes distributed using the scheme of Broder et al. (1994) on a modified input graph. In particular, we obtain an algorithm running in time $\widetilde{O}(n + m)$ which is, in general, more space-efficient than both BFS and DFS. Finally, we show how to fine-tune the Metropolis-type walk so as to match the performance parameters (e.g., average hitting time) of the unbiased random walk for any graph, while preserving a worst-case bound of $\widetilde{O}(n^2)$ on cover time.

6.3.6. The multi-agent rotor-router on the ring: a deterministic alternative to parallel random walks

Participants: Ralf Klasing, Adrian Kosowski, Dominik Pajak.

The *rotor-router mechanism* was introduced as a deterministic alternative to the random walk in undirected graphs. In this model, an agent is initially placed at one of the nodes of the graph. Each node maintains a cyclic ordering of its outgoing arcs, and during successive visits of the agent, propagates it along arcs chosen according to this ordering in round-robin fashion. In [42], we consider the setting in which multiple, indistinguishable agents are deployed in parallel in the nodes of the graph, and move around the graph in synchronous rounds, interacting with a single rotor-router system. We propose new techniques which allow us to perform a theoretical analysis of the multi-agent rotor-router model, and to compare it to the scenario of parallel independent random walks in a graph. Our main results concern the n-node ring, and suggest a strong similarity between the performance characteristics of this deterministic model and random walks.

We show that on the ring the rotor-router with k agents admits a cover time of between $\Theta(n^2/k^2)$ in the best case and $\Theta(n^2/\log k)$ in the worst case, depending on the initial locations of the agents, and that both these bounds are tight. The corresponding expected value of cover time for k random walks, depending on the initial locations of the walkers, is proven to belong to a similar range, namely between $\Theta(n^2/(k^2/\log^2 k))$ and $\Theta(n^2/\log k)$.

Finally, we study the limit behavior of the rotor-router system. We show that, once the rotor-router system has stabilized, all the nodes of the ring are always visited by some agent every $\Theta(n/k)$ steps, regardless of how the system was initialized. This asymptotic bound corresponds to the expected time between successive visits to a node in the case of k random walks. All our results hold up to a polynomially large number of agents ($1 \leq k < n^{1/11}$).

6.3.7. Efficient Exploration of Anonymous Undirected Graphs

Participant: Ralf Klasing.

In [41], we consider the problem of exploring an anonymous undirected graph using an oblivious robot. The studied exploration strategies are designed so that the next edge in the robot's walk is chosen using only local information. We present some current developments in the area. In particular, we focus on recent work on *equitable strategies* and on the *multi-agent rotor-router*.

6.3.8. Gathering radio messages in the path

Participant: Ralf Klasing.

In [19], we address the problem of gathering information in one node (sink) of a radio network where interference constraints are present: when a node transmits, it produces interference in an area bigger than the area in which its message can actually be received. The network is modeled by a graph; a node is able to transmit one unit of information to the set of vertices at distance at most dt in the graph, but when doing so it generates interferences that do not allow nodes at distance up to di ($di \geq dt$) to listen to other transmissions. We are interested in finding a gathering protocol, that is an ordered sequence of rounds (each round consists of non-interfering simultaneous transmissions) such that $w(u)$ messages are transmitted from any node u to a fixed node called the sink. Our aim is to find a gathering protocol with the minimum number of rounds (called *gathering time*). In [19], we focus on the specific case where the network is a path with the sink at an end vertex of the path and where the traffic is unitary ($w(u) = 1$ for all u); indeed this simple case appears to be already very difficult. We first give a new lower bound and a protocol with a gathering time that differ only by a constant independent of the length of the path. Then we present a method to construct incremental protocols. An incremental protocol for the path on $n+1$ vertices is obtained from a protocol for n vertices by adding new rounds and new calls to some rounds but without changing the calls of the original rounds. We show that some of these incremental protocols are optimal for many values of dt and di (in particular when dt is prime). We conjecture that this incremental construction always gives optimal protocols. Finally, we derive an approximation algorithm when the sink is placed in an arbitrary vertex in the path.

6.3.9. Computing Without Communicating: Ring Exploration by Asynchronous Oblivious Robots

Participant: David Ilcinkas.

In [24], we consider the problem of exploring an anonymous unoriented ring by a team of k identical, oblivious, asynchronous mobile robots that can view the environment but cannot communicate. This weak scenario is standard when the spatial universe in which the robots operate is the two-dimensional plane, but (with one exception) has not been investigated before for networks. Our results imply that, although these weak capabilities of robots render the problem considerably more difficult, ring exploration by a small team of robots is still possible. We first show that, when k and n are not co-prime, the problem is not solvable in general, e.g., if k divides n there are initial placements of the robots for which gathering is impossible. We then prove that the problem is always solvable provided that n and k are co-prime, for $k \geq 17$, by giving an exploration algorithm that always terminates, starting from arbitrary initial configurations. Finally, we consider the minimum number $\rho(n)$ of robots that can explore a ring of size n. As a consequence of our positive result we show that $\rho(n)$ is $O(\log n)$. We additionally prove that $\Omega(\log n)$ robots are necessary for infinitely many n.

6.3.10. Worst-case optimal exploration of terrains with obstacles

Participant: David Ilcinkas.

A mobile robot represented by a point moving in the plane has to explore an unknown flat terrain with impassable obstacles. Both the terrain and the obstacles are modeled as arbitrary polygons. We consider two scenarios: the *unlimited vision*, when the robot situated at a point p of the terrain explores (sees) all points q of the terrain for which the segment pq belongs to the terrain, and the *limited vision*, when we require additionally that the distance between p and q is at most 1. All points of the terrain (except obstacles) have to be explored and the performance of an exploration algorithm, called its complexity, is measured by the length of the trajectory of the robot.

For unlimited vision we show in [21] an exploration algorithm with complexity $O(P + D\sqrt{k})$, where P is the total perimeter of the terrain (including perimeters of obstacles), D is the diameter of the convex hull of the terrain, and k is the number of obstacles. We do not assume knowledge of these parameters. We also prove a matching lower bound showing that the above complexity is optimal, even if the terrain is known to the robot. For limited vision we show exploration algorithms with complexity $O(P + A + \sqrt{Ak})$, where A is the area of the terrain (excluding obstacles). Our algorithms work either for arbitrary terrains (if one of the parameters A or k is known) or for c-fat terrains, where c is any constant (unknown to the robot) and no additional knowledge is assumed. (A terrain $\mathcal{T}$ with obstacles is c-fat if $R/r \leq c$, where R is the radius of the smallest disc containing $\mathcal{T}$ and r is the radius of the largest disc contained in $\mathcal{T}$.) We also prove a matching lower bound $\Omega(P + A + \sqrt{Ak})$ on the complexity of exploration for limited vision, even if the terrain is known to the robot.

6.3.11. Exploration of the T-Interval-Connected Dynamic Graphs: the Case of the Ring

Participants: David Ilcinkas, Ahmed Wade.

In [40], we study the T-interval-connected dynamic graphs from the point of view of the time necessary and sufficient for their exploration by a mobile entity (agent). A dynamic graph (more precisely, an evolving graph) is T-interval-connected ($T \geq 1$) if, for every window of T consecutive time steps, there exists a connected spanning subgraph that is stable (always present) during this period. This property of connection stability over time was introduced by Kuhn, Lynch and Oshman (STOC 2010). We focus on the case when the underlying graph is a ring of size n, and we show that the worst-case time complexity for the exploration problem is $2n - T - \Theta(1)$ time units if the agent knows the dynamics of the graph, and $n + \frac{n}{\max\{1, T-1\}}(\delta - 1) \pm \Theta(\delta)$ time units otherwise, where δ is the maximum time between two successive appearances of an edge.

6.3.12. Time vs. space trade-offs for rendezvous in trees

Participant: Adrian Kosowski.

In [23], we consider the rendezvous problem, in which two identical (anonymous) mobile agents start from arbitrary nodes of an unknown tree and have to meet at some node. Agents move in synchronous rounds: in each round an agent can either stay at the current node or move to one of its neighbors. We consider deterministic algorithms for this rendezvous task. We obtain a tight trade-off between the optimal time of completing rendezvous and the size of memory of the agents. For agents with k memory bits, we show that optimal rendezvous time is $\Theta(n + n^2/k)$ in n-node trees. More precisely, if $k \geq c \log n$, for some constant c, we design agents accomplishing rendezvous in arbitrary trees of size n (unknown to the agents) in time $O(n + n^2/k)$, starting with arbitrary delay. We also show that no pair of agents can accomplish rendezvous in time $o(n + n^2/k)$, even in the class of lines of known length and even with simultaneous start. Finally, we prove that at least logarithmic memory is necessary for rendezvous, even for agents starting simultaneously in a n-node line.

CIDRE Project-Team

6. New Results

6.1. Intrusion Detection

6.1.1. Intrusion Detection based on an Analysis of the Flow Control

In 2013, we continue to strengthen our research efforts around intrusion detection parameterized by a security policy.

In [33], we have proposed a language for specifying and composing fine-grained information flow policies. The language used a XML-syntax and has a formal semantic. BSPL enables to precisely specify the expected behavior of applications relatively to their sensitive pieces of information. More precisely it permits to specify where a piece of data owned by an application is allowed to disseminate: in which files or processes.

In [25], we have experimented the previous language (BSPL). We have developed a policy manager for android devices. The manager is able to check the consistency of a policy and to compose two consistent policies. We have also proposed a semi-automatic method for computing information flow policies of applications. We have thus computed some examples of policies and shown that these policies are rich enough to permit benign execution of an application without raising useless alerts and sufficiently restrictive to detect malicious actions induced by a malware.

In [40], we have proposed a new data-structure called System Flow Graph (or SFG in short) that offers a compact representation of how pieces of data flow inside a system. For a given application, the system flow graph describes its external behavior. We have shown that this new data structure suits to represent malware behavior and permits to give an early diagnostic in case of intrusion.

In [36] we have collaborated with Mathieu Jaume from Université de Paris 6 describes a formal framework to draw a correspondence between two types of policy definitions - policies that are defined by properties over states of a system and those that are described by properties over executions of a system.

In [34] and in C.Hauser's PhD desertion, we have extended previous work on kBlare (an IDS that detect illegal flows of information at the kernel level) so as to follow information flows at the network level. To that end, a set of nodes administrated by a single entity can be configured according to a distributed security policy expressed in terms of legal information flows. The different operating systems (kBlare) at each node cooperate by tagging each network packet with a tag that describes the information content of the payload. This way, it is possible to detect illegal information flow of information at the network level. This can be used to detect attacks against confidentiality or integrity of the overall system.

6.1.2. Terminating-Insensitive Non-Interference Verification based on an Information Flow Control

In 2010-2011, we started an informal collaboration with colleagues from CEA LIST laboratory. In 2012, this collaboration has turn into a reality by the funding of a PhD student (Mounir Assaf). This PhD thesis is about the verification of security properties of programs written in an imperative language with pointer aliasing (a subset of C language) by techniques borrowed from the domain of static analysis. One of the property of interest for the security field is called Terminating-Insensitive Non-Interference. Briefly speaking, when verified by a program, this property ensures that the content of any secret variable can not leak into public ones (for any terminating execution). However, this property is too strict in the sense that a large number of programs although perfectly secure are rejected by classical analyzers.

In 2013, Mounir Assaf has studied novel approaches that combine static and dynamic information flow monitoring. These approaches are promising since they enable permissive (accepting a large subset of executions) yet sound (rejecting all insecure executions) enforcement of non-interference. We have investigated a dynamic information flow monitor for a language supporting pointers. Our flow-sensitive monitor relies on prior static analysis in order to soundly enforce non-interference. We have also proposed a program transformation that preserves the behavior of initial programs and soundly inlines our security monitor. This program transformation enables both dynamic and static verification of non-interference in a language supporting pointers. This work has been published in [27] and [45].

6.1.3. Visualization of Security Events

The studies that were performed last year clearly showed that there was an important need for technologies that would allow analysts to handle in a consistent way the various types of log files that they have to study in order to detect intrusion or to perform forensic analysis. Consequently, we proposed this year ELVis, a security-oriented log visualization system that allows the analyst to import its log files and to obtain automatically a relevant representation of their content based on the type of the fields they are made of. First, a summary view is proposed. This summary displays in an adequate manner each field according to its type (i.e. categorical, ordinal, geographical, etc.). Then, the analyst can select one or more fields to obtain some details about it. A relevant representation is then automatically selected by the tool according to the types of the fields that were selected.

ELVis [35] has been presented in VizSec 2013 (part of Vis 2013) in October in Atlanta. A working prototype is currently being tuned in order to perform field trials with our partners in DGA-MI. Next year, we are planing to perform research on how various log files can be combined in the same representation. In the PANOPTESEC project, we will also perform some research on visualization for security monitoring in the context of SCADA systems.

6.2. Privacy

6.2.1. Geoprivacy

With the advent of GPS-equipped devices, a massive amount of location data is being collected, raising the issue of the privacy risks incurred by the individuals whose movements are recorded. In [31], we focus on a specific inference attack called the de-anonymization attack, by which an adversary tries to infer the identity of a particular individual behind a set of mobility traces. More specifically, we propose an implementation of this attack based on a mobility model called Mobility Markov Chain (MMC). A MMC is built out from the mobility traces observed during the training phase and is used to perform the attack during the testing phase. We design several distance metrics quantifying the closeness between two MMCs and combine these distances to build de-anonymizers that can re-identify users in an anonymized geolocated dataset. Experiments conducted on real datasets demonstrate that the attack is both accurate and resilient to sanitization mechanisms such as downsampling. This paper has received the IEEE best student paper award at the conference TrustCom 2013.

In [30], we propose to adopt the MapReduce paradigm in order to be able to perform a privacy analysis on large scale geolocated datasets composed of millions of mobility traces. More precisely, we design and implement a complete MapReduce-based approach to GEPETO. GEPETO (for GEoPrivacy-Enhancing TOolkit) is a flexible software that can be used to visualize, sanitize, perform inference attacks and measure the utility of a particular geolocated dataset. The main objective of GEPETO is to enable a data curator (e.g., a company, a governmental agency or a data protection authority) to design, tune, experiment and evaluate various sanitization algorithms and inference attacks as well as visualizing the following results and evaluating the resulting trade-off between privacy and utility. Most of the algorithms used to conduct an inference attack (such as sampling, k-means and DJ-Cluster) represent good candidates to be abstracted in the MapReduce formalism. These algorithms have been implemented with Hadoop and evaluated on a real dataset. Preliminary results show that the MapReduced versions of the algorithms can efficiently handle millions of mobility traces.

6.2.2. Privacy-enhanced Social Networks

In [38], we have proposed a systematic methodology for evaluating the quality of the privacy proposed by a social networking platform. It is based on an analysis grid organizing a correspondence between a number of design features and properties having an impact on privacy, and a level of distribution. For each property, we consider three possible distribution levels: centralized, decentralized and fully decentralized. For security properties, in particular, we have defined those distribution levels with the help of three different attacker models: an attacker has the ability to compromise either one entity in the system, a pre-defined subset of entites in the system, or the whole set of peers in the system. We argument on the idea that the more powerful the attacker model needed to compromise a property for all users in the system, the higher the privacy level linked to this property. A formal evaluation tool based on lattice structures is then proposed to compare social network systems based on this analysis grid. An example evaluation is also provided, with the thorough analysis of several well-known systems of various kinds, notably leading to the conclusion that some privacy-oriented social networking architectures, presented by their authors as fully distributed, showed centralized characteristics for many privacy-related properties.

6.2.3. Privacy Enhancing Technologies

The development of NFC-enabled smartphones has paved the way to new applications such as mobile payment (m-payment) and mobile ticketing (m-ticketing). However, often the privacy of users of such services is either not taken into account or based on simple pseudonyms, which does not offer strong privacy properties such as the unlinkability of transactions and minimal information leakage. In [26], we introduce a lightweight privacy-preserving contactless transport service that uses the SIM card as a secure element. Our implementation of this service uses a group signature protocol in which costly cryptographic operations are delegated to the mobile phone.

6.2.4. Privacy and Web Services

We have proposed [55] a new model of security policy based for a first part on our previous works in information flow policy and for a second part on a model of Myers and Liskov. This new model of information flow serves web services security and allows a user to precisely define where its own sensitive pieces of data are allowed to flow through the definition of an information flow policy. A novel feature of such policy is that they can be dynamically updated, which is fundamental in the context of web services that allow the dynamic discovery of services. We have also presented an implementation of this model in a web services orchestration in BPEL (Business Process Execution Language).

6.2.5. Privacy-preserving Ad-hoc Routing

6.2.5.1. Proactive Protocol

In [39], we have proposed a *proactive* ad hoc routing protocol that preserves the anonymity of the source and of the destination of the packet flows, and assures the unlinkability of flows between any pair of participants to local observers and to global attackers to a lesser extend. Our solution is based on OLSR and combines Bloom filters and ephemeral identifiers. More specifically, the routing process allows any node to discover the topology of the ad hoc network. Once such a topology is known, a source node can establish beforehand a path to reach any destination node. To conceal the identity of the source and destination nodes, the path may not be the shortest ones nor terminate at the destination node. Then, by including the ephemeral public identifiers of the intermediate nodes into a Bloom filter, the source node is able to specify the nodes that have to rebroadcast packets. Thus, when receiving a packet, a node has simply to check, using its ephemeral private identifier, whether it has to rebroadcast the packet, without knowing the source, the destination, nor the previous and next hop.

6.2.5.2. Reactive Protocol

In [42], we have proposed a classification of privacy preserving properties that ensure privacy in ad hoc network routing. We also proposed a taxonomy of adversary's model to analyse existing privacy preserving ad hoc routing protocols. To improve these protocols and to try address all privacy preserving properties, we proposed NoName [42], a novel privacy-preserving ad hoc routing protocol. Based on trapdoor, virtual switching and partially disjoint multipath using Bloom filter, NoName ensures anonytmity of the source, of the destination and of intermediate nodes. It also ensures unlinkability between source and message and between destination and message.

In [43], we have proposed another anonymous *proactive* ad hoc routing protocol, called APART, based on Gentry's fully homomorphic cryptography. Even though this technology is currently quite inefficient from a computational perspective, especially for an application in ad-hoc networks, the protocol APART is merely a proof of concept showing that an anonymous proactive protocol is possible thanks to it. The main idea is that each node maintains a routing table that contains only encrypted data. When a source node want to communicate with a destination node, it cooperates with its neighbors to discover the node that is the next hop to the destination node. This is done in such a way that the source node does not know the entry in its routing table that corresponds to the destination, and the next hop does only know that it has to rebroadcast the messages coming from that source.

6.2.6. Right to be forgotten

The right to be forgotten has become an investigation topic in itself within the field of privacy protection. In [46], we present the joint research project funded by the ministry of justice between our team and researchers in law and sociology, in order to examine the current state, in society and in technology, of the notion of a right to be forgotten, to identify the forthcoming computing tools capable of implementing the notion, and to evaluate the relevance of an autonomous legislation to define it and regulate it. In association with this study and in the light of the identified state-of-the-art, we have proposed in [47] a new technique to implement a right to be forgotten in the manner of a degradation of the quality of published data in time, associated with a fully distributed ephemeral publication technology. We show how this technique could fit various use cases in geosocial networks.

6.3. Trust

Digital reputation mechanisms have indeed emerged as a promising approach to cope with the specificities of large scale and dynamic systems. Similarly to real world reputation, a digital reputation mechanism expresses a collective opinion about a target user based on aggregated feedback about his past behavior. The resulting reputation score is usually a mathematical object (*e.g.* a number or a percentage). It is used to help entities in deciding whether an interaction with a target user should be considered. Digital reputation mechanisms are thus a powerful tool to incite users to behave trustworthily. Indeed, a user who behaves correctly improves his reputation score, encouraging more users to interact with him. In contrast, misbehaving users have lower reputation scores, which makes it harder for them to interact with other users. To be useful, a reputation mechanism must itself be accurate against adversarial behaviors. Indeed, a user may attack the mechanism to increase his own reputation score or to reduce the reputation of a competitor. A user may also free-ride the mechanism and estimate the reputation of other users without providing his own feedback. From what has been said, it should be clear that reputation is beneficial in order to reduce the potential risk of communicating with almost or completely unknown entities. Unfortunately, the user privacy may easily be jeopardized by reputation mechanisms which is clearly a strong argument to compromise the use of such a mechanism. Indeed, by collecting and aggregating user feedback, or by simply interacting with someone, reputation systems can be easily manipulated in order to deduce user profiles. Thus preserving user privacy while computing robust reputation is a real and important issue that we address in our work [51], [23].

6.4. Other Topics Related to Security and Distributed Computing

6.4.1. Network Monitoring and Fault Detection

Monitoring a system consists in collecting and analyzing relevant information provided by the monitored devices, so as to be continuously aware of the system state (situational awareness). However, the ever growing complexity and scale of systems makes both real time monitoring and fault detection a quite tedious task. Thus the usually adopted option is to focus solely on a subset of information states, so as to provide coarse-grained indicators. As a consequence, detecting isolated failures or anomalies is a quite challenging issue. We propose in [24], [44] to address this issue by pushing the monitoring task at the edge of the network. We present a peer-to-peer based architecture, which enables nodes to adaptively and efficiently self-organize according to their "health" indicators. By exploiting both temporal and spatial correlations that exist between a device and its vicinity, our approach guarantees that only isolated anomalies (an anomaly is isolated if it impacts solely a monitored device) are reported on the fly to the network operator. We show that the end-to-end detection process, *i.e.*, from the local detection to the management operator reporting, requires a logarithmic number of messages in the size of the network.

6.4.2. Metrics Estimation on Very Large Data Streams

In [12], we consider the setting of large scale distributed systems, in which each node needs to quickly process a huge amount of data received in the form of a stream that may have been tampered with by an adversary (*i.e.*, data items ordering can be manipulated by an omniscient adversary [13]). In this situation, a fundamental problem is how to detect and quantify the amount of work performed by the adversary. To address this issue, we propose AnKLe (for Attack-tolerant eNhanced Kullback- Leibler divergence Estimator), a novel algorithm for estimating the KL divergence of an observed stream compared to the expected one. AnKLe combines sampling techniques and information-theoretic methods. It is very efficient, both in terms of space and time complexities, and requires only a single pass over the data stream. Experimental results show that the estimation provided by AnKLe remains accurate even for different adversarial settings for which the quality of other methods dramatically decreases. Considering n as the number of distinct data items in a stream, we show that AnKLe is an (ε, δ)-approximation algorithm with a space complexity $\widetilde{\mathcal{O}}(\frac{1}{\varepsilon} + \frac{1}{\varepsilon^2})$ bits in "most" cases, and $\widetilde{\mathcal{O}}(\frac{1}{\varepsilon} + \frac{n-\varepsilon^{-1}}{\varepsilon^2})$ otherwise. To the best of our knowledge, an approximation algorithm for estimating the Kullback-Leibler divergence has never been analyzed before. We go a step further by considering in [21] the problem of estimating the distance between any two large data streams in small-space constraint. This problem is of utmost importance in data intensive monitoring applications where input streams are generated rapidly. These streams need to be processed on the fly and accurately to quickly determine any deviance from nominal behavior. We present a new metric, the *Sketch* ☆*-metric*, which allows to define a distance between updatable summaries (or sketches) of large data streams. An important feature of the *Sketch* ☆*-metric* is that, given a measure on the entire initial data streams, the *Sketch* ☆*-metric* preserves the axioms of the latter measure on the sketch (such as the non-negativity, the identity, the symmetry, the triangle inequality but also specific properties of the f-divergence or the Bregman one). Extensive experiments conducted on both synthetic traces and real data sets allow us to validate the robustness and accuracy of the *Sketch* ☆*-metric*.

6.4.3. Robustness Analysis of Large Scale Distributed Systems

In the continuation of [53] which proposed an in-depth study of the dynamicity and robustness properties of large-scale distributed systems, in [22], we analyze the behavior of a stochastic system composed of several identically distributed, but non independent, discrete-time absorbing Markov chains competing at each instant for a transition. The competition consists in determining at each instant, using a given probability distribution, the only Markov chain allowed to make a transition. We analyze the first time at which one of the Markov chains reaches its absorbing state. When the number of Markov chains goes to infinity, we analyze the asymptotic behavior of the system for an arbitrary probability mass function governing the competition. We give conditions for the existence of the asymptotic distribution and we show how these results apply to cluster-based distributed systems when the competition between the Markov chains is handled by using a geometric distribution.

6.4.4. Secure Uniform Sampling in Dynamic Systems

In [21], we consider the problem of achieving uniform node sampling in large scale systems in presence of a strong adversary. We first propose an omniscient strategy that processes on the fly an unbounded and arbitrarily biased input stream made of node identifiers exchanged within the system, and outputs a stream that preserves Uniformity and Freshness properties. We show through Markov chains analysis that both properties hold despite any arbitrary bias introduced by the adversary. We then propose a knowledge-free strategy and show through extensive simulations that this strategy accurately approximates the omniscient one. We also evaluate its resilience against a strong adversary by studying two representative attacks (flooding and targeted attacks). We quantify the minimum number of identifiers that the adversary must insert in the input stream to prevent uniformity. To our knowledge, such an analysis has never been proposed before.

CLASSIC Project-Team

5. New Results

5.1. Contributions earlier to 2013 but only published in 2013

Participants: Gérard Biau, Pierre Gaillard, Gilles Stoltz.

We do not discuss here the contributions provided by [14], [12], [13], [16] since they were achieved in 2012 or earlier (but only published this year due to the reviewing and publishing process).

5.2. Approachability with partial monitoring

Participant: Gilles Stoltz.

This line of research has been developed in our team since its creation (see, in particular, the founding article [9] as well as several other publications in the previous reports). Following the earlier contribution on exhibiting an efficient algorithm for approachability with partial monitoring based on some necessary and sufficient dual condition, we study in [15] the primal approach: the statement of the condition and the existence of (efficient or inefficient) algorithms based on it.

5.3. High-dimensional learning and complex data

Participant: Gérard Biau.

We describe four (not so related) contributions on the theme of high-dimensional learning and complex data.

In [17] we address the problem of supervised classification of Cox process trajectories, whose random intensity is driven by some exogenous random covariable. The classification task is achieved through a regularized convex empirical risk minimization procedure, and a nonasymptotic oracle inequality is derived. The results are obtained by taking advantage of martingale and stochastic calculus arguments, which are natural in this context and fully exploit the functional nature of the problem.

The cellular tree classifier model addresses a fundamental problem in the design of classifiers for a parallel or distributed computing world: Given a data set, is it sufficient to apply a majority rule for classification, or shall one split the data into two or more parts and send each part to a potentially different computer (or cell) for further processing? At first sight, it seems impossible to define with this paradigm a consistent classifier as no cell knows the "original data size", n. However, we show in [18] that this is not so by exhibiting two different consistent classifiers.

A new method for combining several initial estimators of the regression function is introduced. Instead of building a linear or convex optimized combination over a collection of basic estimators $r_1, \cdots, r_M$, [19] uses them as a collective indicator of the proximity between the training data and a test observation. This local distance approach is model-free and very fast. More specifically, the resulting collective estimator is shown to perform asymptotically at least as well in the L^2 sense as the best basic estimator in the collective. A companion R package called COBRA (standing for COmBined Regression Alternative) is presented (downloadable on http://cran.r-project.org/web/packages/COBRA/index.html). Substantial numerical evidence is provided on both synthetic and real data sets to assess the excellent performance and velocity of the method in a large variety of prediction problems.

The impact of letting the dimension d go to infinity on the L^p-norm of a random vector with i.i.d. components has surprising consequences, which may dramatically affect high-dimensional data processing. This effect is usually referred to as the *distance concentration phenomenon* in the computational learning literature. Despite a growing interest in this important question, previous work has essentially characterized the problem in terms of numerical experiments and incomplete mathematical statements. In the paper [20], we solidify some of the arguments which previously appeared in the literature and offer new insights into the phenomenon.

5.4. Dimension free principal component analysis

Participants: Olivier Catoni, Ilaria Giulini.

In a work in progress, Ilaria Giulini, as part of her PhD studies, proved the following dimension free inequality, related to Principal Component Analysis in high dimension. Given an i.i.d. sample X_i, $1 \leq i \leq n$ of vector valued random variables $X_i \in \mathbf{R}^d$, there exists an estimator $\widehat{N}$ of the quadratic form $N(\theta) = \mathbf{E}(\langle \theta, X \rangle^2)$ such that for any $n \leq 10^{20}$, with probability at least $1 - 2\epsilon$, for any $\theta \in \mathbf{R}^d$,

$$\mathbf{1}\left(4\mu < 1\right) \left| \frac{\widehat{N}(\theta)}{N(\theta)} - 1 \right| \leq \frac{\mu}{1 - 4\mu},$$

where

$$\mu = \sqrt{\frac{2.07\,(\kappa - 1)}{n} \left[\log\left(\epsilon^{-1}\right) + 4.3 + \frac{1.6\,\|\theta\|^2 \mathbf{Tr}(G)}{N(\theta)} \right]} + \sqrt{\frac{184\,\kappa \|\theta\|^2 \mathbf{Tr}(G)}{nN(\theta)}},$$

where $G = \mathbf{E}\left(XX^\top\right)$ is the Gram matrix and where $\kappa = \sup \left\{ \dfrac{\mathbf{E}\left(\langle \theta, X \rangle^4\right)}{\mathbf{E}\left(\langle \theta, X \rangle^2\right)^2}, \theta \in \mathbf{R}^d \smallsetminus \mathbf{Ker}(G) \right\}$ is some kurtosis coefficient. This result proves that the expected energy in direction θ can be estimated at a rate that is independent of the dimension of the ambient space $\mathbf{R}^d$. It is obtained using PAC-Bayes inequalities with Gaussian parameter perturbations. The same bound holds in a Hilbert space of infinite dimension, opening the possibility of a rigorous mathematical study of kernel principal component analysis of random data, where the data are represented in a possibly infinite dimensional reproducing kernel Hilbert space.

5.5. Statistical models for corpus linguistics

Participants: Olivier Catoni, Thomas Mainguy.

In [21] we describe a language model as the invariant measure of a Markov chain on sentence samples. The kernel of this Markov chain is defined with the help of some context free grammars : from the sentence sample, a random parse model produces a context free grammar with weighted rules, and from this grammar, a new sentence sample is formed by applying the rules randomly. We prove various mathematical properties of this Markov process, related to its computation cost and the fact that it is weakly reversible and therefore ergodic on each of its communicating classes. As a companion to the Markov chain on sentence samples, we can also define a Markov chain on weighted context free grammars. This leads to another type of grammar, that we called Toric Grammars, defined by a family of context tree grammars that can be computed from any of its members as the communicating class of a Markov chain on context free grammars with weighted rules. Preliminary simulations on small data sets are very encouraging, in that they show that this type of model is able to grasp the recursive nature of natural languages.

CLIME Project-Team

6. New Results

6.1. New methods for data assimilation

One major objective of Clime is the conception of 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, focus was on ensemble-variational methods. We introduced a new method known as the iterative ensemble Kalman smoother. It is an ensemble method with an underlying cost function; it does not require the use of the adjoint; and it is flow-dependent. Because of these propreties, the IEnKS outperforms other data assimilation methods when tested with perfect meteorological toy-models. Its potential for parameter estimation has also been demonstrated.

6.1.1. An iterative ensemble Kalman smoother

Participants: Marc Bocquet, Pavel Sakov [BOM, Australia].

The iterative ensemble Kalman filter (IEnKF) was recently proposed to improve the performance of ensemble Kalman filtering with strongly nonlinear geophysical models. IEnKF can be used as a lag-one smoother and extended to a fixed-lag smoother: the iterative ensemble Kalman smoother (IEnKS [12]). IEnKS is an ensemble variational method. It does not require the use of the tangent linear of the evolution and observation models, nor the adjoint of these models: the required sensitivities (gradient and Hessian) are obtained from the ensemble. Looking for the optimal performance, we consider a quasi-static algorithm, out of the many possible extensions. IEnKS is explored on the Lorenz'95 model and on a 2D turbulence model. As a logical extension of IEnKF, IEnKS significantly outperforms standard Kalman filters and smoothers in strongly nonlinear regimes. In mildly nonlinear regimes (typically synoptic scale meteorology), its filtering performance is marginally but clearly better than the standard ensemble Kalman filter, and it keeps improving as the length of the temporal data assimilation window is increased. For long windows, its smoothing performance very significantly outranks the standard smoothers, which is believed to stem from the variational but flow-dependent nature of the algorithm. For very long windows, the use of a multiple data assimilation variant of the scheme, where observations are assimilated several times, is advocated. This paves the way for finer re-analysis freed from the static prior assumption of 4D-Var, but also partially freed from the Gaussian assumptions that usually impede standard ensemble Kalman filtering and smoothing.

6.1.2. Joint state and parameter estimation with an iterative ensemble Kalman smoother

Participants: Marc Bocquet, Pavel Sakov [BOM, Australia].

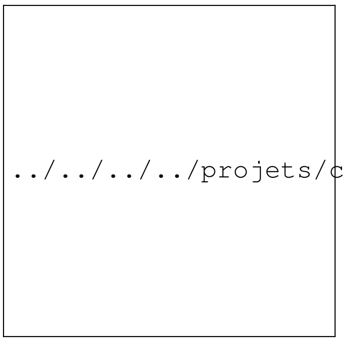
Both ensemble filtering and variational data assimilation methods have proven being useful in the joint estimation of state variables and parameters of geophysical models. Yet, their respective benefits and drawbacks in this task are distinct. An ensemble variational method, known as the iterative ensemble Kalman smoother (IEnKS), has recently been introduced. It is based on an adjoint-free variational but flow-dependent scheme. As such, IEnKS is a candidate tool for joint state and parameter estimation that may inherit the benefits from both the ensemble filtering and variational approaches.

In this study [13], an augmented state IEnKS is tested on the estimation of the forcing parameter of the Lorenz'95 model. Since joint state and parameter estimation is especially useful in applications where the forcings are uncertain but nevertheless determining, typically in atmospheric chemistry, the augmented state IEnKS is tested on a new low-order model that combines the Lorenz'95 model, representing its meteorological part, and the advection diffusion of a tracer for its chemical part. In these experiments, IEnKS is compared to the ensemble Kalman filter, the ensemble Kalman smoother and a 4D-Var method, that are considered choices to solve these joint estimation problems. In this low-order model context, IEnKS is shown to significantly outperform those methods, for any length of the data assimilation window, and for present time analysis as well as retrospective analysis. Besides, the performance of IEnKS is even more striking on parameter estimation, whereas getting close to the same performance with 4D-Var is likely to require both a long data assimilation window and a complex modeling of the background statistics.

6.1.3. Data assimilation applied to air quality at urban scale

Participants: Vivien Mallet, Raphaël Périllat, Anne Tilloy, Fabien Brocheton [Numtech], David Poulet [Numtech], Frédéric Mahé [Airparif], Pierre Pernot [Airparif], Fabrice Joly [Airparif].

Based on Verdandi [14], Polyphemus and the "Urban Air Quality Analysis" software, data assimilation was further developed at urban scale. The Best Linear Unbiased Estimator (BLUE) is computed to merge the outputs of the ADMS Urban model and the observations of a sparse monitoring network [19]. We improved the modeling of the covariance of the model state error. The assimilation was applied for part of Paris (see Fig. 2) and for Paris region, in the context of the PRIMEQUAL project PREQUALIF ("Multidisciplinary Program on Air Quality Research in Île-de-France").



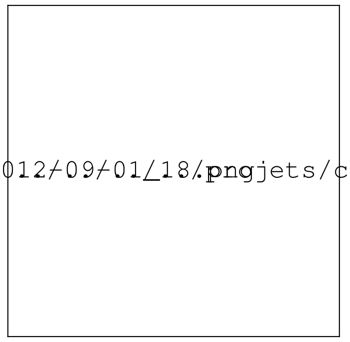


Figure 2. Left: Map of [NO_2] ($\mu g\ m^{-3}$), before assimilation, at a given date in September 2012. Right: Map of [NO_2] ($\mu g\ m^{-3}$), after assimilation of the observations (disks).

It was applied to nitrogen dioxide, particulate matter and black carbon. Specific investigations were carried out to estimate the variance of the a posteriori error and to determine the impact of each monitoring station on the final results.

6.2. Inverse modeling

We continued research on inverse modelling techniques, with a focus on hyperparameter estimation when the statistics are non-Gaussian. We applied these methods to the estimation of the caesium-137 Fukushima source term using heterogenous datasets. We applied similar methods to the estimation of Volatile Organic Compounds (VOC) at the European scale by assimilation of the EMEP VOC observations over one year. We also studied the estimation of several hyperparameters in the context of CO_2 flux inversions.

6.2.1. Estimation of the caesium-137 source term from the Fukushima Daiichi nuclear power plant using a consistent joint assimilation of air concentration and deposition observations

Participants: Victor Winiarek, Marc Bocquet, Nora Duhanyan [CEREA], Yelva Roustan [CEREA], Olivier Saunier [IRSN], Anne Mathieu [IRSN].

To estimate the amount of radionuclides and the temporal profile of the source term released in the atmosphere during the accident of the Fukushima Daiichi nuclear power plant in March 2011, inverse modeling techniques have been used and have proven their ability in this context. In a previous study, the lower bounds of the caesium-137 and iodine-131 source terms were estimated with such techniques, using activity concentration observations. The importance of an objective assessment of prior errors (the observation errors and the background errors) was emphasised for a reliable inversion. In such critical context where the meteorological conditions can make the source term partly unobservable and where only a few observations are available, such prior estimation techniques are mandatory, the retrieved source term being very sensitive to this estimation.

We propose to extend the use of these techniques to the estimation of prior errors when assimilating observations from several data sets [21]. The aim is to compute an estimate of the caesium-137 source term jointly using all available data about this radionuclide, such as activity concentrations in the air, but also daily fallout measurements and total cumulated fallout measurements. It is crucial to properly and simultaneously estimate the background errors and the prior errors relative to each data set. A proper estimation of prior errors is also a necessary condition to reliably estimate the a posteriori uncertainty of the estimated source term. Using such techniques, we retrieve a total released quantity of caesium-137 in the interval $11.6 - 19.3$ PBq with an estimated standard deviation range of $15 - 20\%$ depending on the method and the data sets. The "blind" time intervals of the source term have also been strongly mitigated compared to the first estimations with only activity concentration data.

6.2.2. An inverse modeling method to assess the source term of the Fukushima Nuclear Power Plant accident using gamma dose rate observations

Participants: Olivier Saunier [IRSN], Anne Mathieu [IRSN], Damien Didier [IRSN], Maryline Tombette [IRSN], Denis Quélo [IRSN], Victor Winiarek, Marc Bocquet.

The Chernobyl nuclear accident, and more recently the Fukushima accident, highlighted that the largest source of error on consequences assessment is the source term, including the time evolution of the release rate and its distribution between radioisotopes. Inverse modeling methods, which combine environmental measurements and atmospheric dispersion models, have proven being efficient in assessing source term due to an accidental situation. Most existing approaches are designed to use air sampling measurements and some of them also use deposition measurements [21]. Some studies have been conceived to use dose rate measurements, but none of the developed methods were carried out to assess the complex source term of a real accident situation like the Fukushima accident. However, dose rate measurements are generated by the most widespread measurement system and, in the event of a nuclear accident, these data constitute the main source of measurements of the plume and radioactive fallout during releases. This study [18], [23] proposes a method to use dose rate measurements as part of an inverse modeling approach to assess source terms. The method is proven efficient and reliable when applied to the accident at the Fukushima Daiichi Nuclear Power Plant (FD-NPP). The emissions for the eight main isotopes have been assessed. Accordingly, 105.9 PBq of ^{131}I, 35.8 PBq of ^{132}I, 15.5 PBq of ^{137}Cs and 12,134 PBq of noble gases were released. The events at FD-NPP (such as venting, explosions, etc.) known to have caused atmospheric releases are well identified in the retrieved source term. The estimated source term is validated by comparing simulations of atmospheric dispersion and deposition with environmental observations. In total, it was found that for 80 % of the measurements, simulated and observed dose rates agreed within a factor of 2. Changes in dose rates over time have been overall properly reconstructed, especially in the most contaminated areas to the northwest and south of the FD-NPP. A comparison with observed atmospheric activity concentration and surface deposition shows that the emissions of caesiums and ^{131}I are realistic but that ^{132}I and ^{132}Te are probably underestimated and noble gases are likely overestimated. Finally, an important outcome of this study is that the method proved to be perfectly suited to emergency management and could contribute to improve emergency response in the event of a nuclear accident.

6.2.3. Estimation of volatile organic compound emissions for Europe using data assimilation

Participants: Mohammad Reza Koohkan, Marc Bocquet, Yelva Roustan [CEREA], Yougseob Kim [CEREA], Christian Seigneur [CEREA].

The emissions of non-methane volatile organic compounds (VOCs) over western Europe for the year 2005 are estimated via inverse modeling by assimilation of in situ observations of concentration and they are subsequently compared to a standard emission inventory. The study [16] focuses on fifteen VOC species: five aromatics, six alkanes, two alkenes, one alkyne and one biogenic diene. The inversion relies on a validated fast adjoint of the chemical transport model used to simulate the fate and transport of these VOCs. The assimilated ground-based measurements over Europe are provided by the European Monitoring and Evaluation Programme (EMEP) network. The background emissions errors and the prior observational errors are estimated by maximum likelihood approaches. The positivity assumption on the VOC emission fluxes is pivotal for a successful inversion and this maximum likelihood approach consistently accounts for the positivity of the fluxes. For most species, the retrieved emissions lead to a significant reduction of the bias, which underlines the misfit between the standard inventories and the observed concentrations. The results are validated through a forecast test and a cross-validation test. An estimation of the posterior uncertainty is also provided. It is shown that the statistically consistent non-Gaussian approach, based on a reliable estimation of the errors, offers the best performance. The efficiency in correcting the inventory depends on the lifetime of the VOCs and the accuracy of the boundary conditions. In particular, it is shown that the use of in situ observations using a sparse monitoring network to estimate emissions of isoprene is inadequate because its short chemical lifetime significantly limits the spatial radius of influence of the monitoring data. For species with longer lifetime (a few days), successful, albeit partial, emission corrections can reach regions hundreds of kilometres away from the stations. Domainwide corrections of the emissions inventories of some VOCs are significant, with underestimations on the order of a factor of two for propane, ethane, ethylene and acetylene.

6.2.4. Hyperparameter estimation for uncertainty quantification in mesoscale carbon dioxide inversions

Participants: Lin Wu [LSCE, France], Marc Bocquet, Frédéric Chevallier [LSCE, France], Thomas Lauvaux [Department of Meteorology, Pennsylvania State University, USA], Kenneth Davies [Department of Meteorology, Pennsylvania State University, USA].

Uncertainty quantification is critical in the inversion of CO_2 surface fluxes from atmospheric concentration measurements. We estimate the main hyperparameters of the error covariance matrices for a priori fluxes and CO_2 concentrations, that is, the variances and the correlation lengths, using real, continuous hourly CO_2 concentration data in the context of the Ring 2 experiment of the North American Carbon Program Mid Continent Intensive. Several criteria, namely maximum likelihood (ML), general cross-validation (GCV) and χ^2 test are compared for the first time under a realistic setting in a mesoscale CO_2 inversion. It is shown [22] that the optimal hyperparameters under the ML criterion assure perfect χ^2 consistency of the inverted fluxes. Inversions using the ML error variances estimates rather than the prescribed default values are less weighted by the observations, because the default values underestimate the model-data mismatch error, which is assumed to be dominated by the atmospheric transport error. As for the spatial correlation length in prior flux errors, the Ring 2 network is sparse for GCV and this method fails to reach an optimum. In contrast, the ML estimate (e.g. an optimum of 20 km for the first week of June 2007) does not support long spatial correlations that are usually assumed in the default values.

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, we studied the impact of using lidar observation for particulate matter forecasting.

6.3.1. Assimilation of ground versus lidar observations for PM_{10} forecasting

Participants: Yiguo Wang [CEREA], Karine Sartelet [CEREA], Marc Bocquet, Patrick Chazette [LSCE, France].

This study [20] investigates the potential impact of future ground-based lidar networks on analysis and short-term forecasts of PM_{10}. To do so, an Observing System Simulation Experiment (OSSE) is built for PM_{10} data assimilation using optimal interpolation over Europe for one month in 2001. First, we estimate the efficiency of the assimilation of lidar network measurements in improving PM_{10} concentration analysis and forecast. It is compared to the efficiency of assimilating concentration measurements from the AirBase ground network, which includes about 500 stations in western Europe. It is found that the assimilation of lidar observations is more efficient at improving PM_{10} concentrations in terms of root mean square error and correlation after 12 hours of assimilation than the assimilation of AirBase measurements. Moreover, the spatial and temporal influence of the assimilation of lidar observations is larger and longer. In our experiments, the assimilation of lidar products improves PM_{10} forecast for 108 hours against 60 hours for AirBase assimilation. The results show a potentially powerful impact of the future lidar networks. Secondly, since a lidar is a very costly instrument, a sensitivity study on the number of required lidars is performed to help defining an optimal lidar network for PM_{10} forecast. The results suggest 12 lidar stations over western Europe, because a network with 26 lidar stations is more expensive and offers a limited improvement (less than $1\ \mu g\ m^{-3}$ of root mean square error on average) over the lidar network. A comparison of two networks with 12 lidar stations at different locations does not lead to substantial differences.

6.4. Reduction and emulation

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

- the dimension of the inputs, which can easily be $10^5 - -10^8$ at every time step;
- the dimension of the state vector, which is usually $10^5 - -10^7$;
- the high computational cost.

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 (POD) 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 approaches. Promising results were however obtained with radial basis functions and an adapted kriging-based method.

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 (POD) on the outputs allowed us to drastically reduce their dimension, from 10^4 to just a few scalars. The emulation of the reduced model itself was successfully carried out with radial basis functions or an adapted kriging-based method. The resulting reduced/emulated model exhibited meaningful response to all variables. Its performance compared to observations was the same as the original model. The computational cost of the full model is about 8 minutes on 16 cores (for a single time step), while the reduced/emulated model requires only 50 ms on one core [29].

6.4.3. Motion estimation from images with a waveforms reduced model

Participants: Etienne Huot, Isabelle Herlin, Giuseppe Papari [Lithicon, Norway], Karim Drifi.

Dimension reduction is applied to an image model, composed of Lagrangian constancy of velocity and transport of image brightness. Waveforms basis are obtained on the image domain for subspaces of images, motion fields and divergence-free motion fields, as eigenvectors of quadratic functions. Image assimilation with th reduced model allows to estimate velocity fields satisfying space-time properties defined by user and traduced as a quadratic function. This approach also solves the issue of complex geographical domains and the difficulty of applying boundary conditions on these domains. 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 [31], [26], [25].

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

Participants: Paul Baudin, Vivien Mallet, Gilles Stoltz [CNRS], Laurent Descamps [Météo France].

Nowadays, it is standard procedure to generate an ensemble of simulations for a meteorological forecast. Usually, meteorological centers produce a single forecast, out of the ensemble forecasts, computing the ensemble mean (where every model receives an equal weight). It is however possible to apply aggregation methods. When new observations are available, the meteorological centers also compute analyses. Therefore, we can apply the ensemble forecast of analyses. Ensembles of forecasts for wind velocity and mean sea level pressure, from Météo France, were aggregated. Preliminary results show significant improvements for mean sea level pressure.

6.5.2. Sequential aggregation with uncertainty estimation

Participants: Vivien Mallet, Sergiy Zhuk [IBM research, Ireland], Paul Baudin, Gilles Stoltz [CNRS].

An important 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 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. We also introduced a criterion that the filter results should satisfy if they are representative of the uncertainties [17].

6.6. Uncertainty quantification

Many uncertainties limit the forecast skills of geophysical simulations: limited understanding of physical phenomena, simplified representation of a system state and of the physical processes, inaccurate data and approximate numerical solutions. In many applications, a deterministic forecast or analysis is not enough a result since its uncertainties may be very large. It is of high interest to evaluate the quality of a forecast, before observations are available, and the quality of an analysis at any location, observed or not. An even more desirable result is the full probability density of system state, which can only be derived from a fully stochastic approach.

6.6.1. Sensitivity analysis in the dispersion of radionuclides

Participants: Sylvain Girard [IRSN], Vivien Mallet, Irène Korsakissok [IRSN].

We carried out a sensitivity analysis of the dispersion of radionuclides during Fukushima disaster. We considered the dispersion at regional scale, with the Eulerian transport model Polair3D from Polyphemus. The sensitivities to most input parameters were computed using the Morris method (with 8 levels and 100 trajectories). The influences of 19 scalar parameters were quantified. The scalar parameters were additive terms or multiplicative factors applied to 1D, 2D or 3D fields such as emission rates, precipitations, cloud height, wind velocity. It was shown that, depending on the output quantities of interest (various aggregated atmospheric and ground dose rates), the sensitivity to the inputs may greatly vary. Very few parameters show low sensitivity in any case. The vertical diffusion coefficient, the scavenging factors, the winds and precipitation intensity were found to be the most influential inputs. Most input variables related to the source term (emission rates, emission dates) also had a strong influence.

6.7. 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 location of structures on the observations is used to control the state vector.
- from the image's viewpoint, a model of the dynamics and structures is built from the observations.

6.7.1. 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 either characterized by its vorticity value or by its coefficients on a divergence-free basis and assumed to satisfy the Lagragian constancy hypothesis. An image assimilation method, based on the 4D-Var technic, is defined that estimates motion as a compromise between the evolution equations of vorticity or projection coefficients 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.

6.7.2. Model error and motion estimation

Participants: Dominique Béréziat [UPMC], Isabelle Herlin.

Data assimilation technics are used to retrieve motion from image sequences. These methods require a model of the underlying dynamics, displayed by the evolution of image data. In order to quantify the approximation linked to the chosen dynamic model, an error term is included in the evolution equation of motion and a weak formulation of 4D-Var data assimilation is designed. The cost function to be minimized 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.

The approach has been used to estimate the impact of geophysical forces (gravity, Coriolis, diffusion) and better assess the surface dynamics [24].

6.7.3. *Tracking of structures from an image sequence*

Participants: Yann Lepoittevin, Isabelle Herlin, Dominique Béréziat [UPMC].

The research concerns an approach to estimate velocity on an image sequence and simultaneously segment and track a given structure. It relies on the underlying dynamics' equations of the studied physical system. A data assimilation method is designed to solve evolution equations of image brightness, those of motion's dynamics, and those of the distance map modeling the tracked structures. The method is applied on meteorological satellite data, in order to track tropical clouds on image sequences and estimate their motion, as seen on Fig. 3 .

../../../../projets/clime/IMG/tracking-3.png ../../../../projets/clime/IMG/tracking-9.png ../../../../projets/clime/IMG/

Figure 3. Tracking a tropical cloud. Frames 3, 9, 18 of the sequence.

Quantification is obtained on synthetic experiments by comparing trajectories of characteristic points. The respective position of these points on the last image of the sequence for different methods may be compared to that obtained with ground truth as seen on Fig. 4 .

../../../../projets/clime/IMG/ComPts-17-ell.png

Figure 4. Red point: ground truth. Blue point: our method. Green point: Sun's optical flow. Blue ellipse: our method is the best. Green ellipse: Sun's result is the best. Grey ellipse : results are equivalent.

Data assimilation is performed either with a 4D-Var variational approach [27], [30], [28] or with an ensemble approach. In the last case, computation of the ensemble from optical flow methods of the literature is currently studied.

6.8. Minimax filtering

In minimax filtering for state estimation, the initial state error, the model error and the observation errors are supposed to belong to one joint ellipsoid. It is only assumed that the errors, stochastic or deterministic, are bounded. During the assimilation process, 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.

6.8.1. Retrieval of a continuous image function and 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].

An iterative minimax method is developed for the problem of motion estimation from an image sequence. 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 a continuous image function $\widehat{I}$, 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 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.

Alternatively, point 1) may be used to retrieve a continuous image function from sparse and noisy image snapshots, based on previous motion estimation with a 4D-Var technic as seen on Fig. 5 , that displays ground truth, noisy image observation, image estimation at the end of the studied intervall.

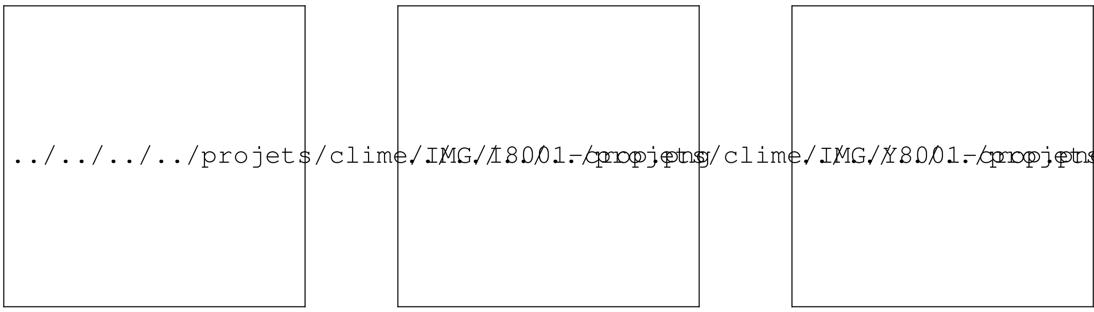


Figure 5. From left to right: Ground truth, image observation, result.

6.9. Fire application

6.9.1. 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 model performance, but currently, simulation errors are not sufficiently qualified and quantified.

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 [15]. One simulation requires mainly the following data: the ignition point, the ground elevation, the vegetation cover and the wind field. See illustration in Fig. 6 . We simulated 80 fires with four physical models, which proved that the most advanced models performed better overall, even though the input data is often inaccurate. We also carried out Monte Carlo simulations to evaluate the impact of the uncertainty in input data. We showed that the Monte Carlo approach led to a reliable forecasting system, which suggests that the probability densities derived from the simulations (see Fig. 6) may be useful information for preventive actions in an operational context.

Figure 6. Left: 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). Right: Burn probability as computed by a Monte Carlo simulation for a wildfire that was observed (red contour) in Corsica in 2003.

COATI Project-Team

6. New Results

6.1. Network Design and Management

Participants: Julio Araújo, Jean-Claude Bermond, Luca Chiaraviglio, David Coudert, Frédéric Giroire, Alvinice Kodjo, Aurélien Lancin, Remigiusz Modrzejewski, Christelle Molle-Caillouet, Joanna Moulierac, Nicolas Nisse, Stéphane Pérennes, Truong Khoa Phan, Ronan Pardo Soares, Issam Tahiri.

6.1.1. Optimization in backbone networks

6.1.1.1. Shared Risk Link Group

The notion of Shared Risk Link Groups (SRLG) has been introduced to capture survivability issues where some links of a network fail simultaneously. In this context, the diverse routing problem is to find a set of pairwise SRLG-disjoint paths between a given pair of end nodes of the network. This problem has been proved NP-complete in general and some polynomial instances have been characterized.

In [33], [32], we investigate the diverse-routing problem in networks where the SRLGs are localized and satisfy the *star property*. This property states that a link may be subject to several SRLGs, but all links subject to a given SRLG are incident to a common node. We first provide counterexamples to the polynomial-time algorithm proposed in the literature for computing a pair of SRLG-disjoint paths in networks with SRLGs satisfying the star property, and then prove that this problem is in fact NP-complete. We have also characterized instances that can be solved in polynomial time or are fixed parameter tractable, in particular when the number of SRLGs is constant, the maximum degree of the vertices is at most 4, and when the network is a directed acyclic graph. Moreover, we have considered the problem of finding the maximum number of SRLG-disjoint paths in networks with SRLGs satisfying the star property. We have proved that such problem is NP-hard and hard to approximate. Then, we have provided exact and approximation algorithms for relevant subcases.

6.1.1.2. Wavelength assignment in WDM networks

Let $\mathcal{P}$ be a family of directed paths in a directed graph G. The load of an arc is the number of directed paths containing this arc. Let $\pi(G,\mathcal{P})$ be the maximum of the load of all the arcs and let $w(G,\mathcal{P})$ be the minimum number of wavelengths (colours) needed to colour $\mathcal{P}$ in such a way that two directed paths with the same wavelength are arc-disjoint. These two parameters correspond respectively to the clique number and the chromatic number of the associated conflict graph, and $\pi(G,\mathcal{P}) \leq w(G,\mathcal{P})$. It was known that there exists directed acyclic graphs (DAGs) such that the ratio between $w(G,\mathcal{P})$ and $w(G,\mathcal{P})$ is arbitrarily large. In [18], solving a conjecture of an earlier article, we show that the same is true for a very restricted class of DAGs, the UPP-DAGs, those for which there is at most one directed path from a vertex to another. We also characterized the DAGs such that $\pi(G,\mathcal{P}) = w(G,\mathcal{P})$ for all families of directed paths.

6.1.1.3. Mutli-operators microwave backhaul networks

In [35], we consider the problem of sharing the infrastructure of a backhaul network for routing. We investigate on the revenue maximization problem for the physical network operator (PNO) when subject to stochastic traffic requirements of multiple virtual network operators (VNO) and prescribed service level agreements (SLA). We use robust optimization to study the tradeoff between revenue maximization and the allowed level of uncertainty in the traffic demands. This mixed integer linear programming model takes into account end-to-end traffic delays as example of quality-of-service requirement in a SLA. To show the effectiveness of our model, we present a study on the price of robustness, i.e. the additional price to pay in order to obtain a feasible solution for the robust scheme, on realistic scenarios.

6.1.2. Energy efficiency

With one third of the world population online in 2013 and an international Internet bandwidth multiplied by more than eight since 2006, the ICT sector is a non-negligible contributor of worldwide greenhouse gases emissions and power consumption. Indeed, power consumption of telecommunication networks has become a major concern for all the actors of the domain, and efforts are made to reduce their impact on the overall figure of ICTs, and to support its foreseen growth in a sustainable way. In this context, the contributors of the European Network of Excellence TREND have developed innovative solutions to improve the energy efficiency of optical networks summarized in [45].

6.1.2.1. Energy aware routing with redundancy elimination

Many studies have shown that energy-aware routing (EAR) can significantly reduce energy consumption of a backbone network. Redundancy Elimination (RE) techniques provide a complementary approach to reduce the amount of traffic in the network. In particular, the GreenRE model combines both techniques, offering potentially significant energy savings.

In [44], we enhance the MIP formulation proposed in [75] for the GreenRE model. We derive cutting planes, extending the well-known cutset inequalities, and report on preliminary computations.

In [37], we propose a concept for respecting uncertain rates of redundant traffic within the GreenRE model, closing the gap between theoretical modeling and drawn-from-life data. To model redundancy rate uncertainty, the robust optimization approach in [73] is adapted and the problem is formally defined as mixed integer linear program. An exemplary evaluation of this concept with real-life traffic traces and estimated fluctuations of data redundancy shows that this closer-to-reality model potentially offers significant energy savings in comparison to GreenRE and EAR.

6.1.2.2. Energy Efficient Content Distribution

The basic protocols of the Internet are point-to-point in nature. However, the traffic is largely broadcasting, with projections stating that as much as 80-90% of it will be video by 2016. This discrepancy leads to an inefficiency, where multiple copies of essentially the same messages travel in parallel through the same links. We have studied approaches to mitigate this inefficiency and reduce the energy consumption of future networks, in particular in [13].

In [29], we study the problem of reducing power consumption in an Internet Service Provider (ISP) network by designing the content distribution infrastructure managed by the operator. We propose an algorithm to optimally decide where to cache the content inside the ISP network. We evaluate our solution over two case studies driven by operators feedback.

Recently, there is a trend to introduce content caches as an inherent capacity of network equipment, with the objective of improving the efficiency of content distribution and reducing network congestion. In [57], [46], [29], we study the impact of using in-network caches and content delivery network (CDN) cooperation on an energy-efficient routing. Experimental results show that by placing a cache on each backbone router to store the most popular content, along with well choosing the best content provider server for each demand to a CDN, we can save up to 23% of power in the backbone.

6.1.3. Distributed systems

6.1.3.1. Distributed Storage systems.

In a P2P storage system using erasure codes, a data block is encoded in many redundancy fragments. These fragments are then sent to distinct peers of the network. In [24], we study the impact of different placement policies of these fragments on the performance of storage systems.

In [39], we propose a new analytical framework that takes into account the correlation between data reconstructions when estimating the repair time and the probability of data loss. The models and schemes proposed are validated by mathematical analysis, extensive set of simulations, and experimentation using the GRID5000 test-bed platform. This new model allows system designers to operate a more accurate choice of system parameters in function of their targeted data durability.

6.1.3.2. P2P Streaming systems

In [41], [68], we propose and analyze a simple localized algorithm to balance a tree. The motivation comes from live distributed streaming systems in which a source diffuses a content to peers via a tree, a node forwarding the data to its children. Such systems are subject to a high churn, peers frequently joining and leaving the system. It is thus crucial to be able to repair the diffusion tree to allow an efficient data distribution. In particular, due to bandwidth limitations, an efficient diffusion tree must ensure that node degrees are bounded. Moreover, to minimize the delay of the streaming, the depth of the diffusion tree must also be controlled. We propose here a simple distributed repair algorithm in which each node carries out local operations based on its degree and on the subtree sizes of its children.

6.1.4. Data Gathering in Radio Networks

We study the problem of gathering information from the nodes of a radio network into a central node. We model the network of possible transmissions by a graph and consider a binary model of interference in which two transmissions interfere if the distance in the graph from the sender of one transmission to the receiver of the other is d_I or less.

In [19], we give an algorithm to construct minimum makespan transmission schedules for data gathering under the following hypotheses: the communication graph G is a tree network, and no buffering is allowed at intermediate nodes and $d_I \geq 2$. In the interesting case in which all nodes in the network have to deliver an arbitrary positive number of packets, we provide a closed formula for the makespan of the optimal gathering schedule. Additionally, we consider the problem of determining the computational complexity of data gathering in general graphs and show that the problem is NP–complete. On the positive side, we design a simple $(1 + 2/d_I)$-factor approximation algorithm for general networks.

In [59], we focus on the gathering and personalized broadcasting problem in grids. We still consider the non-buffering model. In this setting, though the problem of determining the complexity of computing the optimal makespan in a grid is still open, we present linear (in the number of messages) algorithms that compute schedules for gathering with $d_I = 0, 1, 2$. In particular, we present an algorithm that achieves the optimal makespan up to a small additive constant. Note that, the approximation algorithms that we present also provide approximation up to a ratio 2 for the gathering with buffering. All our results are proved in terms of personalized broadcasting.

In [20], we now allow transmission till a distance d_T and buffering in intermediate nodes. We focus on the specific case where the network is a path with the sink at an end vertex of the path and where the traffic is unitary ($w(u) = 1$ for all u); indeed this simple case appears to be already very difficult. We first give a new lower bound and a protocol with a gathering time that differs only by a constant independent from the length of the path. Then we present a method to construct incremental protocols which are optimal for many values of d_T and d_I (in particular when d_T is prime).

In [50], we focus on gathering uncertain traffic demands in mesh networks with multiple sources and sinks. The scheduling is relaxed into the round weighting problem in which a set of pairwise non-interfering links is called a round, and we seek to successively activate rounds in order to get enough capacity on links to route the demand from the set of sources to the set of sinks. We propose a new robust model considering traffic demand uncertainty, efficiently solved by column generation, and quantify the price of robustness, i.e., the additional cost to pay in order to obtain a feasible solution for the robust scheme.

6.1.5. Routing

6.1.5.1. Routing models evaluation

The Autonomous System (AS)-level topology of the Internet that currently comprises more than 40k ASs, is growing at a rate of about 10% per year. In these conditions, Border Gateway Protocol (BGP), the inter-domain routing protocol of the Internet starts to show its limits, among others in terms of the number of routing table entries it can dynamically process and control. To overcome this challenging situation, the design but also the evaluation of alternative dynamic routing models and their comparison with BGP will be performed by means of simulation. However, existing routing models simulators such as DRMSim, the Dynamic Routing Model

Simulator developped in COATI in collaboration with Alcatel-Lucent [72], are limited in terms of the number of routing table entries they can dynamically process and control on a single computer.

In [63], we have conducted a feasibility study of the extension of DRMSim so as to support the Distributed Parallel Discrete Event paradigm. We have studied several distribution models and their associated communication overhead. We have in particular evaluated the expected additional time required by a distributed simulation of BGP (border gate protocol) on topologies with 100k ASes compared to its sequential simulation. We show that such a distributed simulation of BGP is possible with a reasonable time overhead.

6.1.5.2. Complexity of Shortest Path Routing

In telecommunication networks packets are carried from a source s to a destination t on a path that is determined by the underlying routing protocol. Most routing protocols belong to the class of shortest-path routing protocols. For better protection and efficiency, one wishes to use multiple (shortest) paths between two nodes. Therefore the routing protocol must determine how the traffic from s to t is distributed among the shortest paths. In the protocol called OSPF-ECMP (for Open Shortest Path First-Equal Cost Multiple Path) the traffic incoming at every node is uniformly balanced on all outgoing links that are on shortest paths. In [43], [42], we show that the problem of maximizing even a single commodity flow for the OSPF-ECMP protocol cannot be approximated within any constant factor ratio. Besides this main theorem, we derive some positive results which include polynomial-time approximations and an exponential-time exact algorithm.

6.2. Graph Algorithms

Participants: Julio Araújo, Jean-Claude Bermond, David Coudert, Frédéric Havet, Frédéric Giroire, Bi Li, Fatima Zahra Moataz, Christelle Molle-Caillouet, Nicolas Nisse, Ronan Pardo Soares, Stéphane Pérennes.

COATI is also interested in the algorithmic aspects of Graph Theory. In general we try to find the most efficient algorithms to solve various problems of Graph Theory and telecommunication networks. More information on several results presented in this section may be found in R. Soares's thesis [14].

6.2.1. Complexity and Computation of Graph Parameters

We use graph theory to model various network problems. In general we study their complexity and then we investigate the structural properties of graphs that make these problems hard or easy. In particular, we try to find the most efficient algorithms to solve the problems, sometimes focusing on specific graph classes from which the problems are polynomial-time solvable.

6.2.1.1. Parameterized Complexity

Parameterized complexity is a way to deal with intractable computational problems having some parameters that can be relatively small with respect to the input size. This area has been developed extensively during the last decade. More precisely, we consider problems that consist in deciding whether a graph G satisfies some property (i.e., if G belongs to some given family of graphs). For decision problems with input size n and parameter k, the goal is to design an algorithm with running time $f(k).n$, where f depends only on k. Problems for which we can find an optimal algorithm with such time complexity are said to be fixed-parameter tractable (FPT). Equivalently, the goal is to design a polynomial-time algorithm (in k and n) that computes a pair (H, k') where H is a graph (the kernel) with size polynomial in k and $P(G) \leq k$ if and only if $P(H) \leq k'$.

We study the parameterized complexity of the edge-modification problems. Given a graph $G = (V, E)$ and a positive integer k, an edge modification problem for a graph property Π consists in deciding whether there exists a set F of pairs of V of size at most k such that the graph $H = (V, E\Delta F)$ satisfies the property Π. In [25], it is proved that parameterized cograph edge-modification problems have cubic vertex kernels whereas polynomial kernels are unlikely to exist for the P_l-free edge-deletion and the C_l-free edge-deletion problems for $l \geq 7$ and $l \geq 4$ respectively.

We also design a unified parameterized algorithm for computing various widths of graphs (such as branched tree-width, branch-width, cut-width, etc.) [60].

6.2.1.2. Convexity in Graphs

The geodesic convexity of graphs naturally extends the notion of convexity in euclidean metric spaces. A set S of vertices of a graph $G = (V, E)$ is *convex* if any vertex on a shortest path between two vertices of S also belongs to S. The *convex hull* of $S \subset V$ is the smallest convex set containing S. Finally, a *hull set* of a graph is a set of vertices whose convex hull is V. The hull number of a graph G is the minimum size of a hull set in G. In [16] , we prove that computing the hull number is NP-complete in bipartite graphs. We also provide bounds and design various polynomial-time algorithms for this problem in different graph classes such as co-bipartite graphs, P_4-sparse graphs, etc. In [30], we first show a polynomial-time algorithm to compute the hull number of any P_5-free triangle-free graph. Then, we present four reduction rules based on vertices with the same neighborhood. We use these reduction rules to propose a fixed-parameter tractable algorithm to compute the hull number of any graph G, where the parameter is the size of a vertex cover of G or, more generally, its neighborhood diversity. We also use these reductions to characterize the hull number of the lexicographic product of any two graphs.

6.2.1.3. Hyperbolicity

The Gromov hyperbolicity is an important parameter for analyzing complex networks since it expresses how the metric structure of a network looks like a tree. In other words, it provides bounds on the stretch resulting from the embedding of a network topology into a weighted tree. It is therefore used to provide bounds on the expected stretch of greedy-routing algorithms in Internet-like graphs. However, the best known algorithm for computing this parameter has time complexity in $O(n^{3.69})$, which is prohibitive for large-scale graphs. In [36], we proposed a novel algorithm for determining the hyperbolicity of a graph that is scalable for large graphs. The time complexity of this algorithm is output-sensitive and depends on the shortest-path distances distribution in the graph and on the computed value of the hyperbolicity. Although its worst case time complexity is in $O(n^4)$, it is in practice much faster than previous proposals as it uses bounds to cut the search space. This algorithm allowed us for computing the hyperbolicity of all maps of the Internet provided by CAIDA and DIMES.

6.2.2. Graph searching and applications

Pursuit-evasion encompasses a wide variety of combinatorial problems related to the capture of a fugitive residing in a network by a team of searchers. The goal consists in minimizing the number of searchers required to capture the fugitive in a network and in computing the corresponding capture strategy. We investigated several variants of these games.

6.2.2.1. Variants of graph searching.

We study non-deterministic graph searching where the searchers have to capture an invisible fugitive but can see him a bounded number of times. This variant generalizes the notion of pathwidth and treewidth of graphs. In this setting, we provide a polynomial-time algorithm that approximates the minimum number of searchers needed in trees, up to a factor of two [56].

In [34], [61], we define another variant of graph searching, where searchers have to capture an invisible fugitive with the constraint that no two searchers can occupy the same node simultaneously. This variant seems promising for designing approximation algorithms for computing the pathwidth of graphs. The main contribution in [34], [61] is the characterization of trees where k searchers are necessary and sufficient to win. Our characterization leads to a polynomial-time algorithm to compute the minimum number of searchers needed in trees.

We also study graph searching in directed graphs. We prove that the graph processing variant is monotone which allows us to show its equivalence with a particular digraph decomposition [47].

6.2.2.2. Surveillance Game and Fractional Game.

A surprising application of some variant of pursuit-evasion games is the problem for a web-browser to download documents in advance while an internaut is surfing on the Web. In a previous work, we model this problem as a Pursuit-evasion game called Surveillance game. In [40], [67], we continue our study of the Surveillance game. We provide some bounds on the connected and online variants of this game. In particular, we show that, in the online variant (when the searchers discover the graph during the game), the best strategy is the trivial one that consists in downloading the document in the neighborhood of the position of the internaut.

In [69], [48], [52], we define a framework generalizing and relaxing many games (including the Surveillance game) where Players use fractions of their token at each turn. We design an algorithm for solving the fractional games. In particular, our algorithm runs in polynomial-time when the length of the game is bounded by 2 (in contrast, computing the surveillance game is NP-hard even when the game is limited to two turns). For some games, we also prove that the fractional variant provides some good approximation. This direction of research seems promising for solving many open problems related to Pursuit-evasion games.

6.2.2.3. Robots in anonymous networks.

Motivated by the understanding of the limits of distributed computing, we consider a recent model of robot-based computing which makes use of identical, memoryless mobile robots placed on nodes of anonymous graphs. The robots operate in Look-Compute-Move cycles that are performed asynchronously for each robot. In particular, we consider various problems such as graph exploration, graph searching and gathering in various graph classes. We provide a new distributed approach which turns out to be very interesting as it neither completely falls into symmetry-breaking nor into symmetry-preserving techniques. We proposed a general approach [38], [66] to solve the three problems in rings even in case of symmetric initial configurations.

6.2.3. Algorithm design in biology

In COATI, we have recently started a collaboration with EPI ABS (Algorithms Biology Structure) from Sophia Antipolis on “minimal connectivity complexes in mass spectrometry based macro-molecular complex reconstruction” [28], [55]. This problem turns out to be a minimum color covering problem (minimum number of colors to cover colored edges with connectivity constraints on the subgraphs induced by the colors) of the edges of a graph, and is surprizingly similar to a capacity maximization problem in a multi-interfaces radio network we were studying.

6.3. Structural Graph Theory

Participants: Julio Araújo, Jean-Claude Bermond, Frédéric Havet, Nicolas Nisse, Ana Karolinna Maia de Oliveira, Stéphane Pérennes.

6.3.1. Graph colouring and applications

Graph colouring is a central problem in graph theory and it has a huge number of applications in various scientific domains (telecommunications, scheduling, bio-informatics, ...). We mainly study graph colouring problems that model ressource allocation problems.

6.3.1.1. Backbone colouring

A well-known channel assignment problem is the following: we are given a graph G, whose vertices correspond to transmitters, together with an edge-weghting w. The weight of an edge corresponds to the minimum separation between the channels on its endvertices to avoid interferences. (If there is no edge, no separation is required, the transmitters do not interfere.) We need to assign positive integers (corresponding to channels) to the vertices so that for every edge e the channels assigned to its endvertices differ by at least $w(e)$. The goal is to minimize the largest integer used, which corresponds to minimizing the *span* of the used bandwidth.

We studied a particular, yet quite general, case, called *backbone colouring*, in which there are only two levels of interference. So we are given a graph G and a subgraph H, called *the backone*. Two adjacent vertices in H must get integers at least q apart, while adjacent vertices in G must get integers at distance at least 1. The minimum span is this case is called the *q-backbone chromatic number* and is denoted $BBC_q(G,H)$. Backbone forests in planar graphs are of particular interests. In [22], we prove that if G is planar and T is a tree of diameter at most 4, then $BBC_2(G,T) \leq 6$ hence giving an evidence to a conjecture of Broersma et al. [74] stating that the same holds if T has an arbitrary diameter.

6.3.1.2. Weighted colouring

We also studied weighted colouring which models various problems of shared resources allocation. Given a vertex-weighted graph G and a (proper) r -colouring $c = \{C_1, ..., C_r\}$ of G , the *weight* of a colour class C_i is the maximum weight of a vertex coloured i and the *weight* of c is the sum of the weights of its colour classes. The objective of the Weighted Colouring Problem is, given a vertex-weighted graph G , to determine the minimum weight of a proper colouring of G, that is, its *weighted chromatic number*. In [17], we prove that the Weighted Colouring Problem admits a version of Hajós' Theorem and so we show a necessary and sufficient condition for the weighted chromatic number of a vertex-weighted graph G to be at least k, for any positive real k. The Weighted Colouring Problem problem remains NP-complete in some particular graph classes as bipartite graphs. In their seminal paper [77], Guan and Zhu asked whether the weighted chromatic number of bounded tree-width graphs (partial k-trees) can be computed in polynomial-time. Surprisingly, the time-complexity of computing this parameter in trees is still open. We show [58] that, assuming the Exponential Time Hypothesis (3-SAT cannot be solved in sub-exponential time), the best algorithm to compute the weighted chromatic number of n-node trees has time-complexity $n^{\Theta(\log n)}$. Our result mainly relies on proving that, when computing an optimal proper weighted colouring of a graph G, it is hard to combine colourings of its connected components, even when G is a forest.

6.3.1.3. On-line colouring

Since many applications, and in particular channel assignment problems, must be solved on-line, we studied on-line colouring algorithms. The most basic and most widespread of them is the greedy algorithm. The largest number of colours that can be given by the greedy algorithm on some graph. is called its *Grundy number* and is denoted $\Gamma(G)$. Trivially $\Gamma(G) \leq \Delta(G) + 1$, where $\Delta(G)$ is the maximum degree of the graph. In [26], we show that deciding if $\Gamma(G) \leq \Delta(G)$ is NP-complete. We then show that deciding if $\Gamma(G) \geq |V(G)| - k$ is fixed-parameter tractable with respect to the parameter k. We also gave similar complexity results on b-colourings, which is a manner of improving colourings on-line.

In [27], we study a game version of greedy colouring. Given a graph G, two players, Alice and Bob, alternate their turns in choosing uncoloured vertices to be coloured. Whenever an uncoloured vertex is chosen, it is coloured by the least positive integer not used by any of its coloured neighbors. Alice's goal is to minimize the total number of colours used in the game, and Bob's goal is to maximize it. The *game Grundy number* of G is the number of colours used in the game when both players use optimal strategies. It is proved in this paper that the maximum game Grundy number of forests is 3, and the game Grundy number of any partial 2-tree is at most 7.

6.3.1.4. Enumerating edge-colourings and total colourings

With the success of moderately exponential algorithms, there is an increasing interest for enumeration problems, because of their own interest but also because they might be crucial to solve optimization problems. In [21], we are interested in computing the number of edge colourings and total colourings of a connected graph. We prove that the maximum number of k-edge-colourings of a connected k-regular graph on n vertices is $k \cdot ((k-1)!)^{n/2}$. Our proof is constructive and leads to a branching algorithm enumerating all the k-edge-colourings of a connected k-regular graph in time $O^*(((k-1)!)^{n/2})$ and polynomial space. In particular, we obtain a algorithm to enumerate all the 3-edge-colourings of a connected cubic graph in time $O^*(2^{n/2}) = O^*(1.4143^n)$ and polynomial space. This improves the running time of $O^*(1.5423^n)$ of the algorithm of Golovach et al. [76]. We also show that the number of 4-total-colourings of a connected cubic graph is at most $3 \cdot 2^{3n/2}$. Again, our proof yields a branching algorithm to enumerate all the 4-total-colourings of a connected cubic graph.

6.3.2. Directed graphs

Graph theory can be roughly partitioned into two branches: the areas of undirected graphs and directed graphs (digraphs). Even though both areas have numerous important applications, for various reasons, undirected graphs have been studied much more extensively than directed graphs. One of the reasons is that many problems for digraphs are much more difficult than their analogues for undirected graphs.

6.3.2.1. *Finding a subdivision of a digraph*

One of the cornerstones of modern (undirected) graph theory is minor theory of Robertson and Seymour. Unfortunately, we cannot expect an equivalent for directed graphs. Minor theory implies in particular that, for any fixed F, detecting a subdivision of a fixed graph F in an input graph G can be performed in polynomial time by the Robertson and Seymour linkage algorithm. In contrast, the analogous subdivision problem for digraph can be either polynomial-time solvable or NP-complete, depending on the fixed digraph F. In a previous paper, we gave a number of examples of polynomial instances, several NP-completeness proofs as well as a number of conjectures and open problems. In [71], we conjecture that, for every integer k greater than 1, the directed cycles of length at least k have the Erdős-Pósa Property : for every n, there exists an integer t_n such that for every digraph D, either D contains n disjoint directed cycles of length at least k, or there is a set T of t_n vertices that meets every directed cycle of length at least k. This generalizes a celebrated result of Reed, Robertson, Seymour and Thomas which is the case $k = 2$ of this conjecture. We prove the conjecture for $k = 3$. We also show that the directed k-Linkage problem is polynomial-time solvable for digraphs with circumference at most 2. From these two results, we deduce that if F is the disjoint union of directed cycles of length at most 3, then one can decide in polynomial time if a digraph contains a subdivision of F.

6.3.2.2. *Oriented trees in digraphs*

Let $f(k)$ be the smallest integer such that every $f(k)$-chromatic digraph contains every oriented tree of order k. Burr proved $f(k) \leq (k-1)^2$ in general, and he conjectured $f(k) = 2k - 2$. Burr also proved that every $(8k - 7)$-chromatic digraph contains every antidirected tree. We improve both of Burr's bounds. We show [15] that $f(k) \leq k^2/2 - k/2 + 1$ and that every antidirected tree of order k is contained in every $(5k - 9)$-chromatic digraph. We also make a conjecture that explains why antidirected trees are easier to handle. It states that if $|E(D)| > (k - 2)|V(D)|$, then the digraph D contains every antidirected tree of order k. This is a common strengthening of both Burr's conjecture for antidirected trees and the celebrated Erdős-Sós Conjecture. The analogue of our conjecture for general trees is false, no matter what function $f(k)$ is used in place of $k - 2$. We prove our conjecture for antidirected trees of diameter 3 and present some other evidence for it. Along the way, we show that every acyclic k-chromatic digraph contains every oriented tree of order k and suggest a number of approaches for making further progress on Burr's conjecture.

COFFEE Project-Team (section vide)

COMETE Project-Team

6. New Results

6.1. Foundations of information hiding

Information hiding refers to the problem of protecting private information while performing certain tasks or interactions, and trying to avoid that an adversary can infer such information. This is one of the main areas of research in Comète; we are exploring several topics, described below.

6.1.1. Differential privacy with general metrics.

Differential privacy can be interpreted as a bound on the distinguishability of two generic databases, which is determined by their Hamming distance: the distance in the graph determined by the adjacency relation (two databases are adjacent if they differ for one individual).

In [21] we lifted the restriction relative to the Hamming graphs and we explored the implications of differential privacy when the indistinguishability requirement depends on an arbitrary notion of distance. We showed that we can express, in this way, (protection against) kinds of privacy threats that cannot be naturally represented with the standard notion. We gave an intuitive characterization of these threats in terms of Bayesian adversaries, which generalizes the characterization of (standard) differential privacy from the literature. Next, we revisited the well-known result on the non-existence of universally optimal mechanisms for any query other than counting queries. We showed that in our setting, for certain kinds of distances, there are many more queries for which universally optimal mechanisms exist: Notably sum, average, and percentile queries. Finally, we showed some applications in various domains: statistical databases where the units of protection are groups (rather than individuals), geolocation, and smart metering.

6.1.2. Privacy for location-based services.

The growing popularity of location-based services, allowing unknown/untrusted servers to easily collect and process huge amounts of users' information regarding their location, has recently started raising serious concerns about the privacy of this kind of sensitive information. In [19] we studied geo-indistinguishability, a formal notion of privacy for location-based services that protects the exact location of a user, while still allowing approximate information - typically needed to obtain a certain desired service - to be released.

Our privacy definition formalizes the intuitive notion of protecting the user's location within a radius r with a level of privacy that depends on r. We presented three equivalent characterizations of this notion, one of which corresponds to a generalized version [21] of the well-known concept of differential privacy. Furthermore, we presented a perturbation technique for achieving geo-indistinguishability by adding controlled random noise to the user's location, drawn from a planar Laplace distribution. We demonstrated the applicability of our technique through two case studies: First, we showed how to enhance applications for location-based services with privacy guarantees by implementing our technique on the client side of the application. Second, we showed how to apply our technique to sanitize location-based sensible information collected by the US Census Bureau.

6.1.3. Relation between differential privacy and quantitative information flow.

Differential privacy is a notion that has emerged in the community of statistical databases, as a response to the problem of protecting the privacy of the database's participants when performing statistical queries. The idea is that a randomized query satisfies differential privacy if the likelihood of obtaining a certain answer for a database x is not too different from the likelihood of obtaining the same answer on adjacent databases, i.e. databases which differ from x for only one individual.

In [13], we analyzed critically the notion of differential privacy in light of the conceptual framework provided by the Rényi min information theory. We proved that there is a close relation between differential privacy and leakage, due to the graph symmetries induced by the adjacency relation. Furthermore, we considered the utility of the randomized answer, which measures its expected degree of accuracy. We focused on certain kinds of utility functions called “binary”, which have a close correspondence with the Rényi min mutual information. Again, it turns out that there can be a tight correspondence between differential privacy and utility, depending on the symmetries induced by the adjacency relation and by the query. Depending on these symmetries we can also build an optimal-utility randomization mechanism while preserving the required level of differential privacy. Our main contribution was a study of the kind of structures that can be induced by the adjacency relation and the query, and how to use them to derive bounds on the leakage and achieve the optimal utility.

6.1.4. A differentially private mechanism of optimal utility for a region of priors

Differential privacy (already introduced in the previous sections) is usually achieved by using mechanisms that add random noise to the query answer. Thus, privacy is obtained at the cost of reducing the accuracy, and therefore the utility, of the answer. Since the utility depends on the user’s side information, commonly modeled as a prior distribution, a natural goal is to design mechanisms that are optimal for every prior. However, it has been shown in the literature that such mechanisms do not exist for any query other than counting queries.

Given the above negative result, in [22] we considered the problem of identifying a restricted class of priors for which an optimal mechanism does exist. Given an arbitrary query and a privacy parameter, we geometrically characterized a special region of priors as a convex polytope in the priors space. We then derived upper bounds for utility as well as for min-entropy leakage for the priors in this region. Finally we defined what we call the tight-constraints mechanism and we discussed the conditions for its existence. This mechanism has the property of reaching the bounds for all the priors of the region, and thus it is optimal on the whole region.

6.1.5. Compositional analysis of information hiding

Systems concerned with information hiding often use randomization to obfuscate the link between the observables and the information to be protected. The degree of protection provided by a system can be expressed in terms of the probability of error associated to the inference of the secret information. In [14] we considered a probabilistic process calculus to specify such systems, and we studied how the operators affect the probability of error. In particular, we characterized constructs that have the property of not decreasing the degree of protection, and that can therefore be considered safe in the modular construction of these systems. As a case study, we applied these techniques to the Dining Cryptographers, and we derived a generalization of Chaum’s strong anonymity result.

In [26], a similar framework was proposed for reasoning about the degree of differential privacy provided by such systems. In particular, we investigated the preservation of the degree of privacy under composition via the various operators. We illustrated our idea by proving an anonymity-preservation property for a variant of the Crowds protocol for which the standard analyses from the literature are inapplicable. Finally, we made some preliminary steps towards automatically computing the degree of privacy of a system in a compositional way.

6.1.6. Preserving differential privacy under finite-precision semantics

The approximation introduced by finite-precision representation of continuous data can induce arbitrarily large information leaks even when the computation using exact semantics is secure. Such leakage can thus undermine design efforts aimed at protecting sensitive information. For instance, the standard approach to achieve differential privacy (introduced in previous sections) is the addition of noise to the true (private) value. To date, this approach has been proved correct only in the ideal case in which computations are made using an idealized, infinite-precision semantics. In [23], we analyzed the situation at the implementation level, where the semantics is necessarily finite-precision, i.e. the representation of real numbers and the operations on them are rounded according to some level of precision. We showed that in general there are violations of the differential privacy property, and we studied the conditions under which we can still guarantee a limited (but, arguably, totally acceptable) variant of the property, under only a minor degradation of the privacy level. Finally, we illustrated our results on two cases of noise-generating distributions: the standard Laplacian mechanism commonly used in differential privacy, and a bivariate version of the Laplacian recently introduced in the setting of privacy-aware geolocation.

6.1.7. Metrics for differential privacy in concurrent systems

Many protocols for protecting confidential information have involved randomized mechanisms and a nondeterministic behavior (such as the Dining Cryptographers protocol or the Crowds protocol). In [28], we investigate techniques for proving differential privacy in the context of concurrent systems which contain both probabilistic and nondeterministic behaviors. Our motivation stems from the work of Tschantz et al., who proposed a verification method based on proving the existence of a stratified family of bijections between states, that can track the privacy leakage, ensuring that it does not exceed a given leakage budget. We improve this technique by investigating state properties which are more permissive and still imply differential privacy. We consider three pseudometrics on probabilistic automata: The first one is essentially a reformulation of the notion proposed by Tschantz et al. The second one is a more liberal variant, still based on the existence of a family of bijections, but relaxing the relation between them by integrating the notion of amortization, which results into a more parsimonious use of the privacy budget. The third one aims at relaxing the bijection requirement, and is inspired by the Kantorovich-based bisimulation metric proposed by Desharnais et al. We cannot adopt the latter notion directly because it does not imply differential privacy. Thus we propose a multiplicative variant of it, and prove that it is still an extension of weak bisimulation. We show that for all the pseudometrics the level of differential privacy is continuous on the distance between the starting states, which makes them suitable for verification. Moreover we formally compare these three pseudometrics, proving that the latter two metrics are indeed more permissive than the first one, but incomparable with each other, thus constituting two alternative techniques for the verification of differential privacy.

6.1.8. Unlinkability

Unlinkability is a privacy property of crucial importance for several systems (such as RFID or voting systems). Informally, unlinkability states that, given two events/items in a system, an attacker is not able to infer whether they are related to each other. However, in the literature we find several definitions for this notion, which are apparently unrelated and shows a potentially problematic lack of agreement. In [20] we shed new light on unlinkability by comparing different ways of defining it and showing that in many practical situations the various definitions coincide. It does so by (a) expressing in a unifying framework four definitions of unlinkability from the literature (b) demonstrating how these definitions are different yet related to each other and to their dual notion of "inseparability" and (c) by identifying conditions under which all these definitions become equivalent. We argued that the conditions are reasonable to expect in identification systems, and we prove that they hold for a generic class of protocols.

6.1.9. Trust in anonymity networks

Trust metrics are used in anonymity networks to support and enhance reliability in the absence of verifiable identities, and a variety of security attacks currently focus on degrading a user's trustworthiness in the eyes of the other users. In [16] we have presented an enhancement of the Crowds anonymity protocol via a notion of trust which allows crowd members to route their traffic according to their perceived degree of trustworthiness of each other member of the crowd. Such trust relations express a measure of an individual's belief that another user may become compromised by an attacker, either by a direct attempt to corrupt or by a denial-of-service attack. Our protocol variation has the potential of improving the overall trustworthiness of data exchanges in anonymity networks, which cannot normally be taken for granted in a context where users are actively trying to conceal their identities. Using such formalization, in the paper we have then analyzed quantitatively the privacy properties of the protocol under standard and adaptive attacks.

6.2. Foundations of Concurrency

Distributed systems have changed substantially in the recent past with the advent of phenomena like social networks and cloud computing. In the previous incarnation of distributed computing the emphasis was on consistency, fault tolerance, resource management and related topics; these were all characterized by *interaction between processes*. Research proceeded along two lines: the algorithmic side which dominated the Principles Of Distributed Computing conferences and the more process algebraic approach epitomized by CONCUR where the emphasis was on developing compositional reasoning principles. What marks the new era of distributed systems is an emphasis on managing access to information to a much greater degree than before.

6.2.1. Models and Emerging Trends of Concurrent Constraint Programming

The *Concurrent constraint programming (ccp)* paradigm focuses on information access and therefore it is suited for this new era of concurrent systems. Ccp singles out the fundamental aspects of asynchronous systems whose agents (or processes) evolve by accessing information in a global medium, represented as constraints over the variables of the system. Agents communicate by posting and querying partial information in the medium. This covers a vast variety of systems as those arising in biological phenomena, reactive systems, net- centric computing and the advent of social networks and cloud computing. In [17] we surveyed the main applications, developments and current trends of ccp.

6.2.2. Efficient computation of program equivalence for confluent concurrent constraint programming

The development of algorithms and automatic verification procedures for ccp have hitherto been far too little considered. To the best of our knowledge there is only one existing verification algorithm for the standard notion of ccp program (observational) equivalence. In [25] we first showed that this verification algorithm has an exponential-time complexity even for programs from a representative sub-language of ccp; the summation-free fragment (ccp+). We then significantly improved on the complexity of this algorithm by providing two alternative polynomial-time decision procedures for ccp+ program equivalence. Each of these two procedures has an advantage over the other. One has a better time complexity. The other can be easily adapted for the full language of ccp to produce significant state space reductions. The relevance of both procedures derives from the importance of ccp+. This fragment, which has been the subject of many theoretical studies, has strong ties to first-order logic and an elegant denotational semantics, and it can be used to model real-world situations. Its most distinctive feature is that of confluence, a property we exploit to obtain our polynomial procedures.

6.2.3. Abstract Interpretation of Temporal Concurrent Constraint Programs

Timed concurrent constraint programming (tcc) is a declarative model for concurrency offering a logic for specifying reactive systems, i.e. systems that continuously interact with the environment. The universal tcc formalism (utcc) is an extension of tcc with the ability to express mobility. Here mobility is understood as communication of private names as typically done for mobile systems and security protocols. In [15] we considered the denotational semantics for tcc, and we extended it to a "collecting" semantics for utcc based on closure operators over sequences of constraints. Relying on this semantics, we formalized a general framework for data flow analyses of tcc and utcc programs by abstract interpretation techniques. The concrete and abstract semantics we proposed are compositional, thus allowing us to reduce the complexity of data flow analyses. We showed that our method is sound and parametric with respect to the abstract domain. Thus, different analyses can be performed by instantiating the framework. We illustrated how it is possible to reuse abstract domains previously defined for logic programming to perform, for instance, a groundness analysis for tcc programs. We showed the applicability of this analysis in the context of reactive systems. Furthermore, we made also use of the abstract semantics to exhibit a secrecy flaw in a security protocol. We also showed how it is possible to make an analysis which may show that tcc programs are suspension free. This can be useful for several purposes, such as for optimizing compilation or for debugging.

6.2.4. Foundations of Probabilistic Concurrent Systems

In [24] we introduced a formal proof system for compositional verification of probabilistic concurrent processes. Properties are expressed using a probabilistic modal μ-calculus, and the proof system is formulated as a sequent calculus in which sequents are given a quantitative interpretation. A key feature is that the probabilistic scenario is handled by introducing the notion of Markov proof, by which each proof in the system is interpreted as a Markov Decision Process, with the proof only considered valid in the case that the value of the MDP is zero.

COMMANDS Project-Team

5. New Results

5.1. Optimality conditions in Pontryagin form for optimal control problems

Participants: Joseph Frédéric Bonnans, Xavier Dupuis, Laurent Pfeiffer.

5.1.1. Necessary conditions

In the paper [31], we state and prove first- and second-order necessary conditions in Pontryagin form for optimal control problems with pure state and mixed control-state constraints. We say that a Lagrange multiplier of an optimal control problem is a Pontryagin multiplier if it is such that Pontryagin's minimum principle holds, and we call optimality conditions in Pontryagin form those which only involve Pontryagin multipliers. Our conditions rely on a technique of partial relaxation, and apply to Pontryagin local minima.

5.1.2. Sufficient conditions

In the paper [32], we consider sufficient conditions. More precisely, given a reference feasible trajectory of an optimal control problem, we say that the quadratic growth property for bounded strong solutions holds if the cost function of the problem has a quadratic growth over the set of feasible trajectories with a bounded control and with a state variable sufficiently close to the reference state variable. Our sufficient second-order optimality conditions in Pontryagin form ensure this property and ensure *a fortiori* that the reference trajectory is a bounded strong solution. Our proof relies on a decomposition principle, which is a particular second-order expansion of the Lagrangian of the problem.

5.1.3. Shooting Approach to Optimal Control Problems

Participant: Joseph Frédéric Bonnans.

In the paper [24] we give an overview of the shooting technique for solving deterministic optimal control problems. This approach allows to reduce locally these problems to a finite dimensional equation. We first recall the basic idea, in the case of unconstrained or control constrained problems, and show the link with second-order optimality conditions and the analysis or discretization errors. Then we focus on two cases that are now better understood: state constrained problems, and affine control systems. We end by discussing extensions to the optimal control of a parabolic equation.

5.2. Applications of deterministic optimal control problems

5.2.1. Optimization of running strategies based on anaerobic energy and variations of velocity

Participant: Joseph Frédéric Bonnans.

In the report [29] we present new models, numerical simulations and rigorous analysis for the optimization of the velocity in a race. In a seminal paper, Keller [74], [75] explained how a runner should determine his speed in order to run a given distance in the shortest time. We extend this analysis, based on the equation of motion and aerobic energy, to include a balance of anaerobic energy (or accumulated oxygen deficit) and an energy recreation term when the speed decreases. We also take into account that when the anaerobic energy gets too low, the oxygen uptake cannot be maintained to its maximal value. Using optimal control theory, we obtain a proof of Keller's optimal race, and relate the problem to a relaxed formulation, where the propulsive force represents a probability distribution rather than a value function of time. Our analysis leads us to introduce a bound on the variations of the propulsive force to obtain a more realistic model which displays oscillations of the velocity. Our numerical simulations qualitatively reproduce quite well physiological measurements on real runners. We show how, by optimizing over a period, we recover these oscillations of speed. We point out that our numerical simulations provide in particular the exact instantaneous anaerobic energy used in the exercise.

5.2.2. Optimal control of leukemic cell population dynamics

Participant: Xavier Dupuis.

In the paper [33] we discuss the optimal co-administration of two drugs for some acute myeloid leukemias (AML), and we are looking for in vitro protocols as a first step. This issue can be formulated as an optimal control problem. The dynamics of leukemic cell populations in culture is given by age-structured partial differential equations, which can be reduced to a system of delay differential equations, and where the controls represent the action of the drugs. The objective function relies on eigenelements of the uncontrolled model and on general relative entropy, with the idea to maximize the efficiency of the protocols. The constraints take into account the toxicity of the drugs. We present in this paper the modeling aspects, as well as theoretical and numerical results on the optimal control problem that we get.

5.2.3. Contrast imaging problem in nuclear magnetic resonance

Participant: Pierre Martinon.

In collaboration with team McTAO (Sophia), we studied in [25] and [36] the contrast imaging problem in nuclear magnetic resonance, modeled as Mayer problem in optimal control. The optimal solution can be found as an extremal, solution of the Maximum Principle and analyzed with the techniques of geometric control. A first synthesis of locally optimal solutions is given in the single-input case, with some preliminary results in the bi-input case. We conducted a comprehensive numerical investigation of the problem, using a combination of indirect shooting (HAMPATH software) and direct method (BOCOP), with a moment-based (LMI) technique to estimate the global optimum.

Figure 2. Contrast in quantum control for NMR - Oxygenated / deoxygenated blood

5.2.4. Optimizing the anaerobic digestion of microalgae in a coupled process

Participant: Pierre Martinon.

In collaboration with the Inra-Inria team MODEMIC (Montpellier), we studied in [30] a bio-reactor system describing the coupling of a culture of micro-algae and an anaerobic digester. Our aim is to optimize the production of methane in the digester during a certain number of days with respect to the dilution rate (the input flow of micro-algae in the digester). The mathematical model for the dynamics of the two reactors takes into account a periodic day-night model of the light in the culture of micro-algae, and a chemostat model for the digester. We first prove existence and attraction of periodic solutions for a one day period, and we apply Pontryagin's Maximum Principle (PMP) in order to characterize optimal controls. We provide numerical simulations for different light models, by a direct method that we refine using an indirect shooting. We also investigate the dependence of the optimal cost with respect to the ratio of the volumes of the two tanks. Finally, we investigate the optimal strategies over a large number of days without periodic constraints, and compared the mean cost to the optimal cost over one period.

../../../../projets/commands/IMG/attraction.png

Figure 3. Coupled bio-reactor for micro-algae digestion - Attraction property

5.2.5. Design of optimal experiments for parameter estimation of microalgae growth models

Participant: Pierre Martinon.

In collaboration with team BIOCORE (Sophia), we investigated in [27] techniques of Optimal Experiment Design for microalgae growth models. In order to have microalgae growth models that are useful for prediction and process optimization, reliable parameters need to be provided. This reliability implies a careful design of experiments that can be exploited for parameter estimation. OED techniques can provide guidelines for the design of experiments with high informative content that allow an accurate parameter estimation. We study a real experimental device devoted to evaluate the effect of temperature and light on microalgae growth. On the basis of a mathematical model of the experimental system, the optimal experiment design problem was solved as an optimal control problem. E-optimal experiments were obtained by using two discretization approaches, namely sequential and simultaneous. The results showed that an adequate parameterization of the experimental inputs provided optimal solutions very close to those provided by the simultaneous discretization. Simulation results showed the relevance of determining optimal experimental inputs for achieving an accurate parameter estimation.

../../../../projets/commands/IMG/TIP.png

Figure 4. Experimental apparatus for the study of micro-algae growth (Ifremer)

5.2.6. Controllability and optimal strokes for N-link microswimmer

Participant: Pierre Martinon.

In [39] we focus on the N-link swimmer, a generalization of the classical Purcell swimmer. We use the simplification of the Resistive Force Theory to derive the motion equation for the swimmer in a fluid with a low Reynolds number. We prove that the swimmer is controllable in the whole plane when it is composed by more than 3 sticks and for almost every set of stick lengths. As a direct result, we show that there exists an optimal swimming strategy which leads to minimize the time to reach a desired configuration. Numerical experiments on the case of the Purcell swimmer suggest that the optimal strategy is periodic, i.e. composed of a sequence of identical strokes. Our results indicate that this candidate for an optimal stroke indeed gives a better speed than the classical Purcell stroke. Future directions for this work include the design of robotic micro-swimmers, as well as investigation of the movement of swimming micro-organisms.

5.3. Hamilton-Jacobi (HJ) approach

5.3.1. Dynamic programming and error estimates for stochastic control with Max cost

Participants: Olivier Bokanowski, Athena Picarelli, Hasnaa Zidani.

../../../../projets/commands/IMG/swimmer_xf.png

Figure 5. Purcell (3-link) swimmer - Purcell vs optimal stroke

The paper [35] is concerned with stochastic optimal control for a running maximum cost. A direct approach based on dynamic programming techniques is studied leading to the characterization of the value function as the unique viscosity solution of a second order Hamilton-Jacobi-Bellman (HJB) equation with an oblique derivative boundary condition. A general numerical scheme is proposed and a convergence result is provided. Error estimates are obtained for the semi-Lagrangian scheme. These results can apply to the case of lookback options in finance. Moreover, optimal control problems with maximum cost arise in the characterization of the reachable sets for a system of controlled stochastic differential equations. Some numerical simulations on examples of reachable analysis are included to illustrate our approach.

5.3.2. *Optimal feedback control of undamped wave equations by solving a HJB equation*

Participant: Hasnaa Zidani.

An optimal finite-time horizon feedback control problem for (semi linear) wave equations is studied in [42]. The feedback law can be derived from the dynamic programming principle and requires to solve the evolutionary Hamilton-Jacobi-Bellman (HJB) equation. Classical discretization methods based on finite elements lead to approximated problems governed by ODEs in high dimensional space which makes infeasible the numerical resolution by HJB approach. In the present paper, an approximation based on spectral elements is used to discretize the wave equation. The effect of noise is considered and numerical simulations are presented to show the relevance of the approach

5.3.3. *Transmission conditions on interfaces for Hamilton-Jacobi-Bellman equations*

Participants: Hasnaa Zidani, Zhiping Rao.

The works [43], [28] deal with deterministic control problems where the dynamic and the running cost can be completely different in two (or more) complementary domains of the space IR^N. As a consequence, the dynamics and running cost present discontinuities at the interfaces of these domains. This leads to a complex interplay that has to be analyzed among transmission conditions to "glue" the propagation of the value function on the interfaces. Several questions arise: how to define properly the value function(s) and what is (are) the right Bellman Equation(s) associated with this problem?. In the case of a simple geometry (namely when the space IR^N is partitioned into two subdomains separated with an interface which is assumed to be a regular hypersurface without any connectedness requirement), [43] discuss different conditions on the hyperplane where the dynamic and the running cost are discontinuous, and the uniqueness properties of the Bellman problem are studied. In this paper it is used a dynamical approach, namely instead of working with test functions, the accent is put on invariance properties of an augmented dynamics related to the integrated control system. The comparison principle is accordingly based, rather than on (semi)continuity of the Hamiltonian appearing in the Hamilton–Jacobi–Bellman equation, on some weak separation properties of this dynamics with respect to the stratification. A more general situation where the space is partionned on several domains is also analyzed in [28].

5.3.4. *Singular perturbation of optimal control problems on multi-domains*

Participants: Nicolas Forcadel, Hasnaa Zidani.

The goal of the paper[38] is to study a singular perturbation problem in the framework of optimal control on multi-domains. We consider an optimal control problem in which the controlled system contains a fast and a slow variables. This problem is reformulated as an Hamilton-Jacobi-Bellman (HJB) equation. The main difficulty comes from the fact that the fast variable lives in a multi-domain. The geometric singularity of the multi-domains leads to the discontinuity of the Hamiltonian. Under a controllability assumption on the fast variables, the limit equation (as the velocity of the fast variable goes to infinity) is obtained via a PDE approach and by means of the tools of the control theory.

5.3.5. *Optimal control of first order HJ equations with linearly bounded Hamiltonian*

Participant: Philip Graber.

In [40], we consider the optimal control of solutions of first order Hamilton-Jacobi equations, where the Hamiltonian is convex with linear growth. This models the problem of steering the propagation of a front by constructing an obstacle. We prove existence of minimizers to this optimization problem as in a relaxed setting and characterize the minimizers as weak solutions to a mean field game type system of coupled partial differential equations. Furthermore, we prove existence and partial uniqueness of weak solutions to the PDE system. An interpretation in terms of mean field games is also discussed.

5.3.6. Zubov's equation for state-constrained perturbed nonlinear systems

Participant: Hasnaa Zidani.

The paper [41] gives a characterization of the uniform robust domain of attraction for a nite non-linear controlled system subject to perturbations and state constraints. We extend the Zubov approach to characterize this domain by means of the value function of a suitable in nite horizon state-constrained control problem which at the same time is a Lyapunov function for the system. We provide associated Hamilton-Jacobi-Bellman equations and prove existence and uniqueness of the solutions of these generalized Zubov equations.

5.3.7. Numerical methods for chance-constrained stochastic optimal control problems

Participant: Laurent Pfeiffer.

In Laurent Pfeiffer's PhD, we study stochastic optimal control problems with a probability constraint on the final state. This constraint must be satisfied with a probability greater or equal than a given level. We analyse and compare two approaches for discrete-time problems: a first one based on a dynamic programming principle and a second one using Lagrange relaxation. These approaches can be used for continuous-time problems, for which we give numerical illustrations.

COMPSYS Project-Team

6. New Results

6.1. Parameterized Construction of Program Representations for Sparse Dataflow Analysiss

Participants: André Tavares [UFMG, Belo Horizonte, Brazil], Benoit Boissinot [Ex-Compsys, Google Zurich], Fernando Magno Quintão Pereira [UFMG, Belo Horizonte, Brazil], Fabrice Rastello.

Data-flow analysis usually associates information with control flow regions. Informally, if these regions are too small like a point between two consecutive statements, we call the analysis dense. On the other hand, if these regions include many such points, then we call it sparse. This work presents a systematic method to build program representations that support sparse analyses. To pave the way to this framework, we clarify the literature about well-known intermediate program representations. We show that our approach, subsumes, up to parameter choices, many of these representations, such as the SSA, SSI, and e-SSA forms. In particular, our algorithms are faster, simpler and more frugal than the previous techniques used to construct SSI (static single information) form programs. We produce intermediate representations isomorphic to Choi *et al.*'s sparse evaluation graphs (SEG) for the family of data-flow problems that can be partitioned by variables. However, contrary to SEGs, we can handle - sparsely - problems that are not in this family. We have tested our ideas in the LLVM compiler, comparing different program representations in terms of size and construction time.

This work is part of the collaboration with UFMG (see Section 8.4) and has been accepted for presentation and publication at CC'14 (Compiler Construction Conference) [9].

6.2. A Framework for Enhancing Data Reuse via Associative Reordering

Participants: Kevin Stock [OSU, Columbus, USA], Louis-Noël Pouchet [UCLA, Los Angeles, USA], Fabrice Rastello, J. Ramanujam [LSU, Houston, USA], P. Sadayappan [OSU, Columbus, USA].

The freedom to reorder computations involving associative operators has been widely recognized and exploited in designing parallel algorithms and to a more limited extent in optimizing compilers. However, the use of associative reordering for enhancing data locality has not been previously explored to our knowledge.

In this work, we develop a novel framework for utilizing associativity of operations in regular loop computations to enhance register reuse. Stencils represent a particular class of important computations where our optimization framework can be applied to enhance performance. We use a multi-dimensional retiming formalism to characterize the space of valid transformations and to generate the transformed code. Experimental results demonstrate the effectiveness of the framework.

This work has been submitted to PLDI'14 and is part of the collaboration with P. Sadayappan from the University of Columbus (OSU) (see Section 8.4).

6.3. Function Cloning Revisited

Participants: Matheus Vilela [UFMG, Belo Horizonte, Brazil], Guilherme Balena [UFMG, Belo Horizonte, Brazil], Guilherme Marques [UFMG, Belo Horizonte, Brazil], Fernando Magno Quintão Pereira [UFMG, Belo Horizonte, Brazil], Fabrice Rastello.

Compilers rely on two main techniques to implement optimizations that depend on the calling context of functions: inlining and cloning. Historically, function inlining has seen more widespread use, as it tends to be more effective in practice. Yet, function cloning provides benefits that inline leaves behind. In particular, cloning gives the program developer a way to fight performance bugs, because it generates reusable code. Furthermore, it deals with recursion more naturally. Finally, it might lead to less code expansion, the inlining's nemesis.

In this work, we revisited function cloning under the light of these benefits. We discuss four independent code specialization techniques based on function cloning, which, although simple, find wide applicability, even in highly optimized benchmarks, such as SPEC CPU 2006. We claim that our optimizations are easy to implement and to deploy. We use Wu and Larus's well-known static profiling heuristic to measure the profitability of a clone. This metric gives us a concrete way to point out to program developers potential performance bugs, and gives us a metric to decide if we should keep a clone or not. By implementing our ideas in LLVM, we have been able to speed up some of the SPEC benchmarks by up to 6% on top of the -O2 optimization level.

This work is part of the collaboration with UFMG (see Section 8.4) and was also done in the context of the collaboration with Kalray and the ManycoreLabs project (see Section 7.2).

6.4. Register Allocation and Promotion through Combined Instruction Scheduling, Loop Splitting and Unrolling

Participants: P. Sadayappan [OSU, Columbus, USA], Fabrice Rastello, Lukasz Domanaga.

Register allocation is a much studied problem. A particularly important context for optimizing register allocation is within loops, since a significant fraction of the execution time of programs is often inside loop code. A variety of algorithms have been proposed in the past for register allocation, but the complexity of the problem has resulted in a decoupling of several important aspects, including loop unrolling, loop fission, register promotion, and instruction reordering.

In this work, we develop an approach to register allocation and promotion in a unified optimization framework that simultaneously considers the impact of loop unrolling, loop splitting, and instruction scheduling. This is done via a novel instruction tiling approach where instructions within a loop are represented along one dimension and innermost loop iterations along the other dimension. By exploiting the regularity along the loop dimension, and a constrained intra-tile execution order, the problem of optimizing register pressure is cast in a constraint programming formalism. Experimental results are provided from thousands of innermost loops extracted from the SPEC benchmarks, demonstrating improvements over the current state of the art.

This work is part of the collaboration with OSU (see Section 8.4) and was also done in the context of the collaboration with Kalray and the ManycoreLabs project (see Section 7.2). It contributes to the developments of the Tirex toolbox (see 5.17). It has also been submitted to PLDI'14.

6.5. Beyond Reuse Distance Analysis: Dynamic Analysis for Characterization of Data Locality Potential

Participants: Naznin Fauzia [OSU, Columbus, USA], Venmugil Elango [OSU, Columbus, USA], Mahesh Ravishankar [OSU, Columbus, USA], J. (ram) Ramanujam [LSU, Houston, USA], Fabrice Rastello, Atanas Rountev [OSU, Columbus, USA], Louis-Noël Pouchet [UCLA, Los Angeles, USA], P. Sadayappan [OSU, Columbus, USA].

Emerging computer architectures will feature drastically decreased flops/byte (ratio of peak processing rate to memory bandwidth) as highlighted by recent studies on Exascale architectural trends. Further, flops are getting cheaper while the energy cost of data movement is increasingly dominant. The understanding and characterization of data locality properties of computations is critical in order to guide efforts to enhance data locality.

Reuse distance analysis of memory address traces is a valuable tool to perform data locality characterization of programs. A single reuse distance analysis can be used to estimate the number of cache misses in a fully associative LRU cache of any size, thereby providing estimates on the minimum bandwidth requirements at different levels of the memory hierarchy to avoid being bandwidth bound. However, such an analysis only holds for the particular execution order that produced the trace. It cannot estimate potential improvement in data locality through dependence preserving transformations that change the execution schedule of the operations in the computation.

In this work, we develop a novel dynamic analysis approach to characterize the inherent locality properties of a computation and thereby assess the potential for data locality enhancement via dependence preserving transformations. The execution trace of a code is analyzed to extract a computational directed acyclic graph (CDAG) of the data dependences. The CDAG is then partitioned into convex subsets, and the convex partitioning is used to reorder the operations in the execution trace to enhance data locality. The approach enables us to go beyond reuse distance analysis of a single specific order of execution of the operations of a computation in characterization of its data locality properties. It can serve a valuable role in identifying promising code regions for manual transformation, as well as assessing the effectiveness of compiler transformations for data locality enhancement. We demonstrate the effectiveness of the approach using a number of benchmarks, including case studies where the potential shown by the analysis is exploited to achieve lower data movement costs and better performance.

This work is part of the collaboration with OSU (see Section 8.4) and has been accepted for publication at ACM TACO [2].

6.6. Characterizing the Inherent Data Movement Complexity of Computations via Lower Bounds

Participants: P. Sadayappan [OSU, Columbus, USA], Venmugil Elango [OSU, Columbus, USA], J. (ram) Ramanujam [LSU, Houston, USA], Louis-Noël Pouchet [UCLA, Los Angeles, USA], Fabrice Rastello.

Technology trends will cause data movement to account for the majority of energy expenditure and execution time on emerging computers. Therefore, computational complexity will no longer be a sufficient metric for comparing algorithms, and a fundamental characterization of data access complexity will be increasingly important. Although the problem of characterizing data access complexity has been modeled previously using the formalism of Hong & Kung's red/blue pebble game [27], applicability of previously-developed approaches has been extremely limited. We improve on prior work in several ways: 1) we develop an approach to composing lower bounds from arbitrary decompositions of computational directed acyclic graphs, thereby eliminating a significant limitation of previous approaches that required homogeneity of analyzed computations, 2) we develop a complementary graph min-cut based strategy to Hong & Kung's S-partitioning approach, and 3) we develop an automated approach to generate concrete I/O lower bounds of arbitrary, possibly irregular computational directed acyclic graphs. We provide experimental results demonstrating the utility of the developed approach.

This work has been submitted to PLDI'14 and is part of an informal collaboration with P. Sadayappan from the University of Columbus (OSU) (see Section 8.4).

6.7. Enhancing the Compilation of Synchronous Dataflow Programs

Participants: Paul Feautrier, Abdoulaye Gamatié [LIRMM, Montpellier], Laure Gonnord.

In this work [12], which is an extension of [26], we propose an enhancement of the compilation of synchronous programs with a combined numerical-Boolean abstraction. While our approach applies to synchronous dataflow languages in general, here, we consider the SIGNAL language for illustration. In the new abstraction, every signal in a program is associated with a pair of the form (`clock`, `value`), where `clock` is a Boolean function and `value` is a Boolean or numeric function. Given the performance level reached by recent progress in satisfiability modulo theory (SMT), we use an SMT solver to reason on this abstraction. Through sample examples, we show how our solution is used to determine absence of reaction captured by empty clocks; mutual exclusion captured by two or more clocks whose associated signals never occur at the same time; or hierarchical control of component activations via clock inclusion. We also show that the analysis improves the quality of the code generated automatically by a compiler, e.g., a code with smaller footprint, or a code executed more efficiently thanks to optimizations enabled by the new abstraction. The implementation of the whole approach includes a translator of synchronous programs towards the standard input format of SMT solvers, and an ad hoc SMT solver that integrates advanced functionalities to cope with the issues of interest in this work. These results have been published in 2013 (but considered as published in 2012) in the CSI Journal of Computing [24].

6.8. Synthesis of Ranking Functions using Extremal Counter-Examples

Participants: David Monniaux [Verimag, Grenoble], Lucas Séguinot [Student at ENS Cachan Bretagne], Laure Gonnord.

In [14], we presented a new algorithm adapted from scheduling techniques to synthesize (multi-dimensional) affine functions from general flowcharts programs. But, as for other methods, our algorithm tried to solve linear constraints on each control point and each transition, which can lead to quasi-untractable linear programming instances.

In contrast to these approaches, we proposed a new algorithm based on the following observations:

- Searching for ranking functions for loop headers is sufficient to prove termination.
- Furthermore, there exist loops such that there is a linear lexicographic ranking function that decreases along each path inside the loop, from one loop iteration to the next, but such that there is no lexicographic linear ranking function that decreases at each step along these paths. For these reasons, it is tempting to treat each path inside a loop as a single transition.

Unfortunately the number of paths may be exponential in the size of the program, thus the constraint system may become very large, even though it features fewer variables. To face this theoretical complexity, even though the number of paths may be large, we argue that, in practice, few of them actually matter in the constraint system (we formalize this concept by giving a characterization as geometric extremal points). Our algorithm therefore builds the constraint system lazily, taking paths into account *on demand*.

We are currently testing our preliminary implementation and submitting a paper on these new results.

6.9. Data-Aware Process Networks

Participants: Christophe Alias, Alexandru Plesco.

The following results concern the applied research activities directly linked to the Zettice start-up (see Section 7.3), which aims at applying polyhedral techniques to high-level circuit synthesis (HLS). Following the guidelines of Inria DTI, as this research aims to be transferred, these results are not published before being “protected” or exploited. An Inria patent deposit is currently processed.

- **Data-aware process networks (DPN).** This is the intermediate representation of the HLS flow. DPN is a parallel execution model fitting the hardware constraints of circuit synthesis, in which the data transfer and the synchronizations are made explicit. We formally described the DPN model and a translation scheme from C programs, and we showed the consistency in the meaning where any terminating sequential program is translated to an equivalent DPN, guaranteed to be deadlock free.
- **Front-end analysis.** We designed many program analyses to produce a quality DPN from a C program:
  - *Throughput optimization.* A I/O scheme has been designed, with the corresponding compiler analysis, to minimize the I/O traffic with the external memory. This allows us to balance efficiently the spilling of temporary value to the memory, and the local buffer size. This scheme impacts the DPN structure itself.
  - *Communication vectorization.* The matrix structure of the memory allows us to load data by chunks. A polyhedral analysis has been designed to solve this issue.
  - *Synchronization scheme.* As parallel units need to communicate intermediate results, synchronizations must be ensured.Unlike KPN, DPN do not use FIFO, but buffers, which required an efficient synchronization mechanism.
- **Back-end analysis.** Once generated, a DPN must be mapped to an FPGA. This raises many interesting issues:
  - *Pipeline completion.* Data paths make an extensive use of pipelined operators, which delays the signal. An algorithm has been designed to enforce the time coherence of signals.
  - *Polyhedral units.* DPNs make an extensive use of piece-wise affine functions, which must be mapped properly to ensure the efficiency of the whole system. A preliminary algorithm has been designed to reach a correct trade-off between critical path size and LUT usage.

All these analyses have been fully implemented. The tool Dcc (DPN C Compiler) implements all the front-end analyses. The tool IceGEN implements the back-end analysis.

6.10. Program Equivalence Modulo A/C (Associativity/Commutativity)

Participants: Guillaume Iooss [PhD student], Christophe Alias, Sanjay Rajopadhye [Colorado State University].

Program equivalence is a well-known problem with a wide range of applications, such as algorithm recognition, program verification, and program optimization. This problem is also known to be undecidable if the class of programs is rich enough, in which case semi-algorithms are commonly used. We focus on programs represented as a system of affine recurrence equations (SARE), defined over parametric polyhedral domains, a well-known formalism for the *polyhedral model*, which includes as a proper subset, the class of affine control loop programs. Several semi-algorithms for program equivalence have already been proposed for this class. A few of them take into account algebraic properties such as associativity and commutativity. However, to the best of our knowledge, none of them is able to manage reductions, i.e., accumulations of a parametric number of sub-expressions using an associative and commutative operator.

Our contributions are:

- An equivalence checking algorithm able to manage associativity and commutativity properties. Our method subsumes the previous approaches and is, to the best of our knowledge, the first one able to manage these properties over a parametric number of expressions.
- A semi-algorithm to construct a perfect matching problem on a parametric bipartite graph. We partially solve this problem through a heuristic based on the augmenting path algorithm. This heuristic is able to find a set of non-interfering augmenting paths to improve a proposed maximum matching, as long as these augmenting paths do not have a parametric length.

A preliminary implementation is under development. This work has been submitted to ESOP'14.

6.11. Constant Aspect-Ratio Parametric Tiling

Participants: Guillaume Iooss [PhD student], Sanjay Rajopadhye [Colorado State University], Christophe Alias, Yun Zou [PhD student, Colorado State University].

Parametric tiling is a well-known transformation that is widely used to improve locality, parallelism, and granularity. However, parametric tiling is also a non-linear transformation and this prevents polyhedral analysis or further polyhedral transformation after parametric tiling. It is therefore generally applied during the code generation phase.

This result consists on a method to stay polyhedral in a special case of parametric tiling, where all the dimensions are tiled and all the tile sizes are constant multiples of a single tile size parameter. We call this *Constant Aspect Ratio Tiling*. We show how to mathematically transform a polyhedron and an affine function into their tiled counterpart and show how to obtain good generated code.

This work has been accepted for publication at IMPACT'14 [8].

6.12. Parametric Tiling with Inter-Tile Data Reuse

Participants: Alain Darte, Alexandre Isoard.

Loop tiling is a loop transformation widely used to improve spatial and temporal data locality, increase computation granularity, and enable blocking algorithms, which are particularly useful when offloading kernels on platforms with small memories. When hardware caches are not available, data transfers must be software-managed: they can be reduced by exploiting data reuse between tiles and, this way, avoid some useless external communications. An important parameter of loop tiling is the sizes of the tiles, which impact the size of the necessary local memory. However, for most analyzes that involve several tiles, which is the case for inter-tile data reuse, the tile sizes induce non-linear constraints, unless they are numerical constants. This complicates or prevents a parametric analysis. In this work, we showed that, actually, parametric tiling with inter-tile data reuse is nevertheless possible.

Our solution is the first parametric solution for generating the memory transfers needed when a kernel is offloaded to a distant accelerator, tile by tile after loop tiling, and when all intermediate results are stored locally on the accelerator. For such computations, there is a complete decoupling between loads and stores, and when a value has been defined in a previous tile, it has to be loaded from the local memory and not from the distant memory as this memory is not yet up-to-date. In other words, inter-tile reuse is mandatory. This also saves external communications. Our solution is parametric in the sense that we derive the set of loads and stores from and to the distant memory with the tile sizes as parameters. Although the direct formulation is quadratic, we can still solve it in an affine way by developing techniques that consider, in the analysis, all (unaligned) possible tiles obtained by translation and not just those that belong to a tiling (partitioning) of the iteration space. We were able to use a similar technique to also parameterize the computations of local memory sizes, thanks to parametric lifetime analysis and folding with modulos, even for pipeline schedules similar to double buffering. Our method is currently implemented with the `iscc` calculator of `ISL`, a library for the manipulation of integer sets defined with Presburger arithmetic.

Also, the whole analysis can handle approximations thanks to the introduction of the concept of pointwise functions, well suited to deal with unaligned tiles. We believe that this technique can be used for other applications linked to the extension of the polyhedral model as it turns out to be fairly powerful. Our future work will be to derive efficient approximation techniques, either because the program cannot be fully analyzable, or because approximations can speed-up or simplify the results of the analysis without losing much in terms of memory transfers and/or memory sizes.

This work has been accepted for publication at IMPACT'14 [5].

6.13. Data Races in the Parallel Language X10

Participants: Tomofumi Yuki [Colorado State University and Inria/IRISA], Paul Feautrier, Sanjay Rajopadhye [Colorado State University], Vijay Saraswat [IBM Research].

Parallel programmers are now required to efficiently utilize the massive amount of parallelism provided by multi-core and many-core systems. Parallel programming is difficult, and the existing tools are mostly low-level extensions to sequential languages or libraries. As an effort to improve this situation, several groups have initiated the design of parallel programming languages, mostly based on the partitioned global address space (PGAS) paradigm. One of these languages is X10, which is developed at IBM Research by a team led by Vijay Saraswat.

While such languages hide the low-level details of parallel programming, they cannot guarantee that the object code will be correct by construction. Parallelism introduces two new types of bugs: non-determinism and deadlocks, and experience shows that it is possible to guarantee the absence of one type but not both. X10 programs are guaranteed deadlock-free but may have non-determinism. Non-determinism can be detected at runtime, but this approach cannot give absolute guarantees. However, it is possible, at least for a restricted class of X10 programs, to check for non-determinism at compile time.

The first step in this direction is to define the *polyhedral fragment* of X10, in which the only control constructs are `for` loops with affine bounds, and the only data structures are arrays with affine subscripts. X10 has many parallel constructs: as a first effort, we focused on `async`, which creates an activity (lightweight thread) and `finish`, which waits for termination of all impending activities. The execution order (or *happens-before relation*) of such a program is an incomplete lexicographic order, in which terms relating operations in different activities are removed. The dataflow analysis method of [23] has to be adapted to a partial execution order, which may have many extrema instead of a unique maximum. Multiple extrema denote data races, thus non-determinism. A detector along these lines has been implemented and presented at PPoPP'13 (Symposium on Principles and Practice of Parallel Programming) [10].

X10 other parallel programming primitive directives are *clocks* and `atomic`. The `at` construct allows down-loading a computation to another *place*. Clocks are a dynamic version of barriers. Their analysis involves counting their instances. For polyhedral programs, this can be done using the Ehrhart and Barvinok theories; the results are polynomials. Checking whether clocks remove non-determinism involves finding integer roots and hence is undecidable. However, modern SMT solvers are able to solve most of these problems. The resulting paper [13] has been submitted to the ECOOP conference.

6.14. Clock Removal in X10

Participants: Paul Feautrier, Eric Violard [Inria/Camus], Alain Ketterlin [Inria/Camus].

In the light of the previous work on the determinism of X10, a natural question is: are the parallel programming directives of X10 redundant? The answer is yes, at least for static control programs, i.e., programs in which the set of operations and their execution order do not depend on the input data. The basic idea is that the synchronization which occurs when several activities execute an advance is similar to the synchronization at the end of a finish. If one is able to count advances, one may construct a front by gathering all operations with the same advance count. Each front is executed inside one finish, and fronts are executed sequentially in order of increasing counts. For polyhedral programs, advance counting can be done at compile time. If the counts are affine functions, the restructuring can be done by classical polyhedral code generators like CLooG, and no overhead is incurred. For polynomial counts, one overall enclosing loop must be added, but the resulting program can usually be optimized by simple loop transformations, e.g., pushing guards into enclosing loop bounds. For arbitrary programs, the counts have to be computed dynamically; this is possible only if the program has static control.

This result does not contradict the previous undecidability proof (Section 6.13), as the translation of a polyhedral program is usually not polyhedral. Application of the method to a set of simple kernels has shown significant speedups. The interpretation of this result is that, at least in the present state of the X10 runtime, the implementation of the `async` primitive is more mature than the implementation of clocks. A paper on this topic has been accepted at CC'14 (Compiler Construction Conference) [7].

6.15. Static Analysis of OpenStream Programs

Participants: Albert Cohen [Inria, Parkas], Alain Darte, Paul Feautrier.

The objective of the collaboration between the Compsys and Parkas teams in the ManycoreLabs project (Section 7.2) is to evaluate the possibility of applying polyhedral techniques to the parallel language OpenStream, which is developed by Inria Parkas. When applicable, these techniques are invaluable for compile-time debugging and for improving the target code for a better adaptation to the target architecture.

OpenStream is a two-level language, in which a sequential control code directs the initialization of parallel task instances that communicate through *streams*. OpenStream programs are deterministic by construction, but may have deadlocks. If the control code is polyhedral, one may statically compute, for each task instance, its read and write indices for each stream. These indices may be polynomials of arbitrary degree. When linear, the full power of the polyhedral model may be brought to bear for dependence and dataflow analysis, scheduling and deadlock detection, and program transformations.

In the general case, one can think of two approaches: the first one consists in over-approximating dependences until problems become linear. In the second approach, one first leverages modern developments in SMT solvers, which allow them to solve polynomial problems, albeit with no guarantee of success. Furthermore, the task index functions have special properties that may be used to construct original analysis algorithms. Three preliminary results in this direction:

- the proof that deadlock detection is undecidable in general, thanks to an adaptation of the proof designed for X10 (Section 6.13),
- a characterization of deadlocks in terms of dependence graphs, which implies that streams can be safely bounded as soon as a schedule exists with such sizes,
- a preliminary analysis of some solvable cases.

A document is available as Deliverable 2.5.3 for the ManycoreLabs project.

6.16. Array Contraction in Parallel Programs

Participants: Alain Darte, Alexandre Isoard.

Array contraction is a technique to reuse array elements when they are dead, in a form of array folding. A standard technique for array contraction is to use affine remappings with modulos. When the modulo is equal to 1, this corresponds to the removal of the corresponding array dimension. Array contraction is well-known for sequential programs, after element-wise array liveness analysis. It has also been customized for parallel codes obtained through affine schedules by Lefebvre and Feautrier, and Quilleré-Rajopadhye, both frameworks being generalized by the lattice-based memory allocation framework of Darte, Schreiber, and Villard [17] and the construction of the set of conflicting array indices. We showed how the same framework can be used for a larger range of parallel programs, including programs with outer parallel loops, programs exhibiting pipelining, a subset of X10, etc. The optimality of the construction can be shown, despite a related (but actually non-contradictory here) NP-completeness result for worst-case of register pressure in the context of register allocation. A research report on this topic is in preparation.

CONTRAINTES Project-Team

6. New Results

6.1. A Stronger Necessary Condition for the Multistationarity of Chemical Reaction Networks

Participant: Sylvain Soliman.

In the last thirty years, the conjecture of Thomas on the necessary presence of a positive circuit for the occurrence of multistationarity has opened a whole field of research, allowing better modeling and understanding of biochemical networks, especially in the emerging field of systems biology. However, if that aspect is striking in the field of discrete modeling of gene regulatory networks, it did not have the same impact in the Ordinary Differential Equations (ODE) based modeling community. This is mostly due to the fact that this necessary condition, the existence of a positive loop in the Jacobian of the ODE system, is almost always satisfied.

In [5] we improve on the ten years old proof by Soulé, using the structural information from the stoichiometric matrix of a biochemical reaction system. This allows us to state a more strict version of the famous Thomas' necessary condition for multistationarity. In particular, the obvious cases where Thomas' condition was trivially satisfied, mutual inhibition due to a multimolecular reaction and mutual activation due to a reversible reaction, can now easily be ruled out. The new condition makes it possible to use circuit analysis as an useful tool in the arsenal of the computational biologist, together with other structural methods.

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

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

The Thesis of Faten Nabli [1] marks our achievements on the static analysis of biochemical reaction networks using Petri Net concepts and Constraint Logic Programming algorithms. This Thesis presents 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. The performances of these methods are compared 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. Faten Nabli has been hired with a Post-Doc position at Sanofi Paris.

6.3. Structural Model Reduction: CLP and SAT Solvers for Computing Subgraph Epimorphisms

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

This year, in [8], we have developed and compared CLP and SAT solvers on the NP-complete problem of deciding the existence of a subgraph epimorphism between two graphs. Our interest in this variant of graph matching problem stems from the study of model reductions in systems biology, where large systems of biochemical reactions can be naturally represented by bipartite digraphs of species and reactions. In this setting, model reduction can be formalized as the existence of a sequence of vertices, species or reaction, deletion and merge operations which transforms a first reaction graph into a second graph [2]. This problem is in turn equivalent to the existence of a subgraph (corresponding to delete operations) epimorphism (i.e. surjective homomorphism, corresponding to merge operations) from the first graph to the second. We show how subgraph epimorphism problems can be modeled as Boolean constraint satisfaction problems, and we compare CLP and SAT solvers on a large benchmark of reaction graphs from systems biology.

6.4. Quantitative Model Reduction: a CLP Solver for Computing Tropical Equilibrations

Participants: François Fages, Sylvain Soliman.

Model reduction is a central topic in computational systems biology and dynamical systems theory, for reducing the complexity of quantitative models, finding important parameters, and developing multi-scale models for instance. While perturbation theory is a standard mathematical tool to analyze the different time scales of a dynamical system, and decompose the system accordingly, tropical methods provide a simple algebraic framework to perform these analyses systematically in polynomial systems. The crux of these tropicalization methods is in the computation of tropical equilibrations. In [13], we show that constraint-based methods, using reified constraints for expressing the equilibration conditions, make it possible to numerically solve non-linear tropical equilibration problems, out of reach of standard computation methods. We illustrate this approach first with the reduction of simple biochemical mechanisms such as the Michaelis-Menten and Goldbeter-Koshland models, and second, with performance figures obtained on a large scale on the model repository `biomodels.net`.

This work is done in collaboration with Ovidiu Radulescu, Univ. Montpellier, in the context of a larger project about symbolic methods in systems biology with François Boullier, LIFL and Andras Weber, Univ. Bonn.

6.5. Species Minimization in Biochemical Reaction Computing

Participants: Hui-Ju Chiang, François Fages.

Engineering biochemical reactions for computational purposes is a common pursue in synthetic biology. In such design tasks, molecular species have to be carefully engineered to ensure modularity and orthogonality, and are scarce resources. Minimizing the number of involved molecular species is crucial to accomplish a complex computation within a confined biochemical environment.

In [10], we investigate an approach to species minimization by reusing modular and regular reactions in an asynchronous time-multiplexed fashion. Our method enhances not only species utility, but also re-programmability and robustness in realizing various logic circuits. A case study demonstrates the ease of design in realizing general logic computation, and simulation confirms the feasibility and robustness of the proposed method.

This work is done in collaboration with Jie-Hong Jiang and Katherine Chiang from NTU Taiwan in the context of a common project about biochemical programming.

6.6. Hybrid Composition and Simulation of Heterogeneous Biochemical Models

Participants: Hui-Ju Chiang, François Fages, Sylvain Soliman.

[2] Steven Gay, Sylvain Soliman, François Fages. A Graphical Method for Reducing and Relating Models in Systems Biology. Bioinformatics, 26(18):i575–i581, 2010.

Models of biochemical systems presented as a set of formal reaction rules with kinetic expressions can be interpreted with different semantics: as either deterministic Ordinary Differential Equations, stochastic continuous-time Markov Chains, Petri nets or Boolean transition systems. While the formal composition of reaction models can be syntactically defined as the (multiset) union of the reactions, the hybrid composition of models in different formalisms is a largely open issue.

In [7], we show that the combination of reaction rules with conditional events, as the ones already present in SBML, does provide the expressive power of hybrid automata and can be used in a non standard way to give meaning to the hybrid composition of heterogeneous models of biochemical processes. In particular, we show how hybrid differential-stochastic and hybrid differential-Boolean models can be compiled and simulated in this framework, through the specification of a high-level interface for composing heterogeneous models. This is illustrated by a hybrid stochastic-differential model of bacteriophage T7 infection, and by a reconstruction of the hybrid model of the mammalian cell cycle regulation of Singhania et al. as the composition of a Boolean model of cell cycle phase transitions and a differential model of cyclin activation.

6.7. Composition and Abstraction of Logical Influence Networks: Application to Multi-Cellular Systems

Participant: Grégory Batt.

Logical (Boolean or multi-valued) modelling is widely employed to study regulatory or signalling networks. Even though these discrete models constitute a coarse, yet useful, abstraction of reality, the analysis of large networks faces a classical combinatorial problem. In [4], we proposed to take advantage of the intrinsic modularity of inter-cellular networks to set up a compositional procedure that enables a significant reduction of the dynamics, yet preserving the reachability of stable states. To that end, we relied on process algebras, a well-established computational technique for the specification and verification of interacting systems.

We developed a novel compositional approach to support the logical modelling of interconnected cellular networks. First, we formalised the concept of logical regulatory modules and their composition. Then, we made this framework operational by transposing the composition of logical modules into a process algebra framework. Importantly, the combination of incremental composition, abstraction and minimisation using an appropriate equivalence relation (here the safety equivalence) yields huge reductions of the dynamics. We illustrated the potential of this approach with two case-studies: the Segment-Polarity and the Delta-Notch modules.

6.8. Identification of Biological Models from Single Cell Data: a Comparison between Mixed-Effects and Moment-based Inference

Participants: Grégory Batt, Andres Mauricio Gonzalez Vargas, Pascal Hersen, Artémis Llamosi, Jannis Uhlendorf.

Experimental techniques in biology such as microfluidic devices and time-lapse microscopy allow tracking of the gene expression in single cells over time. So far, few attempts have been made to fully exploit these data for modeling the dynamics of biological networks in cell populations. In [9], we compare two modeling approaches capable to describe cell-to-cell variability: Mixed-Effects (ME) models and the Chemical Master Equation (CME). We discuss how network parameters can be identified from experimental data and use real data of the HOG pathway in yeast to assess model quality.

For CME we rely on the identification approach proposed by Zechner et al. (PNAS, 2012), based on moments of the probability distribution involved in the CME. ME and moment-based (MB) inference will be also contrasted in terms of general features and possible uses in biology.

6.9. STL-based Analysis of TRAIL-induced Apoptosis Challenges the Notion of Type I/Type II Cell Line Classification

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. These differences are attributed to the activation of one of two pathways, leading to classification of cell lines into two groups: type I and type II.

In [6] we challenge this type I/type 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. After having solved a few inconsistencies using STL-guided parameter search, 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 these three distinguishing behaviors should be merely considered as type I/II features rather than cell-type defining criteria. On the methodological side, this work illustrates the biological relevance of STL-diagrams, STL population data, and STL-guided parameter search implemented in the tool Breach. Such tools are well-adapted to the ever-increasing availability of heterogeneous knowledge on complex signal transduction pathways.

6.10. Single Cell Models and Models of Populations: A Mixed Effect Approach

Participants: Grégory Batt, Andres Mauricio Gonzalez Vargas, Pascal Hersen, Artémis Llamosi.

For a long time, experiments and models of gene expression were mainly based on the mean behavior of a population of cells. Although observed early, it is only recently that experimental technique allowed detailed investigation of variability in this process. Since the pioneering work of Elowitz and colleagues, a distinction is drawn between what is called intrinsic and extrinsic variability or noise. Intrinsic noise originates in the randomness of chemical reactions within a cell whether extrinsic noise is the variation in between cells at a given time. Extrinsic variability is associated with population heterogeneity in the concentrations of ribosomes or other molecular players or processes relevant to gene expression (RNAPolII concentration, degradation and dilution rates etc.).

In this work, we propose a modelling framework for gene expression based on a system of ODEs with random parameters following a distribution across the population of cells. In this context, each cell has its own identity which is represented by the value of its parameters. With this model we ask how much of the long term variability can be explained by extrinsic variability alone. We produced long term, time lapse and single-cell data of repeated gene induction in Saccharomyces cerevisiae. One experiment was treated as learning set whereas two were used as test sets. From the learning set, we are able to infer single cell parameters and population distributions which represent accurately in terms of mean and variance the variability in the population. These learned population distributions allowed good predictions on both the learning and test sets.

Our study demonstrates also that the way inference of single cell parameters and distributions is performed is crucial to achieve good performance. Best results being found by joint estimation of the parameters for single cells and for the whole population. With this technique, we noted that very decent fits of the population dynamics can be obtained by estimating only on a very limited number of cells. Concerning the quality of single cell parameters inferred, we validated the presence of an expected significant correlation between the dilution rate and the measured single cell growth rate. This motivates the use of this tool in order to investigate the origins of extrinsic noise, by correlating single cell parameters with measured candidate factors of gene expression variability such as cell density, cell size or age.

6.11. 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 important 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 pursued the investigation of the effect of transcription inhibition during mitosis, as a reverse coupling from the cell cycle to the circadian clock. We use quantitative 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. We are defining a series of common temporal logic patterns and *ad hoc* schemes for computing their validity domain on a given trace, more efficiently than by the generic method implemented in BIOCHAM.

6.12. Solving Mixed Shapes Packing Problems by Continuous Optimization with the CMA Evolution Strategy

Participants: François Fages, Thierry Martinez, Lumadaiara Do Nascimento Vitorino.

Bin packing is a classical combinatorial optimization problem which has a wide range of real-world applications in industry, logistics, transport, parallel computing, circuit design and other domains. While usually presented as discrete problems, in [12] we consider continuous packing problems including curve shapes, and model these problems as continuous optimization problems with a multi-objective function combining non-overlapping with minimum bin size constraints. More specifically, we consider the covariance matrix adaptation evolution strategy (CMA-ES) with a nonoverlapping and minimum size objective function in either two or three dimensions. Instead of taking the intersection area as measure of overlap, we propose other measures, monotonic with respect to the intersection area, to better guide the search. In order to compare this approach to previous work on bin packing, we first evaluate CMA-ES on Korf's benchmark of consecutive sizes square packing problems, for which optimal solutions are known, and on a benchmark of circle packing problems. We show that on square packing, CMA-ES computes solutions at typically 14% of the optimal cost, with the time limit given by the best dedicated algorithm for computing optimal solutions, and that on circle packing, the computed solutions are at 2% of the best known solutions. We then consider generalizations of this benchmark to mixed squares and circles, boxes, spheres and cylinders packing problems, and study a real-world problem for loading boxes and cylinders in containers. These hard problems illustrate the interesting trade-off between generality and efficiency in this approach.

6.13. Railway Time Tabling Optimization with CMA-ES and Greedy Heuristics

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

The problem of reducing energy consumption in public transportation has received increasing attention over the last years. Most metros have energy regenerative braking systems, which allow them to produce electric energy when they brake. We study the problem of optimizing the energy consumption of a metro line by modifying the timetable, in order to maximize the actual reuse of the regenerative energy. This is achieved by synchronizing the braking and acceleration phases of the metros, through slight modifications of the stopping times in stations. In an article in preparation, we present a constraint-based model of the electric network of the line, which is used to evaluate the energy consumption at each instant, and to compute a distribution matrix for approximating the potential energy transfers between metros. The optimization of the timetable is then performed by an evolutionary algorithm using the Covariance Matrix Adaptation Evolution Strategy (CMA-ES from Nikolaus Hansen, EPI TAO). On real data, this strategy shows energy savings ranging from 2.38% to 4.54%. Furthermore, these savings are shown to be robust with respect to perturbations of the dwell times.

CONVECS Project-Team

6. New Results

6.1. New Formal Languages and their Implementations

LNT is a next generation formal description language for asynchronous concurrent systems, which attempts to combine the best features of imperative programming languages and value-passing process algebras. LNT is increasingly used by CONVECS for industrial case studies and applications (see § 6.5) and serves also in university courses on concurrency, in particular at ENSIMAG (Grenoble) and at Saarland University.

6.1.1. Translation from LNT to LOTOS

Participants: Hubert Garavel, Frédéric Lang, Wendelin Serwe.

The LNT2LOTOS, LNT.OPEN, and LPP tools convert LNT code to LOTOS, thus allowing the use of CADP to verify LNT descriptions. These tools have been used successfully for many different systems (see § 6.5 and § 9.1).

In 2013, in addition to 15 bug fixes, the following enhancements have been brought to these tools:

- The list of predefined functions that can be generated automatically for list and set types has been enriched, so as to include all operations commonly found in programming languages.
- A new "sorted set" type was added to LNT, where the automatically generated insertion function preserves the invariant that all elements in the set are sorted in ascending order and have a single occurrence.
- The use of range and predicate types has been facilitated, by translating explicit type annotations by type conversions if necessary.
- An implicit type conversion is applied by assignments to a variable; this helps the type-checker to solve overloaded definitions.
- The generated LOTOS and C code has been modified to avoid spurious warning messages from the LOTOS and C compilers.
- The demo examples demo_19, demo_30, and demo_35 have been enhanced with an LNT version.

6.1.2. Translation from LOTOS to Petri nets and C

Participants: Hubert Garavel, Wendelin Serwe.

The LOTOS compilers CAESAR and CAESAR.ADT, which were once the flagship of CADP, now play a more discrete role since LNT (rather than LOTOS) has become the recommended specification language of CADP. Thus, CAESAR and CAESAR.ADT are mostly used as back-end translators for LOTOS programs automatically generated from LNT or other formalisms such as Fiacre, and are only modified when this appears to be strictly necessary.

In 2013, in addition to fixing four bugs, the type checking algorithm of CAESAR and CAESAR.ADT was entirely revised to display less and better messages in case of typing errors, avoiding cascading error messages, e.g., when an undefined variable or constant is used, or when an overloaded function is improperly used in a context where a unique type is expected.

Also, the CAESAR compiler found a new usefulness as a means to easily produce large-size, realistic Petri nets that can be used as benchmarks by the Petri net community. To make this possible, a new option was added to CAESAR to export the BPN (*Basic Petri Net*) file generated from a LOTOS specification. The definition of the BPN format was made more precise by adding semantic constraints. The CAESAR.BDD tool of CADP was enhanced with two new options, one that checks whether a BPN file satisfies all semantic constraints, and another one that converts a BPN file into PNML (*Petri Net Markup Language*) format.

This work has been done in coordination with Fabrice Kordon and Lom-Messan Hillah (UPMC/LIP6, Paris, France) for the MCC (*Model Checking Contest*) workshop [5]. H. Garavel was in charge of redesigning the model forms used for this contest. One Petri net generated using CAESAR was selected as a benchmark for MCC'2013 and five Petri nets generated using CAESAR have been submitted to MCC'2014.

6.1.3. Translation from an Applied Pi-Calculus to LNT

Participants: Radu Mateescu, Gwen Salaün.

The π-calculus is a process algebra defined by Milner, Parrow, and Walker two decades ago for describing concurrent mobile processes. Despite a substantial body of theoretical work in this area, only a few verification tools have been designed for analysing π-calculus specifications automatically. Our objective is to provide analysis features for the π-calculus by reusing the verification technology available for process algebras without mobility. For this purpose, we extended the original polyadic π-calculus with the data types and functions of LNT. This yields a general-purpose applied π-calculus, which is suitable for specifying mobile value-passing concurrent systems belonging to various application domains. Our approach is based on a novel translation from the finite control fragment of π-calculus to LNT, making possible the analysis of applied π-calculus specifications using all verification tools of CADP. This translation is fully automated by the PIC2LNT translator (see § 5.3).

In 2013, we continued our work on the applied π-calculus and its translation to LNT. This resulted in a new version PIC2LNT 3.0 of the tool, which fixes several bugs and brings the following improvements:

- A bounded replication operator was added to the language, which expresses the parallel execution of a fixed number of π-calculus agents. This operator is translated into LNT by instantiating the appropriate number of corresponding processes.
- A type Chan representing channel names was implemented, which can be freely combined with ordinary data types. This increases the versatility of the language by allowing, e.g., the definition of agents parametrized by sets of channel names.
- Several options were added to the tool for enhancing its ergonomy and tuning the state space generation (specify the set of private channels that can be created, generate the state space of a particular agent).

A paper describing this work has been published in an international conference [16].

6.1.4. Translation from EB3 to LNT

Participants: Frédéric Lang, Radu Mateescu.

In collaboration with Dimitris Vekris (University Paris-Est Créteil), we considered a translation from the EB3 language [39] for information systems to LNT. EB3 has a process algebraic flavor, but has the particularity to contain so-called *attribute functions*, whose semantics depend on the history of events. We have proposed a formal translation scheme, which ensures the strong equivalence between the LTSs corresponding to an EB3 specification and to the LNT code generated. A prototype translator has been developed at University Paris-Est Créteil, which enables EB3 specifications to be formally verified using CADP.

In 2013, a paper has been published in an international conference [19].

6.1.5. Coverage Analysis for LNT

Participants: Gwen Salaün, Lina Ye.

In the classic verification setting, the designer has a specification of a system in a value-passing process algebra, a set of temporal properties to be verified on the corresponding LTS model, and a data set of examples (test cases) for validation purposes. At this stage, building the set of validation examples and debugging the specification is a complicated task, in particular for non-experts.

[5] http://mcc.lip6.fr

In 2013, we proposed a new framework for debugging value-passing process algebraic specifications by means of coverage analysis and we illustrated our approach with LNT. We define several coverage notions before showing how to instrument the specification without affecting its original behavior. Our approach helps the specifier to find dead code, ill-formed conditional structures, and other errors in the specification, but also to improve the quality of a data set of examples used for validation purposes. We have implemented a prototype tool, named CAL, for automating the verification of coverage analysis, and we applied it to several real-world case studies in different application areas. A paper has been submitted to an international conference.

6.1.6. Other Compiler Developments

Participants: Soraya Arias, Hubert Garavel, Frédéric Lang, Wendelin Serwe.

- In co-operation with Jérôme Hugues (ISAE, Toulouse), we investigated the translation of AADL (*Architecture Analysis and Design Language*) into LNT. An AADL example was manually tackled, leading to the conclusion that LNT could be a suitable target language for translating a large fragment of AADL.

  In co-operation with Holger Hermanns (Saarland University, Germany) and Joost-Pieter Katoen (RWTH Aachen, Germany), we prepared a contribution for the AADL standardization committee to detail semantics issues of the GSPN (*Generalized Stochastic Petri Nets*) model.
- We continued our work on the FLAC tool, which translates the Fiacre intermediate language into LOTOS to enable verification using CADP. In 2013, we eliminated spurious compilation warnings, we removed the definitions of integer operations *div* and *mod*, which have been added to a standard LOTOS library, and we improved the encoding of integer numbers. These changes have led to revisions 76 to 79 of the FLAC code, which is available on the development forge dedicated to Fiacre compilers [6].
- In co-operation with Holger Hermanns, we started studying the PseuCo language that is being defined and implemented at Saarland University. Developed from an educational perspective as a means to teach concurrency theory to bachelor students, PseuCo combines features from Java and Go, the language promoted by Google for concurrent programming. PseuCo supports both message-passing and shared-memory concurrency in a way that is easy to use and that can readily be transferred to Java, Go, or other mainstream languages. PseuCo has been awarded with the 2013 German national "*Preis des Fakultätentages Informatik*" for its innovative role in undergraduate education.

  In 2013, we undertook the manual translation of various PseuCo sample programs into LNT and started enhancing LNT with features that would enable automated PseuCo-to-LNT translation. We also reviewed a PseuCo-to-CCS translator recently developed at Saarland University and wrote an evaluation report for this software.

6.2. Parallel and Distributed Verification

6.2.1. Manipulation of Partitioned LTSs

Participants: Hubert Garavel, Radu Mateescu, Wendelin Serwe.

For distributed verification, CADP provides the PBG format, which implements the theoretical concept of *Partitioned LTS* [45] and provides a unified access to an LTS distributed over a set of remote machines.

In 2013, we continued the development of the prototype tool PBG_OPEN, which is an OPEN/CAESAR-compliant compiler for the PBG format, enabling the use of all CADP on-the-fly verification tools on a partitioned LTS. The main advantage of PBG_OPEN is that it can use the memory of several machines to store the transition relation of a partitioned LTS. Therefore, PBG_OPEN can explore on-the-fly large partitioned LTSs that could not be explored using other tool combinations. To reduce the amount of communications, PBG_OPEN can use a cache to store already encountered states, together with their outgoing transitions.

[6] http://gforge.enseeiht.fr/projects/fiacre-compil

We also developed another prototype tool, named PBG_INVERT, which changes the storage of the transitions of a partitioned LTS, transforming a partitioned LTS where each fragment stores the transitions leading to the states of the fragment (as generated by DISTRIBUTOR) into a partitioned LTS where each fragment stores the transitions going out from the states of the fragment. Adding this transformation step yields a reduction of up to 25% of the overall execution time, when verifying the partitioned LTS with PBG_OPEN. We experimented all these tools on the Grid'5000 computing infrastructure [31] using up to 512 distributed processes. These experiments confirmed the good scalability of our distributed LTS manipulation approach. A paper describing this work has been published in an international conference [13].

6.2.2. Distributed Code Generation for LNT

Participants: Hugues Evrard, Frédéric Lang.

Rigorous development and prototyping of a distributed verification algorithm in LNT involves the automatic generation of a distributed implementation. For the latter, a protocol realizing process synchronization is required. As far as possible, this protocol must itself be distributed, so as to avoid the bottleneck that would inevitably arise if a unique process would have to manage all synchronizations in the system. A particularity of such a protocol is its ability to support *branching synchronizations*, corresponding to situations where a process may offer a choice of synchronizing actions (which themselves may nondeterministically involve several sets of synchronizing processes) instead of a single one. Therefore, a classical barrier protocol is not sufficient and a more elaborate synchronization protocol is needed.

In 2013, we formally modelled and verified several existing synchronization protocols. This revealed an error in one of them, which led to a publication in an international conference [12]. Based on this study, we selected a suitable protocol and adapted it to the LNT synchronization operators.

Using this protocol, we developed a prototype distributed code generator, taking as input the model of a distributed system, described as a set of LNT processes and their parallel composition written in EXP. The LNT.OPEN and CAESAR tools are used to obtain the sequential implementation of each LNT process, and the EXP.OPEN tool is used to compute the possible interactions between processes. Then, our prototype generates the corresponding implementation of the distributed synchronization protocol and all necessary glue code between processes and the protocol. Our prototype automatically performs all these steps, such that a complete and runnable distributed implementation can easily be obtained from the original model.

So far, our prototype manages synchronizations with no data or data of enumerated types only, in which case the implementation checks that data values and types match before allowing a synchronization.

6.3. Timed, Probabilistic, and Stochastic Extensions

Participants: Hubert Garavel, Frédéric Lang, Radu Mateescu.

Process calculi provide a suitable formal framework for describing and analyzing concurrent systems, but need to be extended to model refined aspects of these systems. For instance, it may be necessary to represent probabilistic choices (in addition to deterministic and nondeterministic choices) as well as delays and latencies governed by probability laws. Many such extensions have been proposed in the literature, some of which have been implemented in software tools and applied to nontrivial problems. In particular, two of these extensions (namely, *Interactive Markov Chains* and *Interactive Probabilistic Chains*) are implemented in CADP. Despite these achievements, the state of the art is not satisfactory as the extended languages primarily focus on the probabilistic and stochastic aspects, leaving away the expressive and user-friendly features that process calculi provide for describing conventional concurrent systems.

In 2013, we did the following steps to progress our agenda of bridging the gap between functional verification and quantitative evaluation:

- We equipped CADP with a new tool named BCG_CMP, which enables to compare quantitative models modulo probabilistic and stochastic variants of strong bisimulation and branching bisimulation. Such comparison relations were not available in the BISIMULATOR tool that already existed in CADP.

- We investigated the feasibility of creating interconnections between mainstream verification tools for probabilistic and stochastic systems. In a first step, we focused on the DTMC (*Discrete-Time Markov Chain*) model and on three mainstream tools: CADP (Grenoble), MRMC (Aachen), and PRISM (Birmingham-Oxford).

  We developed translation tools to perform conversions between the various formats of these tools (".aut" and ".bcg" for CADP, ".tra/.sta/.lab" for MRMC, ".pm" and ".tra/.sta/.lab" for PRISM). So doing, we reported one bug in MRMC and five minor issues in PRISM. By discussing with Dave Parker (University of Birmingham), we contributed to the introduction in PRISM 4.1 of two new options "-importmodel" and "-exportmodel" that greatly simplify exchanges of models between PRISM and other tools.

  We developed a generator of random DTMCs in CADP, MRMC, and PRISM formats, and undertook the construction of a collection of DTMCs, which we used to compare the performance and scalability of CADP and PRISM.
- We started to investigate the evaluation of temporal logic properties on extended DTMCs, in which transitions are labeled with probabilities and optional actions. For this purpose, we developed a new prototype XTL library (consisting of XTL and C code) encoding the PCTL (*Probabilistic CTL*) temporal logic [50]. This new PCTL library enables the specifier to combine data-based, discrete-time, and probabilistic properties of DTMCs in a uniform way.

6.4. Component-Based Architectures for On-the-Fly Verification

6.4.1. Compositional Model Checking

Participants: Frédéric Lang, Radu Mateescu.

We have continued our work on partial model checking following the approach proposed in [26]. Given a temporal logic formula φ to be evaluated on a set S of concurrent processes, partial model checking consists in transforming φ into another equivalent formula φ' to be evaluated on a subset of S. Formula φ' is constructed incrementally by choosing one process P in S and incorporating into φ the behavioral information corresponding to P — an operation called quotienting. Simplifications must be applied at each step, so as to maintain formulas at a tractable size.

In 2013, we extended the approach to handle fairness operators of alternation depth two, and we conducted new experiments. This resulted in a new version of the PMC prototype tool (see § 5.4) supporting all features of the input language of EXP.OPEN 2.1. An article has been published in an international journal [5].

6.4.2. On-the-Fly Test Generation

Participants: Radu Mateescu, Wendelin Serwe.

In the context of the collaboration with STMicroelectronics (see § 6.5.1 and § 7.1), we studied techniques for testing if an implementation is conform to a formal model written in LNT. Our approach is inspired by the theory of conformance testing [68], as implemented for instance in TGV [53] and JTorX [30].

We developed two prototype tools. The first tool implements a dedicated OPEN/CAESAR-compliant compiler for the particular asymmetric synchronous product of the model and the test purpose, and uses slightly extended generic components for graph manipulation (τ-compression, τ-confluence reduction, determinization) and resolution of Boolean equation systems. The second tool generates the complete test graph, which can be used to extract concrete test cases or to drive the test of the implementation. The principal advantage of our approach compared to existing tools is the use of LNT for test purposes, facilitating the manipulation of data values.

In 2013, we continued the development of these tools, with a focus on reducing execution time. We also implemented a prototype tool to extract from a complete test graph one or all test cases of minimal depth. We experimented with these tools on two case-studies, namely the ACE coherence protocol (see § 6.5.1) and the EnergyBus (see § 6.5.5).

6.4.3. Equivalence Checking

Participant: Frédéric Lang.

Equivalence relations can be used for verification in two complementary ways: for the minimization of an LTS and the comparison of two LTSs.

In 2013, we worked along the following lines:

- We added observational equivalence (following a request from LAAS-CNRS) as well as divergence-sensitive branching bisimulation (together with its stochastic and probabilistic variants) in BCG_MIN.
- We improved the speed of BCG_MIN in the case of branching reduction applied to a graph with a high branching factor and many internal transitions, by correcting a function that has a quadratic complexity instead of a linear one.
- We added the new tool BCG_CMP, which takes as input two BCG graphs and checks whether they are equivalent modulo a relation chosen among strong and branching bisimulation (and their stochastic and probabilistic variants), divergence-sensitive branching bisimulation, or observational equivalence. BCG_CMP checks equivalence using the partition-refinement algorithm of BCG_MIN. We compared BCG_CMP and BISIMULATOR on the VLTS benchmark suite [7], showing that BCG_CMP is generally slightly less efficient than BISIMULATOR for comparisons yielding a FALSE result, but much more efficient than BISIMULATOR for comparisons yielding a TRUE result.
- The new tool BCG_CMP as well as the new equivalence relations added to BCG_MIN have been added to the EUCALYPTUS graphical user interface and to the SVL scripting language.

6.4.4. Other Software Developments

The OPEN/CAESAR environment was enhanced with a new generic library (named CAESAR_CACHE_1) for manipulating hierarchical caches, with 15 built-in replacement strategies and the possibility to define new ones.

We also maintained the CADP toolbox, taking into account the feedback received from numerous users in the world. In addition to fixing 41 bugs, we evolved CADP to support the latest versions of Windows, Cygwin, Mac OS X, and their corresponding C compilers. The documentation for installing CADP has been updated and shortened. Finally, support for Sparc, Itanium, and PowerPC processors was dropped at the end of 2013 based on the observation that these architectures are almost no longer used among the CADP user community.

6.5. Real-Life Applications and Case Studies

6.5.1. ACE Cache Coherency Protocol

Participants: Abderahman Kriouile, Radu Mateescu, Wendelin Serwe.

In the context of a CIFRE convention with STMicroelectronics, we studied system-level cache coherency, a major challenge faced in the current system-on-chip architectures. Because of their increasing complexity (mainly due to the significant number of computing units), the validation effort using current simulation-based techniques grows exponentially. As an alternative, we study formal verification.

We focused on the ACE (*AXI Coherency Extensions*) cache coherency protocol, a system-level coherency protocol proposed by ARM [25]. In a first step, we developed a formal LNT model (about 3200 lines of LNT) of a system consisting of an ACE-compliant cache coherent interconnect, processors, and a main memory. The model is parametric and can be instantiated with different configurations (number of processors, number of cache lines, number of memory lines) and different sets of supported elementary ACE operations (currently, a representative subset of 15 operations), including an abstract operation that represents any other ACE operation. We handled the global requirements of the ACE specification using a constraint oriented programming style, i.e., by representing each global requirement as a dedicated process observing the global behaviour and inhibiting incorrect executions.

[7] http://cadp.inria.fr/resources/vlts

In a second step, we generated for several configurations the corresponding LTS (up to 100 million states and 350 million transitions). We wrote two liveness properties in MCL expressing that each read (respectively write) transaction is executed until its termination. We also wrote two properties expressing cache coherence and data integrity. This required to transform state-based properties into action-based properties, by adding information about the cache state to actions executed by the cache. For all considered configurations, we checked these properties using parametric SVL scripts (about 100 lines) and EVALUATOR. For some scenarios without the processes representing the global requirements, EVALUATOR generated counterexamples for the cache coherence and data integrity. We are currently using these counterexamples to derive test cases for the architecture under design at STMicroelectronics.

This work led to publications [21], [15].

6.5.2. Choreography-based Communicating Systems

Participants: Radu Mateescu, Gwen Salaün, Lina Ye, Kaoutar Hafdi.

Choreographies are contracts specifying interactions among a set of services from a global point of view. These contracts serve as reference for the further development steps of the distributed system. Therefore, their specification and analysis is crucial to avoid issues (e.g., deadlocks) that may induce delays and additional costs if identified lately in the design and development process.

In 2013, we have obtained the following results:

- In collaboration with Meriem Ouederni (University of Toulouse) and Tevfik Bultan (University of California at Santa Barbara), we have proposed a branching definition of the synchronizability property, which identifies systems whose interaction behavior remains the same when asynchronous communication is replaced with synchronous communication. We have also shown how these results can be used for checking the compatibility of a set of asynchronously communicating components [17].
- In collaboration with Matthias Güdemann (Systerel), we have defined sufficient conditions for checking the repairability property, which indicates whether realizability can be enforced for choreography-based communicating systems using distributed controllers. A paper has been submitted to an international conference.
- We have proposed an approach for computing the degree of parallelism of BPMN processes using model checking techniques. A paper has been submitted to an international conference.
- In collaboration with Pascal Poizat (University of Paris Ouest Nanterre), we have been working on the development of the VerChor platform, which aims at assembling all the verification techniques and tools automating the analysis of choreography specifications [14].

6.5.3. Deployment and Reconfiguration Protocols for Cloud Applications

Participants: Rim Abid, Gwen Salaün.

We collaborated with Noël de Palma and Fabienne Boyer (University Joseph Fourier), Xavier Etchevers and Thierry Coupaye (Orange Labs, Meylan, France) in the field of cloud computing applications, which are complex distributed applications composed of interconnected software components running on distinct virtual machines. Setting up, (re)configuring, and monitoring these applications involves intricate management protocols, which fully automate these tasks while preserving application consistency as well as some key architectural invariants.

In 2013, we focused on the reliability of the self-configuration protocol [23]. This protocol always succeeds in deploying a cloud application, even when facing a finite number of virtual machine or network failures. Designing such highly parallel management protocols is difficult, therefore formal modelling techniques and verification tools were used for validation purposes. These results were accepted for publication in an international conference [11]. Also, an experience export on the verification tasks for such (re)configuration protocols has been published in an international journal [8].

We have also worked on the design and verification of a reconfiguration protocol, where virtual machines interact altogether using a publish-subscribe messaging system. The verification of this protocol with CADP helped to refine several parts of the protocol and correct subtle bugs. These results have been published in an international conference [10]. In collaboration with Francisco Durán (University of Málaga), we have also worked on the design of a variant of this reconfiguration protocol, where the virtual machines interact via FIFO buffers. A paper has been submitted to an international conference.

6.5.4. Networks of Programmable Logic Controllers

Participants: Hubert Garavel, Fatma Jebali, Jingyan Jourdan-Lu, Frédéric Lang, Eric Léo, Radu Mateescu.

In the context of the Bluesky project (see § 8.1.2.1), we study the software applications embedded on the PLCs (*Programmable Logic Controllers*) manufactured by Crouzet Automatismes. One of the objectives of Bluesky is to enable the rigorous design of complex control applications running on several PLCs connected by a network. Such applications are instances of GALS (*Globally Asynchronous, Locally Synchronous*) systems composed of several synchronous automata embedded on individual PLCs, which interact asynchronously by exchanging messages. A formal analysis of these systems can be naturally achieved by using the formal languages and verification techniques developed in the field of asynchronous concurrency.

For describing the applications embedded on individual PLCs, Crouzet provides a dataflow language with graphical syntax and synchronous semantics, equipped with an ergonomic user interface that facilitates the learning and use of the language by non-experts. To equip the PLC language of Crouzet with functionalities for automated verification, the solution adopted in Bluesky was to translate it into a pivot language that will enable the connection to testing and verification tools covering the synchronous and asynchronous aspects. Our work focuses on the translation from the pivot language to LNT, which will provide a direct connection to all verification functionalities of CADP, in particular model checking and equivalence checking.

In 2013, we studied the existing approaches and languages that address formal modeling and verification of GALS systems. We concluded that the current landscape lacks general-purpose, flexible, and formal representation of GALS systems suitable for efficient verification. To fulfill this requirement, we have designed GRL (*GALS Representation Language*), a language with user-friendly syntax and formal semantics, to efficiently model GALS systems for the purpose of formal verification. GRL targets GALS systems consisting of networks of synchronous systems interacting with their environments and communicating via asynchronous media. GRL draws mainly from two foundations. Regarding asynchronous concurrency, GRL builds upon process calculi (in particular LNT). Thereby, it leverages process calculi expressiveness, versatility, and verification efficiency. Regarding synchronous features, GRL holds a dataflow-oriented model based on the dataflow diagram model (also called block-diagram model). The GRL synchronous model inherits from the simplicity and modularity of the block-diagram model.

We defined the lexical and the abstract syntax of GRL (about 80 grammar rules), its static semantics (about 150 binding, typing, and initialization rules), and its dynamic semantics (about 20 structured operational semantics rules). Using the SYNTAX and LOTOS NT compiler construction technology, we started the development of a prototype translator GRL2LNT (about 8000 lines). The tool currently performs the lexical and syntactic analysis of GRL programs, together with some static semantic checks. A database containing about 30 examples of GRL programs has been constructed and used for non-regression testing of GRL2LNT. A reference manual for GRL (130 pages up to now) containing the definition of the language and its translation to LNT has been written. A paper presenting the GRL language has been submitted to an international conference.

Regarding the analysis of PLC networks by equivalence checking, we defined variants of classic equivalence relations (strong, τ*.a, and branching) for comparing the Mealy machine corresponding to a PLC network with the Moore machine corresponding to its external behaviour. We reformulated the verification problem as the resolution of a Boolean equation system, and we developed a prototype tool, based on the CAESAR_SOLVE_1 library, for the on-the-fly comparison of a Mealy and a Moore machine modulo the strong or the τ*.a equivalences.

6.5.5. EnergyBus Standard for Connecting Electric Components

Participants: Hubert Garavel, Wendelin Serwe.

The EnergyBus [8] is an upcoming industrial standard for electric power transmission and management, based on the CANopen field bus. It is developed by a consortium assembling all major industrial players (such as Bosch, Panasonic, and Emtas) in the area of light electric vehicles (LEV); their intention is to ensure interoperability between all electric LEV components. At the core of this initiative is a universal plug integrating a CAN-Bus[9] with switchable power lines. The central and innovative role of the EnergyBus is to manage the safe electricity access and distribution inside an EnergyBus network.

In the framework of the European FP7 project SENSATION (see § 8.2.1.1) a formal specification in LNT of the main EnergyBus protocols is being developed by Alexander Graf-Brill and Holger Hermanns at Saarland University [49], with the active collaboration of CONVECS.

In 2013, CONVECS provided help in modelling using the LNT language and the TGV tool, and enhanced the CADP toolbox to address a number of issues reported by Saarland University. At present, this LNT specification (1670 lines) is used for generating test suites using the TGV tool [53]. The formal modelling prompted for modifications in the EnergyBus standard and the generated test suites revealed three unknown bugs in an industrial CANopen implementation.

6.5.6. Graphical User-Interfaces and Plasticity

Participants: Hubert Garavel, Frédéric Lang, Raquel Oliveira.

In the context of the Connexion project (see § 8.1.1.2) and in close co-operation with Gaëlle Calvary, Eric Ceret, and Sophie Dupuy-Chessa (IIHM team of the LIG laboratory), we study the formal description and validation of graphical user-interfaces using the most recent features of the CADP toolbox. The case study assigned to LIG in this project is a prototype graphical user-interface [35] designed to provide human operators with an overview of a running nuclear plant. Contrary to conventional control rooms, which employ large desks and dedicated hardware panels for supervision, this new-generation interface uses standard computer hardware (i.e., smaller screen(s), keyboard, and mouse), thus raising challenging questions on how to best provide synthetic views of status information and alarms resulting from faults, disturbances, or unexpected events in the plant. Another challenge is to introduce plasticity in such interface, so as to enable several supervision operators, including mobile ones outside of the control room, to get accurate information in real time.

In 2013, CONVECS contributed to the following results. Based upon the available information published by EDF, a formal specification in LNT of this new-generation interface was developed (2600 lines). This specification not only encompasses the usual components traditionally found in graphical user-interfaces, but also a model of the physical world (namely, a nuclear reactor with various fault scenarios) and a cognitive model of a human operator in charge of supervising the plant. Also, a few desirable properties of the interface have been expressed in the MCL language of CADP and verified on the LNT model.

So doing, three main difficulties have been faced. The description of the prototype available in the published literature is not exhaustive, which required us to provide those missing details needed to obtain a realistic model. Quite often, we faced a combinatorial explosion in the number of states of the model, which forced us to restrict the complexity of operator behaviour and fault models. Finally, this case study revealed several LNT-specific issues, which triggered enhancements in the LNT language and tools.

[8] http://www.energybus.org
[9] http://www.can-cia.org

COPRIN Project-Team

6. New Results

6.1. Robotics

6.1.1. Cable-driven parallel robots (CDPR)

6.1.1.1. Analysis of Cable-driven parallel robots

Participants: Laurent Blanchet, Jean-Pierre Merlet [correspondant], Yves Papegay, Rémy Ramadour.

We are still investigating the extremely complex analysis of the kinematics [24] of CDPRs assuming either rigid [21] [20], elastic or sagging cables.

We have also started an analysis of *cable configuration* of redundantly actuated CDPRs for control purposes. Indeed we have shown that for robot with rigid cables it is impossible to have, in a given pose, more than 6 cables in tension simultaneously: the set of cables under tension is called the cable configuration. However at a pose there may be different sextuplets of under tension cables that satisfy the kinematico-static equations. Each of these sextuplets exhibits different performances (e.g. maximal tension in the cables or sensitivity of the positioning to errors in the cable lengths). Hence it may be interesting for control purposes to select one of the sextuplet that is optimal with respect to a performance criteria and to enforce this configuration by letting *voluntary* the cables that are not in the sextuplet being slack (i.e. adjusting their lengths to be larger than the one required for the pose).

We have generalized this approach for a trajectory of a 4 cables CDPR with all cables attached to the same point of the platform. In that case only up to 3 cables may be under tension at the same time. We have designed an algorithm that determine the optimal cable configuration on the whole trajectory.

Simultaneously we have addressed part of an ambitious goal: a full simulation tool for CDPR. We assume a high level motion planning loop that calculate a motion order every Δt_1 second and send this command to an inner motor control loop that execute it by sending a command to the motor every Δt_2 second. Then we have a continuous time model of the motor that determine its velocity. The whole purpose is to calculate the pose of the platform together with the tensions in the cables. This simulation is extremely demanding and cannot be performed with classical software because of the changes in the cable configuration that have to be detected for determining the platform pose and cable tensions. We have succeeded for CDPR with rigid and elastic cables, furthermore introducing random errors in the cable length measurements. This tool has allowed us to show that cable tensions are very sensitive: for example a high level loop that is designed to minimize $\sum \tau_j^2$, where τ are the cable tensions, exhibits large difference with the objective as soon as discrete time-control is taken into account.

6.1.1.2. Certified Calibration of a Cable-Driven Robot Using Interval Contractor Programming

Participants: Julien Alexandre Dit Sandretto, David Daney, Gilles Trombettoni.

An interval based approach is proposed to rigorously identify the model parameters of a parallel cable-driven robot. The studied manipulator follows a parallel architecture having 8 cables to control the 6 DOFs of its mobile platform. This robot is complex to model, mainly due to the cable behavior. To simplify it, some hypotheses on cable properties (no mass and no elasticity) are done. An interval approach can take into account the maximal error between this model and the real one. This allows us to work with a simplified although guaranteed interval model. In addition, a specific interval operator makes it possible to manage outliers. A complete experiment validates our method for robot parameter certified identification and leads to interesting observations [9], [16], [15].

6.1.1.3. Tool for Agencement Analysis and Synthesis of CDPRs

Participants: Laurent Blanchet, Jean-Pierre Merlet [correspondant].

In the frame of FP7 project CABLEBOT, we are developing a methodology to analyze or synthesize a Cable Driven Parallel Robots configuration i.e. either to determine the performances of a given CDPR (e.g. maximal wire tensions over a given workspace) or, being given a list of requirements, to determine what what are **all** possible CDPR geometries that are guaranteed to satisfy the requirements. This tool relies heavily on our analysis of the CDPRs and on interval analysis.

To illustrate this approach we have developed a software that can be used to illustrate the workings/operating procedure of interval analysis through a 3D visualization. This software sets up a scenario of a CDPR in a warehouse and computes in real time its workspace under different constraints.

6.1.1.4. Visual-servoing of a parallel cable-driven robot

Participants: Rémy Ramadour, Jean-Pierre Merlet [correspondant], François Chaumette [correspondant].

`MARIONET-ASSIST` is a parallel cable-driven robot designed to move through large rooms in order to provide services such as walking-aid, lifting people or manipulating heavy loads. In order to experiment, a full-scaled flat with a crane robot has been built. Adding one or several low-cost cameras (the cost being here a fundamental constraint), visual-servoing control is used to provide a whole new set of useful services such as grasping objects in order to bring them to the end-user (if they are too heavy, too far, high or low), or cleaning the table after lunch. Using a parallel crane robot, we are able to cover a large workspace, the vision-control allowing us to obtain the precision required by the manipulation of daily-life objects. The collaborative implementation of the vision and the kinematic control of the robot gives us a way to make best use of the advantages of both parts, while overcoming their respective drawbacks.

This project is supported by the large-scale initiative PAL.

Experimentation showed that we are able to provide a much better accuracy and repeatability using visual-servoing. However, the velocity of the process is slowed because of several encountered problems :

- when there are changes in the distribution of tension between the wires, oscillations are occurring on the end-effector, affecting the movement of the camera in such a way that we can not rely on the measurements
- the methods used to first detect the object are not satisfactory. Also, the actual segmentation is not robust to luminance changes, the target may thus be lost during the process.

In order to overcome the first problem, we are working on an algorithm able to determinate the best sequence of configurations (distribution of tension) : we can avoid singularities and provide a more stable trajectory. The second problem has yet to be solved : we are at the moment looking into several methods, using for example k-nearest neighbors algorithms with different color spaces, gradient-based information and morphological preprocessing.

Finally, we experimented our device with others technologies developed within the context of PAL, in a full-scaled apartment located in Nancy (Loria-Inria).

6.1.2. Assistance robotics

This is now the core of our activity and our work on CDPR is deeply connected to this field as they are an efficient solution for mobility assistance, a high priority for the elderly, helpers and medical community. We have presented our vision of assistance robotics in several occasions [22], [23], [19].

6.1.2.1. Assessment of elderly frailty

Participants: Karim Bakal, Jean-Pierre Merlet.

The assessment of elderly frailty is a difficult concept because it involves the physical capacities of a person and its environment (health-care services, families, funds...). To evaluate the physical abilities, biomechanics tests can be underwent on the upper limb, lower limb or the whole body. In particularly, the motricity of the upper limb can be measured in terms of range of motion, velocity, acceleration or forces.

To analyze the velocity of the loads in the upper limb, a polytope interpretation is used. Currently the force polytope at the hand is calculated from the torques τ measured at each joint (shoulder, elbow and wrist) by a dynamometer (Biodex III, Biodex Medical Systems). But because of the redundancy of the upper limb (7 degrees of freedom), the dynamic equation ($\tau = \mathbf{J}^T F$) is difficult to solve. To find the minimal and maximal forces F that can be exerted at the hand from the measured torques, we may use the jacobian pseudo-inverse with the method of Chiacchio but this method is not well suited to manage the large uncertainties in the measurements. In the a reverse approach, the force at the hand will be measured by a 6-axis load sensor and the minimal and maximal joint torques will be computed by using interval analysis and compared with the measurements of the Biodex.

Moreover this analysis of the force capacities in the upper limb need to be connected to the daily activities or usual motion test monitored by the medical services. Therefore, a review of tests and questionnaires regularly used to measure the physical capacities has been performed. This review gather the type of mark, the exercises and the used sensors that can be employed in future experimentation. Also, this review will be discussed with medical staff to highlight relevant activities.

6.1.2.2. Walking analysis

Participants: Claire Dune, Ting Wang, Jean-Pierre Merlet [correspondant].

In the period 2009-2013 we have conducted in collaboration with Nice hospital a large experiment involving 54 subjects (30 elderly and 24 young adults) for determining walking pattern of elderly people using our instrumented walker `ANG-light`. We have started the processing of this large amount of data we some interesting results [25]:

- a classical walking test is the 10 meter walking test: the subject is asked to perform a 10m straight line trajectory and the result is the total time. Such test may have large consequences as it is used to determine the autonomy level and the resulting financial aid. Our test has surprisingly shown that when using a walker elderly people are usually faster that young adults
- on the other hand the maximal deviation with respect to the desired trajectory is much smaller for young adults than for elderly one. Furthermore few elderly have the same deviation and it may be considered as a signature of the walking pattern that is worth measuring

Our objective is now to analyze the maneuvers (half-turn and round-about) and to compare/complement the data with the one obtained with a Kinect. A long term objective is also to implement a model of a human walking with a walker and to use this model for an inverse calculation: measuring walking patterns indicators with the walker and calculating these indicators when not using the walker.

6.1.3. Experimental calibration of a high-accuracy space telescope

Participants: Thibault Gayral, David Daney, Jean-Pierre Merlet.

A collaborative work began in October 2010 with Thales Alenia Space on the calibration of the mechanical structure of a space telescope. Its architecture is based on a parallel manipulator (type active wrist 6-PUS) used to correct the relative position of two mirrors. The aim is to reach a micrometer accuracy in order to obtain a suitable quality of the images provided by the telescope. Thus, a complete model of the space telescope needs to be developed and validated through calibration. Since high velocity is not required in such an application, the dynamic effects can be neglected and only geometric and/or static calibration has to be considered.

For the geometric models, measurements for calibration were performed in a clean room under controlled pressure, temperature and humidity conditions to minimize the influence of the non-geometric errors. Thus, two possible static inaccuracy sources were identified and modeled: one from the deformation of the mobile platform and the other resulting from the behavior of the flexure joints. Three incremental models of the flexure joints were developed and compared: a spherical joint model, a model issued from the beam theory and a stiffness model. Results of calibration using an accurate measurement system of photogrammetry showed that the flexure joints can be modeled by perfect spherical joints due to the small workspace of the telescope. Concerning the mobile platform deformation, two models were developed. With those models, a positioning accuracy of some micrometers was finally reached after calibration with only position and orientation measurements of the mobile platform.

Then, opto-mechanical models were developed considering experimental measurements by imaging on the prototype of the space telescope. The optical defects were analyzed considering Zernike polynomials. The aim of optical calibration was to minimize the coefficients of the Zernike polynomials in order to improve the optical properties of the space telescope. Results of calibration were studied in order to perform a proper choice of the opto-mechanical models. Finally, the optical quality was improved after calibration. This validates the fact that the telescope can be calibrated directly in space, after its deployment, with only the provided information. A second campaign of measurements by imaging was programmed to finely adjust the opto-mechanical model parameters.

6.2. Miscellaneous results

6.2.1. Symbolic tools for modeling and simulation

Participant: Yves Papegay.

This activity is the main part of a long-term ongoing collaboration with Airbus whose goal is to directly translate the conceptual work of aeronautics engineers into digital simulators to accelerate aircraft design.

An extensive modeling and simulation platform has been designed which includes a dedicated modeling language for the description of aircraft dynamics models in term of formulae and algorithms, and a symbolic compiler producing as target an efficient numerical simulation code ready to be plugged into a flight simulator, as well as a formatted documentation compliant with industrial requirements of corporate memory.

Implementation of this platform is a modeling and simulation environment based on symbolic computation tools. It contains several components :

- a model editor, that makes it possible and easy to enter the whole set of equations describing large and complex industrial models,
- an highly interactive and modular evaluation workbench allowing to simulate the models and to visualize the results inside the modeling environment with the benefits for the designer of being able to directly use all its computational functionnalities.
- a C code generator which, using these models, automatically generates the numerical real-time simulation engines
- a technical documentation generator

Technology demonstrated by our prototype has been transferred to our industrial partner in 2012 when final version of our modeling and simulation environment has been delivered to Airbus in November 2012.

However, in 2013, we have worked on several enhancements and extension of functionnalities, namely to ease the integration of our environment into the airbus toolbox. Developer level know-how has been transferred to a software company in charge of industrialization and maintenance of the modeling and simulation environment.

6.2.2. Multi-agent aircraft design

Participant: Yves Papegay.

The modeling environment described in the previous section is used, in collaboration with other teams at Airbus, in the framework of the ID4CS project founded by ANR and dedicated to multi-agent optimization of large scale system.

Several models of aircraft engines and of aircrafts have been developed as user cases for the project.

2013 is the last year of the project when agent code based on models has been used to solve several practical optimization problems based on these models.

6.2.3. Equilibrium strategies for linked Electricity and CO2 markets

Participant: Odile Pourtallier.

In collaboration with M. Bossy (Inria -TOSCA Team) and N. Maïzi (CMA - Mines Paristech) O. Pourtallier we have pursued our work on CO2 and electricity market coupling.

The aim of this work is to develop analytic tools, in order to design a relevant mechanism for carbon markets, where relevant refers to emission reduction. In the context of electricity, the number of producers is limited, a standard game theory approach applies. The producers are considered as players behaving on the two financial markets represented here by carbon and electricity. We establish a Nash equilibrium for this non-cooperative J-player game through a coupling mechanism between the two markets.

The original idea comes from the French electricity sector, where the spot electricity market is often used to satisfy peak demand. Producers behavior is demand driven and linked to the maximum level of electricity production. Each producer strives to maximize its market share. In the meantime, it has to manage the environmental burden associated with its electricity production through a mechanism inspired by the EU ETS (European Emission Trading System) framework : each producer emission level must be balanced by a permit or through the payment of a penalty. Emission permit allocations are simulated through a carbon market that allows the producers to buy the allowances at an auction.

Based on a static elastic demand curve (referring to the times stages in an organized electricity market, mainly day-ahead and intra-day), we solve the local problem of establishing a non-cooperative Nash equilibrium for the two coupled markets.

CORIDA Project-Team

5. New Results

5.1. Analysis and control of fluids and of fluid-structure interactions

In [47], we analyze the system fluid-rigid body in the case of where the rigid body is a ball of "small radius". More precisely, we consider the limit system as the radius goes to zero. We recover the Navier-Stokes system with a particle following the the velocity of the fluid. We consider in [45] a model of vesicle moving into a viscous incompressible fluid. Such a model, based on a phase-field approach was derived by researchers in Physics, and is quite difficult to study. By considering some approximation, we prove some result of existence of solutions for such a system.

By acting on a part of the fluid domain or on a part of the exterior boundary, we aim at controlling the fluid velocity, the rigid velocity and the position of the rigid body. It can be a control in open loop or in closed loop. We have studied both problems in the 1D case. In this case, the study benefits some simplifications, but can also be more difficult since the fluid domain is no more connected. As a consequence, if one wants to control by using only one input, on one part of the fluid domain, the fluid on the other side of the particle is only controlled by the motion of the structure.

We introduce a new method for controllability of nonlinear parabolic system allowing to deal with this problem and we solve it in ([24]). We also obtain the local stabilization of such system around a stationary state in [41].

We study the Cauchy problem corresponding to a similar 1D system without viscosity in [40]. In that case, we have to deal with the interaction between the particle and shock waves or relaxation waves. In [44], we analyze a numerical scheme for the method of observers used to reconstruct the initial data of hyperbolic systems such as wave equation. We add some numerical viscosity in the scheme in order to have a uniform decay of the error between the reconstructed solution and the real one.

In [30], a Lagrange-Galerkin method is introduced to approximate a two dimensional fluid-structure interaction problem for deformable solids. The new numerical scheme we present is based on a characteristics function mapping the approximate deformable body at the discrete time level t_{k+1} into the approximate body at time t_k.

The aim of [25] is to tackle the time optimal controllability of an $(n+1)$-dimensional nonholonomic integrator with state constraints. A full description of an optimal control together with the corresponding optimal trajectories are explicitly obtained. The optimal trajectories we construct, are composed of arcs of circle lying in a 2-dimensional plane.

In [26], controllability results are obtained for a low Reynolds number swimmer composed by a spherical object which is undergoing radial and axi-symmetric deformations in order to propel itself in a viscous fluid governed by the Stokes equations. A time optimal control problem is also solved for a simplified model and explicit optimal solutions are constructed.

5.2. Frequency domain methods for the analysis and control of systems governed by PDE's

With a numerical viscosity terms in the approximation scheme of second order evolution equations, we show in [11] the exponential or polynomial decay of the discrete scheme when the continuous problem has such a decay and when the spectrum of the spatial operator associated with the undamped problem satisfies the generalized gap condition. We further show the convergence of the discrete solution to the continuous one.

In [19], we propose a strategy to determine the Dirichlet-to-Neumann (DtN) operator for infinite, lossy and locally perturbed hexagonal periodic media, using a factorization of this operator involving two non local operators. The first one is a DtN type operator and corresponds to a half-space problem, while the second one is a Dirichlet-to-Dirichlet (DtD) type operator related to the symmetry properties of the problem.

In [22], we generalize to the case of acoustic penetrable scatterers the results derived by Hazard and Ramdani [54] for sound hard scatterers. In particular, we provide a justification of the DORT method in this case and we show that each small inhomogeneity gives rise to $3d+1$ eigenvalues of the time reversal operator. The selective focusing of the corresponding eigenfunctions is also proved.

In [17], we consider the inverse problem of determining the potential in the dynamical Schrödinger equation on the interval by the measurement on the boundary. We use the Boundary Control Method to recover the spectrum of the problem from the observation at either left or right end points. Using the specificity of the one-dimensional situation we recover the spectral function, reducing the problem to the classical one which could be treated by known methods. We also consider the case where only a finite number (N) of eigenvalues are available and we prove the convergence of the reconstruction method as N tends to infinity.

We give some spectral and condition number estimates of the acoustic single-layer operator for low-frequency multiple scattering in dense [15] and dilute [16] media.

5.3. Use of geometric techniques for the control of finite and infinite dimensional systems

The paper [31] deals with the design of high gain observers for a class of continuous dynamical systems with discrete-time measurements. The new idea of the this work is to synthesize an observer requiring the less knowledge as possible from the output measurements. This is done by using an updated sampling time observer.

In [12], it is shown that, for a bilinear system, the property of observability is preserved after sampling provided that the controls take their values in a compact space and do not vary too quickly.

In the note [18] two notions of controllability are studied, called respectively radial controllability and directional controllability. It is proven that for families of linear vector fields, the two notions are actually equivalent.

We used operators theory to obtain some new estimates of the energy of an infinite dimensional bilinear quantum systems. These results were presented in [34].

Robust control of bilinear Schrödinger equation was investigated in [35]. The use of sharp finite dimensional energy estimates (in the spirit of [34]) allows to obtain the first approximate ensemble controllability results for infinite dimensional quantum systems, also in presence of mixed spectrum for the free Hamiltonian.

The above energy questions, together with a their relation with some open question in the control of bilinear quantum systems, were gathered in the survey [32].

Our team is heavily involved in the optimization of driving strategy, and especially in the effective implementation in the prototype build in ESSTIN. MPC related methods have been tested and successfully improved as described in [37].

CORTEX Team

6. New Results

6.1. Understanding embodied neural systems

Participants: Dominique Martinez, Carlos Carvajal-Gallardo, Georgios Detorakis.

6.1.1. Bio-physical modeling and embodied olfaction

Our understanding of the computations that take place in the human brain is limited by the extreme complexity of the cortex, and by the difficulty of experimentally recording neural activities, for practical and ethical reasons. The Human Genome Project was preceded by the sequencing of smaller but complete genomes. Similarly, it is likely that future breakthroughs in neuroscience will result from the study of smaller but complete nervous systems, such as the insect brain or the rat olfactory bulb. These relatively small nervous systems exhibit general properties that are also present in humans, such as neural synchronization and network oscillations. Our goal has been therefore to understand the role of these phenomena by combining biophysical modelling and experimental recordings, before applying this knowledge to humans. In the last year, we have extended our neuronal model of the insect olfactory system. This model is capable of reproducing and explaining the stereotyped multiphasic firing pattern observed in pheromone sensitive antennal lobe neurons [10].

Using this model in robotic experiments and insect antennae as olfactory sensors, we related these multiphasic responses to action selection. The efficiency of the model for olfactory searches was demonstrated in driving the robot toward a source of pheromones. Two different classes of strategies are possible for olfactory searches, those based on a spatial map, e.g. Infotaxis, and those where the casting-and-zigzagging behaviour observed in insects is purely reactive, without any need for an internal memory, representation of the environment, or inference [15]. Our goal was to investigate this question by implementing infotactic and reactive search strategies in a robot and test them in real environmental conditions. We previously showed that robot Infotaxis produces trajectories that feature zigzagging and casting behaviours similar to those of moths, is robust and allows for rapid and reliable search processes. We have implemented infotactic and reactive search strategies in a cyborg using the antennae of a tethered moth as sensors, since no articial sensor for pheromone molecules is presently known [10].

6.1.2. Somato-sensory cortex

In a joint work with the Mnemosyne team, we have investigated the formation and maintenance of ordered topographic maps in the primary somatosensory cortex as well as the reorganization of representations after sensory deprivation or cortical lesion. We consider both the critical period (postnatal) where representations are shaped and the post-critical period where representations are maintained and possibly reorganized. We hypothesize that feed-forward thalamocortical connections are an adequate site of plasticity while cortico-cortical connections are believed to drive a competitive mechanism that is critical for learning. We model a small skin patch located on the distal phalangeal surface of a digit as a set of 256 Merkel ending complexes (MEC) that feed a computational model of the primary somatosensory cortex (area 3b). This model is a two-dimensional neural field where spatially localized solutions (a.k.a. bumps) drive cortical plasticity through a Hebbian-like learning rule. Simulations explain the initial formation of ordered representations following repetitive and random stimulations of the skin patch. Skin lesions as well as cortical lesions are also studied and results confirm the possibility to reorganize representations using the same learning rule and depending on the type of the lesion. For severe lesions, the model suggests that cortico-cortical connections may play an important role in complete recovery [11], [19], [7].

6.1.3. K-cells in visuomotor tasks

In another joint work with the Mnemosyne team, we have explored the role of the thalamus in visuomotor tasks implicating non-standard ganglion cells. Such cells in the retina have specific loci of projection in the visuomotor systems and particularly in the thalamus and the superior colliculus. In the thalamus, they feed the konio pathway of the LGN. Exploring the specificities of that pathway, we discovered it could be associated to the matrix system of thalamo-cortical projections, known to allow for diffuse patterns of connectivity and to play a major role in the synchronization of cortical regions by the thalamus. An early model led to the design of the corresponding information flows in the thalamo-cortical system, that we expanded, in the framework of the Keops project, to be applied to real visuomotor tasks [13].

We proposed to implement the computational principles raised by the study on the K-cells of the retina using a variational specification of the visual front-end, with an important consequence. In such a framework, the GC are not to be considered individually, but as a network, yielding a mesoscopic view of the retinal process. Given natural image sequences, fast event-detection properties appear to be exhibited by the mesoscopic collective non-standard behavior of a subclass of the so-called dorsal and ventral konio-cells (K-cells) that correspond to specific retinal output. We considered this visual event detection mechanism to be based on image segmentation and specific natural statistical recognition, including temporal pattern recognition, yielding fast region categorization. We discussed how such sophisticated functionalities could be implemented in the biological tissues as a unique generic two-layered non-linear filtering mechanism with feedback. We used computer vision methods to propose an effective link between the observed functions and their possible implementation in the retinal network. The available computational architecture is a two-layers network with non-separable local spatio-temporal convolution as input, and recurrent connections performing non-linear diffusion before prototype based visual event detection [17].

6.2. Neuro-inspired computational models

Participants: Yann Boniface, Benoît Chappet de Vangel, Bernard Girau, Patrick Hénaff.

6.2.1. Motion detection

We develop bio-inspired neural architectures to extract and segment the direction and speed components of the optical flow from sequences of images. Following this line, we have built additional models to code and distinguish different visual sequences. The structure of these models takes inspiration from the course of visual movement processing in the human brain, such as in area MT (middle temporal) that detects patterns of movement, or area FBA where neurons have been found to be sensitive to single spatio-temporal patterns. This work has been extended to complex movements: to fight, to wave, to clap, using real-world video databases [9].

6.2.2. Multimodal learning through joint dynamic neural fields

We have developed a coherent multimodal learning for a system with multiple sensory inputs. To this aim, we modified the BCM synaptic rule, a local learning rule, to obtain the self organization of our neuronal inputs maps and we used a CNFT based competition to drive the BCM rule. In practice, we introduced a feedback modulation of the learning rule, representing multimodal constraints of the environment. We also introduced an unlearning term in the BCM equation to solve the problem of the different temporalities between the raise of the activity within modal maps and the multimodal learning of the organization of the maps [12].

6.2.3. Adaptive sensori-motor loop

We develop bio-inspired neural controllers to control humanoids robot when they interact physically (or socially) with the human. We focus on the role of rythmicity in the interaction: how the phenomena of coupling, synchrony or others are involved in the interaction between humans? what models of neural structures can incorporate rythmicity intrinsically, and can include learning or adaptive mechanisms of the rythmicity.

6.2.4. Randomly spiking dynamic neural fields

We have defined a new kind of spiking neural field that is able to use only local links while transmitting spikes through the map by succesive random propagations. Such a model is able to be mapped onto FPGAs, while maintaining most properties of neural fields. This model has been validated from a behavioral point of view, and a fully scalable hardware implementation has been designed with several thousands of neurons on-chip. These first results are the object of an article that is currently reviewed after requested revisions.

CQFD Project-Team

6. New Results

6.1. Nonparametric estimation of the jump rate for non-homogeneous marked renewal processes.

Participants: Romain Azais, François Dufour, Anne Gégout-Petit.

Non-homogeneous marked renewal process, nonparametric estimation, jump rate estimation, Nelson-Aalen estimator, asymptotic consistency, ergodicity of Markov chains

This work is devoted to the nonparametric estimation of the jump rate and the cumulative rate for a general class of non-homogeneous marked renewal processes, defined on a separable metric space. In our framework, the estimation needs only one observation of the process within a long time. Our approach is based on a generalization of the multiplicative intensity model, introduced by Aalen in the seventies. We provide consistent estimators of these two functions, under some assumptions related to the ergodicity of an embedded chain and the characteristics of the process. A numerical example illustrates our theoretical results.

It has been published in Ann. Inst. H. Poincaré Probab. Statist. [16].

6.2. Nonparametric estimation of the conditional distribution of the inter-jumping times for piecewise-deterministic markov processes

Participants: Romain Azais, François Dufour, Anne Gégout-Petit.

Piecewise-deterministic Markov process, ergodicity of Markov chains, nonparametric estimation, jump rate estimation, Nelson-Aalen estimator, asymptotic consistency

In this work, we present a nonparametric method for estimating the conditional density associated to the jump rate of a piecewise-deterministic Markov process. In our framework, the estimation needs only one observation of the process within a long time interval. Our method relies on a generalization of Aalen?s multiplicative intensity model. We prove the uniform consistency of our estimator, under some reasonable assumptions related to the primitive characteristics of the process. A simulation study illustrates the behavior of our estimator.

It has been accepted for publication in Scandinavian Journal of Statistics [17].

6.3. Finite Linear Programming Approximations of constrained discounted Markov decision processes

Participant: François Dufour.

Constrained Markov decision processes, linear programming approach to control problems, quantization, approximation of Markov decision processes

We consider a Markov decision process (MDP) with constraints under the total expected discounted cost optimality criterion. We are interested in proposing approximation methods of the optimal value of this constrained MDP. To this end, starting from the linear programming (LP) formulation of the constrained MDP (on an infinite-dimensional space of measures), we propose a finite state approximation of this LP problem. This is achieved by suitably approximating a probability measure underlying the random transitions of the dynamics of the system. Explicit convergence orders of the approximations of the optimal constrained cost are obtained. By exploiting convexity properties of the class of relaxed controls, we reduce the LP formulation of the constrained MDP to a finite-dimensional static optimization problem, that can be used to obtain explicit numerical approximations of the corresponding optimal constrained cost. A numerical application illustrates our theoretical results.

These results have been obtained in collaboration with Tomas Prieto-Rumeau, Department of Statistics and Operations Research, UNED, Madrid, Spain.

It has been published in SIAM Journal of Control and Optimization [25].

6.4. Stochastic Approximations of Constrained Discounted Markov Decision Processes

Participant: François Dufour.

Constrained Markov decision processes; Linear programming approach to control problems; Approximation of Markov decision processes

We consider a discrete-time constrained Markov decision process under the discounted cost optimality criterion. The state and action spaces are assumed to be Borel spaces, while the cost and constraint functions might be unbounded. We are interested in approximating numerically the optimal discounted constrained cost. To this end, we suppose that the transition kernel of the Markov decision process is absolutely continuous with respect to some probability measure μ. Then, by solving the linear programming formulation of a constrained control problem related to the empirical probability measure μ_n of μ, we obtain the corresponding approximation of the optimal constrained cost. We derive a concentration inequality which gives bounds on the probability that the estimation error is larger than some given constant. This bound is shown to decrease exponentially in n. Our theoretical results are illustrated with a numerical application based on a stochastic version of the Beverton-Holt population model.

These results have been obtained in collaboration with Tomas Prieto-Rumeau, Department of Statistics and Operations Research, UNED, Madrid, Spain.

It has been accepted for publication in Journal of Mathematical Analysis and Applications [26].

6.5. The expected total cost criterion for Markov decision processes under constraints

Participant: François Dufour.

Markov decision process, expected total cost criterion, constraints, linear programming, occupation measure

In this work, we study discrete-time Markov decision processes (MDPs) with constraints when all the objectives have the same form of expected total cost over the infinite time horizon. Our objective is to analyze this problem by using the linear programming approach. Under some technical hypotheses, it is shown that if there exists an optimal solution for the associated linear program then there exists a randomized stationary policy which is optimal for the MDP, and that the optimal value of the linear program coincides with the optimal value of the constrained control problem. A second important result states that the set of randomized stationary policies provides a sufficient set for solving this MDP. It is important to notice that, in contrast with the classical results of the literature, we do not assume the MDP to be transient or absorbing. More importantly, we do not impose the cost functions to be non-negative or to be bounded below. Several examples are presented to illustrate our results.

These results have been obtained in collaboration with Alexey Piunovskiy from Department. of Mathematical Sciences.

It has been published in Advances in Applied Probability [24].

6.6. Optimal stopping for piecewise-deterministic Markov processes and applications

Participants: Adrien Brandejsky, Benoîte de Saporta, François Dufour, Huilong Zhang.

We worked further on numerical methods for optimal stopping of PDMPs. On the one hand, we applied our numerical method to compute an optimal maintenance date to the test case of the heated hold-up tank. The system consists of a tank containing a fluid whose level is controlled by three components: two inlet pumps and one outlet valve. A thermal power source heats up the fluid. The failure rates of the components depends on the temperature, the position of the three components monitors the liquid level in the tank and the liquid level determines the temperature. Therefore, this system can be modeled by a hybrid process where the discrete (components) and continuous (level, temperature) parts interact in a closed loop. We model the system by a piecewise deterministic Markov process, propose and implement a numerical method to compute the optimal maintenance date to repair the components before the total failure of the system. This work is published in [30].

On the other hand, we investigated the optimal stopping problem under partial observations for PDMPs. We first obtain a recursive formulation of the optimal filter process and derive the dynamic programming equation of the partially observed optimal stopping problem. Then, we propose a numerical method, based on the quantization of the discrete-time filter process and the inter-jump times, to approximate the value function and to compute an ϵ-optimal stopping time. We prove the convergence of the algorithms and bound the rates of convergence. This work is published in [20].

6.7. Stochastic control for underwater optimal trajectories

Participants: Benoîte de Saporta, François Dufour, Huilong Zhang.

This work aims to compute optimal trajectories for underwater vehicles evolving in a given environment to accomplish some tasks. This is an optimal control problem. In real context, available inputs are not perfectly known. Hence a stochastic approach seems to be needed, coupled with the outputs of the tracking algorithms. Markov decision processes (MDPs) constitute a general family of controlled stochastic processes suitable for the modeling of sequential decision-making problems. The analysis of MDPs leads to mathematical and computational problems. The corresponding theory has reached a rather high degree of maturity, although the classical tools (such as value iteration, policy iteration, linear programming, and their various extensions) are generally hardly applicable in practice. Hence, solving MDPs numerically is an awkward and important problem. The method is applied to control a submarine which wants to well detect one or several targets and only has the information given by the tracking algorithms from the sonar observations [47].

6.8. Modeling of cell division data

Participants: Benoîte de Saporta, Anne Gégout-Petit.

This work is in collaboration with Laurence Marsalle (Univ. Lille 1).

A rigorous methodology is proposed to study cell division data consisting in several observed genealogical trees of possibly different shapes. The procedure takes into account missing observations, data from different trees, as well as the dependence structure within genealogical trees. Its main new feature is the joint use of all available information from several data sets instead of single data set estimation, to avoid the drawbacks of low accuracy for estimators or low power for tests on small single trees. The data is modeled by an asymmetric bifurcating autoregressive process and possibly missing observations are taken into account by modeling the genealogies with a two-type Galton-Watson process. Least-squares estimators of the unknown parameters of the processes are given and symmetry tests are derived. Results are applied on real data of Escherichia coli division and an empirical study of the convergence rates of the estimators and power of the tests is conducted on simulated data. This work is to appear in [29].

We have also presented a new model of asymmetric bifurcating autoregressive process with random coefficients. We couple this model with a Galton-Watson tree to take into account possibly missing observations. We propose least-squares estimators for the various parameters of the model and prove their consistency, with a convergence rate, and asymptotic normality. We use both the bifurcating Markov chain and martingale approaches and derive new results in both these frameworks. This work is to appear in [28].

6.9. Numerical method for the filtering of Markov jump linear systems

Participants: Benoîte de Saporta, Eduardo Costa.

We are interested in efficient pre-computations of the solutions of Markov switching Riccati equations. These equations are matrix-valued and naturally arise in control or filtering problems for Markov jump linear systems. It is crucial for applications to be able to compute the filter in real time, although the solutions to Riccati equations are slow to compute. Hence the need for pre-computations, taking into account the random possible changes of regimes. We propose a numerical method based on the discretization by quantization of the underlying Markov chain.

6.10. Optimization of the assembly line of the future European launcher

Participants: Benoîte de Saporta, François Dufour, Christophe Nivot.

In collaboration with Astrium space transportation, we have started working on the optimization of the assembly line of the future European launcher. We have started with a simplified model with five components to be assembled in workshops liable to breakdowns. We have modeled the problem using the Markov Decision Processes (MDP) framework and built a simulator of the process in order to run an optimization procedure

6.11. A variable clustering approach for the typology of units: a survey on farming and environment

Participants: Jérôme Saracco, Marie Chavent.

A survey on farming and environment dealing with the current transformations of the farmer job is considered. We propose to replace the usual data mining strategy which consists of applying Multiple Correspondence Analysis by a variable clustering approach. Clustering of variables aims at lumping together variables which are strongly related to each other and thus bring the same information. The ClustOfVar approach used in this paper provides at the same time groups of variables and their associated synthetic variables. In this algorithm, the homogeneity criterion of a cluster is defined by the squared Pearson correlation for the quantitative variables and by the correlation ratio for the qualitative variables. The step of variable clustering enables to get synthetic variables that can be read as a gradient. In our case study, values correspond to some relevant groupings of categories. This enables to interpret and name easily the synthetic variables. Trends in the opinion of farmers are thus highlighted with the variable clustering approach. Then we clarify these first results by applying a clustering method on the scores of the individuals measured by the synthetic variables. At the sociological level, the supply provided by the synthetic variables to interpret the clusters of farmers is obvious.

These results have been obtained in collaboration with Vanessa Kuentz from Irstea, UR ADBX.

They have been published in Journal de la Société Française de Statistique [31].

6.12. Multiple Facctorial Analysis for mixed data type

Participants: Jérôme Saracco, Marie Chavent, Amaury Labenne.

Multiple Factor Analysis (MFA) originally proposed by Escofier and Pages in 1982 is a method dedicated to the study of a set of n individuals described by groups of quantitative variables. Later, this method was extended to take into account groups of qualitative variables (Pages, 1983) then simultaneously quantitative groups and qualitative groups (Pages, 2002). However, this method does not currently take into account mixed groups, that is to say containing both quantitative and qualitative variables. The aim of our study is to propose sustainable development indicators by integrating the aspect of quality of life. For that, we are confronted with the analysis of groups of variables with quantitative and qualitative variables. In this work, we propose an extension of the MFA method, called MFAMIX, for the multiple factor analysis of mixed groups of variables. This approach relies on a combination of AFM and PCAMIX method that allows the analysis of mixed data. MFAMIX method is presented using a singular value decomposition and illustrated on socio-economic data about the quality of life.

These results have been obtained in collaboration with Vanessa Kuentz from Irstea, UR ADBX.

They have been have been presented in two national conferences [43], [41].

6.13. Detecting mental states of alertness with genetic algorithm variable selection

Participants: Marie Chavent, Laurent Vézard.

The objective of the present work is to develop a method able to automatically determine mental states of vigilance; i.e., a person's state of alertness. Such a task is relevant to diverse domains, where a person is expected or required to be in a particular state. For instance, pilots or medical staffs are expected to be in a highly alert state, and this method could help to detect possible problems. In this paper, an approach is developed to predict the state of alertness ("normal" or "relaxed") from the study of electroencephalographic signals (EEG) collected with a limited number of electrodes. The EEG of 58 participants in the two alertness states (116 records) were collected via a cap with 58 electrodes. After a data validation step, 19 subjects were retained for further analysis. A genetic algorithm was used to select an optimal subset of electrodes. Common spatial pattern (CSP) coupled to linear discriminant analysis (LDA) was used to build a decision rule and thus predict the alertness of the participants. Different subset sizes were investigated and the best result was obtained by considering 9 electrodes (correct classification rate of 73.68

These results have been obtained in collaboration with Pierrick Legrand from Alea Inria team and Leonardo Trujillo from Instituto Tecnologico de Tijuana.

This work has been presented in a international IEEEl conference [38].

6.14. ClustOfVar : an R package for dimension reduction via clustering of variables. Application in supervised classification and variable selection in gene expressions data

Participants: Marie Chavent, Jérôme Saracco.

The main goal of this work is to tackle the problem of dimension reduction for high-dimensional supervised classication. The motivation is to handle gene expression data. The proposed method works in 2 steps. First, one eliminates redundancy using clustering of variables, based on the R-package ClustOfVar. This first step is only based on the exploratory variables (genes). Second, the synthetic variables (summarizing the clusters obtained at the first step) are used to construct a classifier (e.g. logistic regression, LDA, random forests). We stress that the first step reduces the dimension and gives linear combinations of original variables (synthetic variables). This step can be considered as an alternative to PCA. A selection of predictors (synthetic variables) in the second step gives a set of relevant original variables (genes). Numerical performances of the proposed procedure are evaluated on gene expression datasets. We compare our methodology with LASSO and sparse PLS discriminant analysis on these datasets.

This work is a collaboration with Robin Genuer from SISTM Inria team and Benoit Liquet from University of Queensland.

This work has been presented in a international workshop on Statistical Methods for (post)-Genomics Data (SMPGD 2013) [42].

6.15. A sliced inverse regression approach for data stream

Participants: Jérôme Saracco, Marie Chavent.

This work is in collaboration with Stéphane Girard (Inria Grenoble Alpes), Benoît Liquet (MRC, Cambridge University), Vanessa Kuentz (Irstea) and Thi Mong Gnoc Nguyen (Univ. de Strasbourg).

In this work, we focus on data arriving sequentially by blocks in a stream. A semiparametric regression model involving a common EDR (Effective Dimension Reduction) direction is assumed in each block. Our goal is to estimate this direction at each arrival of a new block. A simple direct approach consists of pooling all the observed blocks and estimating the EDR direction by the SIR (Sliced Inverse Regression) method. But in practice, some disadvantages appear such as the storage of the blocks and the running time for large sample sizes. To overcome these drawbacks, we propose an adaptive SIR estimator of based on the optimization of a quality measure. The corresponding approach is faster both in terms of computational complexity and running time, and provides data storage benefits. The consistency of our estimator is established and its asymptotic distribution is given. An extension to multiple indices model is proposed. A graphical tool is also provided in order to detect changes in the underlying model, i.e., drift in the EDR direction or aberrant blocks in the data stream. A simulation study illustrates the numerical behavior of our estimator. Finally, an application to real data concerning the estimation of physical properties of the Mars surface is presented.

This work is to appear in [21].

6.16. Comparison of Kernel Density Estimators with Assumption on Number of Modes

Participant: Jérôme Saracco.

This work is in collaboration with Bernard Bercu (Univ. Bretagne Sud) and Thi Mong Gnoc Nguyen (Univ. de Strasbourg).

In this work, we investigate the asymptotic behavior of the Nadaraya-Watson estimator for the estimation of the regression function in a semiparametric regression model. On the one hand, we make use of the recursive version of the sliced inverse regression method for the estimation of the unknown parameter of the model. On the other hand, we implement a recursive Nadaraya-Watson procedure for the estimation of the regression function which takes into account the previous estimation of the parameter of the semiparametric regression model. We establish the almost sure convergence as well as the asymptotic normality for our Nadaraya-Watson estimator. We also illustrate our semiparametric estimation procedure on simulated data.

This work is to appear in [19].

6.17. Comparison of Kernel Density Estimators with Assumption on Number of Modes

Participants: Jérôme Saracco, Raphaël Coudret.

This work is in collaboration with Gilles Durrieu (Univ. Bretagne Sud).

A data-driven bandwidth choice for a kernel density estimator called critical bandwidth is investigated. This procedure allows the estimation to have as many modes as assumed for the density to estimate. Both Gaussian and uniform kernels are considered. For the Gaussian kernel, asymptotic results are given. For the uniform kernel, an argument against these properties is mentioned. These theoretical results are illustrated with a simulation study which compare the kernel estimators that rely on critical bandwidth with another one which uses a plug-in method to select its bandwidth. An estimator that consists in estimates of density contour clusters and takes assumptions on number of modes into account is also considered. Finally, the methodology is illustrated using environment monitoring data.

This work is to appear in [22].

6.18. Comparison of sliced inverse regression approaches for underdetermined cases

Participants: Jérôme Saracco, Raphaël Coudret.

This work is in collaboration with Benoît Liquet (MRC, Cambridge University).

Among methods to analyze high-dimensional data, the sliced inverse regression (SIR) is of particular interest for non-linear relations between the dependent variable and some indices of the covariate. When the dimension of the covariate is greater than the number of observations, classical versions of SIR cannot be applied. Various upgrades were then proposed to tackle this issue such as regularized SIR (RSIR) and sparse ridge SIR (SR-SIR), to estimate the parameters of the underlying model and to select variables of interest. In this paper, we introduce two new estimation methods respectively based on the QZ algorithm and on the Moore-Penrose pseudo-inverse. We also describe a new selection procedure of the most relevant components of the covariate that relies on a proximity criterion between submodels and the initial one. These approaches are compared with RSIR and SR-SIR in a simulation study. Finally we applied SIR-QZ and the associated selection procedure to a genetic dataset in order to find markers that are linked to the expression of a gene. These markers are called expression quantitative trait loci (eQTL).

This work was presented in a national conference [23] and is to appear in [37].

6.19. Conditional Quantile Estimationthrough Optimal Quantization

Participants: Jérôme Saracco, Isabelle Charlier.

This work is in collaboration with Davy Paindaveine (Univ. Libre de Bruxelles).

In this work, we construct a nonparametric estimator of conditional quantiles of Y given X via optimal quantization. In a first step, we propose to approximate conditional quantiles thanks to optimal quantization in L^p-norm, consisting in discretizing X and Y thanks to some optimal grids of size N. We state a result of convergence of this approximation toward the true conditional quantile. The estimator was implemented in **R** in order to evaluate its numerical behavior and to compare it with existing estimators. A simulation study illustrates the good behavior of our estimator. The practical choice of N is discussed. We apply our approach to a real data set.

This work was presented in a national conference [35].

6.20. Estimation of water consumption based on survey techniques using an automatic meter reading sample

Participant: Jérôme Saracco.

This work is in collaboration with Karim Claudio (LyRE), Vincent Couallier (Univ. de Bordeaux) and Yves Le Gat (Irstea).

Automatic water meters reading are, nowadays, the best technology for real time knowledge of water consumption. At an hydraulic sector scale, a complete equipment permits to know the total consumption of a finite size population, for a time scale as small as the hour. However its cost for generalization is sometimes unbearable for the collectivity, for whom sampling techniques have to be set up. In a objective of a total consumption estimation, this article describes and compares standard methods of survey techniques and propose to retain a methodology for implementation of an operational sample and to calibrate the corresponding total estimator.

This work was presented in a national conference [36] and an associated paper is currently in revision.

6.21. Hidden Markov Model for the detection of a degraded operating mode of optronic equipment

Participants: Camille Baysse, Anne Gégout-Petit, Jérôme Saracco.

This work is in collaboration with Didier Bihannic (Thales Optronics) and Michel Prenat (Thales Optronics).

As part of optimizing the reliability, Thales Optronics now includes systems that examine the state of its equipment. The aim of this work is to use hidden Markov Model to detect as soon as possible a change of state of optronic equipment in order to propose maintenance before failure. For this, we carefully observe the dynamic of a variable called “cool down time” and noted Tmf, which reflects the state of the cooling system. Indeed, the Tmf is an observation of the hidden state of the system. This one is modeled by a Markov chain and the Tmf is a noisy function of it. Thanks to filtering equations, we obtain results on the probability that an appliance is in degraded state at time t, knowing the history of the Tmf until this moment. We have evaluated the numerical behavior of our approach on simulated data. Then we have applied this methodology on our real data and we have checked that the results are consistent with the reality. This method can be implemented in a HUMS (Health and Usage Monitoring System). This simple example of HUMS would allow the Thales Optronics Company to improve its maintenance system. This company will be able to recall appliances which are estimated to be in degraded state and do not control too early those estimated in stable state.

This work is to appear in [18].

6.22. A new sliced inverse regression method for multivariate response

Participants: Jérôme Saracco, Raphaël Coudret.

This work is in collaboration with Stéphane Girard (Inria Grenoble Alpes).

We consider a semiparametric regression model of a q-dimensional multivariate response y on a p-dimensional covariate x. In this paper, a new approach is proposed based on sliced inverse regression for estimating the effective dimension reduction (EDR) space without requiring a prespecified parametric model. The convergence at rate $\sqrt{n}$ of the estimated EDR space is shown. We discuss the choice of the dimension of the EDR space. The numerical performance of the proposed multivariate SIR method is illustrated on a simulation study. Moreover, we provide a way to cluster components of y related to the same EDR space. One can thus apply properly multivariate SIR on each cluster instead of blindly applying multivariate SIR on all components of y. An application to hyperspectral data is provided.

This work is currently under revision, see [48].

6.23. An introduction to dimension reduction in nonparametric kernel regression

Participant: Jérôme Saracco.

This work is in collaboration with Stéphane Girard (Inria Grenoble Alpes).

Nonparametric regression is a powerful tool to estimate nonlinear relations between some predictors and a response variable. However, when the number of predictors is high, nonparametric estimators may suffer from the curse of dimensionality. In this chapter, we show how a dimension reduction method (namely Sliced Inverse Regression) can be combined with nonparametric kernel regression to overcome this drawback. The methods are illustrated both on simulated datasets as well as on an astronomy dataset using the **R** software.

This work was presented in ”School in Astrostatistics” (Annecy, October, 21-25, 2013) and is to appear as a chapter in book intilted *Methods and Applications of Regression in Astrophysics* in 2014.

CRYPT Team (section vide)

DAHU Project-Team

5. New Results

5.1. Specification and Verification of Database Driven Systems

Participants: Serge Abiteboul, Luc Segoufin, Victor Vianu.

We continued our investication on the verification of database driven systems using an automata model with registers. We have exhibited new classes of decidable scenarios using nominal set theory [25]. These new classes contain the previously known relational cases but also the some semistructered ones.

We introduce in [24] and study a model of collaborative data-driven workflows. In a local-as-view style, each peer has a partial view of a global instance that remains purely virtual. Local updates have side effects on other peers' data, defined via the global instance. We also assume that the peers provide (an abstraction of) their specifications, so that each peer can actually see and reason on the specification of the entire system. We study the ability of a peer to carry out runtime reasoning about the global run of the system, and in particular about actions of other peers, based on its own local observations. A main contribution is to show that, under a reasonable restriction (namely, key-visibility), one can construct a finite symbolic representation of the infinite set of global runs consistent with given local observations. Using the symbolic representation, we show that we can evaluate in pspace a large class of properties over global runs, expressed in an extension of first-order logic with past linear-time temporal operators, PLTL-FO. We also provide a variant of the algorithm allowing to incrementally monitor a statically defined property, and then develop an extension allowing to monitor an infinite class of properties sharing the same temporal structure, defined dynamically as the run unfolds. Finally, we consider an extension of the language, augmeting work-flow control with PLTL-FO formulas. We prove that this does not increase the power of the workflow specification language, thereby showing that the language is closed under such introspective reasoning.

5.2. Distributed data management

Participants: Serge Abiteboul, Émilien Antoine, Cristina Sirangelo.

We have studied the feasibility of query answering in the presence of incomplete information in data. In particular we have investigated when it is the case that classical query evaluation techniques, which are commonly used over complete data, suffice to answer queries also in the presence of incompleteness [26]. These results allowed to find syntactic classes of queries that can be answered efficiently under many well known semantics of incompleteness, using query answering techniques which are already implemented (and optimized) in classical database systems.

The management of Web users' personal information is increasingly distributed across a broad array of applications and systems, including online social networks and cloud-based services. While users wish to share and integrate data using these systems, it is increasingly difficult to avoid the risks of unintended disclosures or unauthorized access by applications.

In [21], [20], we propose a novel access control model that operates within a distributed data management framework based on datalog. Using this model, users can control access to data they own and control applications they run. They can conveniently specify access control policies providing flexible tuple-level control derived using provenance information. We present a formal specification of the model, a theoretical analysis, and an implementation. We show that the computational cost of access control is acceptable.

5.3. Query Processing for the Web

Participants: Johann Brault-Baron, Arnaud Durand, Nadime Francis, Wojciech Kazana, Luc Segoufin, Cristina Sirangelo.

In many applications the output of a query may have a huge size and enumerating all the answers may already consume too many of the allowed resources. In this case it may be appropriate to first output a small subset of the answers and then, on demand, output a subsequent small numbers of answers and so on until all possible answers have been exhausted. To make this even more attractive it is preferable to be able to minimize the time necessary to output the first answers and, from a given set of answers, also minimize the time necessary to output the next set of answers - this second time interval is known as the *delay*. We have shown that this was doable with a linear preprocessing time and constant enumeration delay for first-order queries over structures of bounded expansion [27] and for monadic second-order queries over structures of bounded tree-width [15]. We also presented a survey about this work at the Intl. Conf. on Database Theory (ICDT) [19].

Web data is often structured in the XML format. In [18] we have surveyed results about static analysis of pattern-based queries over XML documents. These queries are analogs of conjunctive queries, their unions and Boolean combinations, in which tree patterns play the role of atomic formulae. These can be viewed as both queries and incomplete documents, and thus static analysis problems can also be viewed as answering queries over such documents. We looked at satisfiability of patterns under schemas, containment of queries for various features of XML used in queries, query answering, and applications of pattern-based queries in reasoning about schema mappings for data exchange.

DANTE Team

6. New Results

6.1. Probabilistic resource management

Participants: Paulo Gonçalves [correspondant], Thomas Begin, Shubhabrata Roy, Thibaud Trolliet.

This contribution is part of the PhD work of S. Roy (Dec. 2010 – March 2014) on probabilistic resource management in the context of highly volatile workloads. We proposed a Markovian model that can reproduce the workload volatility occurring in real-life VoD systems, such as Video On Demand (VoD). We derived an original MCMC based identification procedure to calibrate model on real data. We assess the accuracy of the proposed procedure in terms of bias and variance through several numerical experiments, and we compared its outcome with a former ad-hoc method that we had designed. We also compared the performance of our approach to that of other existing models examining the goodness-of-fit of the steady state distribution and of the autocorrelation function of real workload traces. Results show that the combination of out model and its MCMC based calibration clearly outperforms the existing state-of-the art. (See [17], [18])

peta

6.2. Semi-supervised machine learning

Participant: Paulo Gonçalves [correspondant].

This contribution is part of the PhD work of M. Sokol (EPI MAESTRO, Oct. 2009 – May 2014), co-supervised with K. Avrachenkov and Ph. Nain, on the classification of content and users in peer-to-peer networks using graph-based semi-supervised learning methods. Semi-supervised learning methods constitute a category of machine learning methods which use labelled points together with unlabelled data to tune the classifier. The main idea of the semi-supervised methods is based on an assumption that the classification function should change smoothly over a similarity graph, which represents relations among data points. This idea can be expressed using kernels on graphs such as graph Laplacian. Different semi-supervised learning methods have different kernels which reflect how the underlying similarity graph influences the classification results. In a recent work, we analysed a general family of semi-supervised methods, provided insights about the differences among the methods and gave recommendations for the choice of the kernel parameters and labelled points. In particular, it appeared that it was preferable to choose a kernel based on the properties of the labelled points. We illustrated our general theoretical conclusions with an analytically tractable characteristic example, clustered preferential attachment model and classification of content in P2P networks. (See [8])

6.3. Analysis of heart beat rate variability

Participant: Paulo Gonçalves [correspondant].

Intrapartum fetal heart rate monitoring constitutes an important stake aiming at early acidosis detection. Measuring heart rate variability is often considered a powerful tool to assess the intrapartum health status of fetus and has has been envisaged using various techniques. In the present contribution, the power of scale invariance parameters, such as the Hurst exponent and the global regular- ity exponent, estimated from wavelet coefficients of intrapartum fetal heart rate time series, to evaluate the health status of fetuses is quantified from a case study database, constituted at a French Academic Hospital in Lyon. Notably, the ability of such parameters to discriminate subjects incorrectly classified according to FIGO rules as abnormal will be discussed. Also, the impact of the occurrence of decelerations identified as complicated by obstetricians on the values taken by Hurst parameter is investigated in detail. (See [7])

6.4. Hierarchical Modeling of IEEE 802.11 Multi-hop Wireless Networks

Participants: Thiago Wanderley Matos de Abreu, Thomas Begin, Isabelle Guérin Lassous.

IEEE 802.11 is implemented in many wireless networks, including multi-hop networks where communications between nodes are conveyed along a chain. We present a modelling framework to evaluate the performance of flows conveyed through such a chain. Our framework is based on a hierarchical modelling composed of two levels. The lower level is dedicated to the modelling of each node, while the upper level matches the actual topology of the chain. Our approach can handle different topologies, takes into account Bit Error Rate and can be applied to multi-hop flows with rates ranging from light to heavy workloads. We assess the ability of our model to evaluate loss rate, throughput, and end-to-end delay experienced by flows on a simple scenario, where the number of nodes is limited to three. Numerical results show that our model accurately approximates the performance of flows with a relative error typically less than 10%. (See [6])

6.5. Available Bandwidth Estimation in GPSR for VANETs

Participant: Isabelle Guérin Lassous.

We propose an adaptation of the collision probability used in the measure of the available bandwidth designed for Mobile Ad hoc Networks (MANETs) and which is used in ABE. Instead, we propose a new ABE+ that includes a new function to estimate the probability of losses. This new function has been specially designed for Vehicular Ad hoc Networks, to be suited to the high mobility and variable density in vehicular environments. In this new solution, we do not only consider the packet size, but also other metrics, such as, density and speed of the nodes. We include the ABE+ algorithm in the forwarding decisions of the GBSR-B protocol, which is an improvement of the well-known GPSR protocol. Finally through simulations, we compare the performance of our new ABE+ compared to the original ABE. These results show that ABE+ coupled with GBSR-B achieves a good trade-off in terms of packet losses and packet end-to-end delay. (See [19])

6.6. Reduced complexity in $M/Ph/c/N$ queues

Participant: Thomas Begin [correspondant].

This contribution stems from a long-existing collaboration with Pr. Brandwajn (UCSC), which is devoted to innovative numerical solution of classical queueing systems. Many real-life systems can be modelled using the classical $M/G/c/N$ queue. A frequently-used approach is to replace the general service time distribution by a phase-type distribution since the $M/Ph/c/N$ queue can be described by familiar balance equations. The downside of this approach is that the size of the resulting state space suffers from the "dimensionality curse", *i.e.*, exhibits combinatorial growth as the number of servers and/or phases increases. To circumvent this complexity issue, we propose to use a reduced state description in which the state of only one server is represented explicitly, while the other servers are accounted for through their rate of completions. The accuracy of the resulting approximation is generally good and, moreover, tends to improve as the number of servers in the system increases. Its computational complexity in terms of the number of states grows only linearly in the number of servers and phases. (See [9])

6.7. Throughput maximisation in multi-radio wireless networks

Participants: Isabelle Guérin Lassous, Busson Anthony.

Wireless mesh network offers a simple and costless solution to deploy wireless based infrastructure network. They are particularly suitable when the network is deployed temporarily, such as substitution networks (studied in the ANR RESCUE project). In order to ensure an important capacity, the mesh nodes may be equipped with several 802.11 network interfaces. The classical approach to assign 802.11 channels to these interfaces aim to minimise global interference, *i.e.* minimise the conflict graph. Our proposition is two folds. We define a new benefit function that describes the network capacity rather than interference/conflicts. Also, we derive an efficient algorithm that maximises this function. Simulation results show that the proposed function is very close to the measured end-to-end throughputs, empirically proving that it is the good function to optimise. Moreover, the channel assignation algorithm based on this optimisation presents an important throughput increase compared to the classical approaches.

6.8. Aggregation of temporal contact series into graph series

Participants: Christophe Crespelle, Eric Fleury, Yannick Léo.

We consider the problem of aggregating a temporal contact series into a series of graph. This consists in slicing time into time-windows of equal length and forming for each window the graph of the contacts occurred within it. The length chosen for the windows has a great impact on the properties of the graph series obtained. Then the key question that arises is: how one should choose the length of aggregation windows? In the master insternship of Yannick Léo (spring 2013), we designed a method to do so, by using the occupation rate of paths in the graph series. We have applied this method on several real-world data and obtained very good results. The method has also greatly beneficiated of a new notion of shortest dynamic paths that we developed during the master internship of Pierre-Alain Scribot (spring 2013).

6.9. Dynamic Contact Network Analysis in Hospital Wards

Participants: Christophe Crespelle, Eric Fleury, Lucie Martinet.

We analysed a huge and very precise trace of contact data collected during 6 months on the entire population of a rehabilitation hospital. We investigated the graph structure of the average daily contact network, and we unveiled striking properties of this structure in the considered hospital, as a very strong introversion of services, the key role of the contacts between patients and staff in connecting those introverted services all together, and very different pattern of contacts during one day between patients and staffs. The methodology we designed to lead these analysis is very general and can be applied for analysing any dynamic complex network where nodes are classified into groups. Those results are part of Lucie Martinet's PhD thesis.

6.10. A Linear-Time Algorithm for Computing the Prime Decomposition of a Directed Graph with Regard to the Cartesian Product

Participant: Christophe Crespelle.

We design the first linear-time algorithm for computing the prime decomposition of a digraph G with regard to the cartesian product. A remarkable feature of our solution is that it computes the decomposition of G from the decomposition of its underlying undirected graph, for which there exists a linear-time algorithm. First, this allows our algorithm to remain conceptually very simple and in addition, it provides new insight into the connexions between the directed and undirected versions of cartesian product of graphs [11]

6.11. Linear-time Constant-ratio Approximation Algorithm and Tight Bounds for the Contiguity of Co-graphs

Participant: Christophe Crespelle.

We consider a graph parameter called *contiguity* which aims at encoding a graph by a linear ordering of its vertices. The purpose is to obtain very compact encoding of a graph which still answers in optimal time to neighbourhood queries on the graph (*i.e.* list the neighbours of a given vertex). This allows to deal with very large instances of graphs by loading them entirely into the memory, without penalising the running time of algorithms treating those instances. We designed a linear time algorithm for computing a constant-ratio approximation of the contiguity of an arbitrary co-graph. Our algorithm does not only give an approximation of the parameter, but also provides an encoding of the co-graph realising this value [10]

6.12. Model for Time-Varying Graphs.

Participant: Éric Fleury.

We propose a novel model for representing finite discrete Time-Varying Graphs (TVGs). The major application of such a model is for the modelling and representation of dynamic networks. In our proposed model, an edge is able to connect a node u at a given time instant t_a to any other node v (u possibly equal to v) at any other time instant t_b (t_a possibly equal to t_b), leading to the concept that such an edge can be represented by an ordered quadruple of the form (u, t_a, v, t_b). Building upon this basic concept, our proposed model defines a TVG as an object $H = (V, E, T)$, where V is the set of nodes, $E \subseteq V \times T \times V \times T$ is the set of edges, and T is the finite set of time instants on which the TVG is defined. We show how key concepts, such as degree, path, and connectivity, are handled in our model. We also analyse the data structures used for the representation of dynamic networks built following our proposed model and demonstrate that, for most practical cases, the asymptotic memory complexity of our TVG representation model is determined by the cardinality of the set of edges. (See [20])

DEDUCTEAM Exploratory Action

6. New Results

6.1. Dedukti

The version 2.0 of the Dedukti system, developed by Ronan Saillard, has been released in July 2013. It is based on an improved version of the $\lambda\Pi$-calculus modulo where rewrite rules are explicitly added [31], and where the conditions for typing the rewrite rules are weakened.

This version is fully written in OCaml. It is smaller, far more efficient than the previous version, and permits to type-check much bigger files.

New features include a better reporting of errors, an interactive mode, an export functionality from Dedukti to the MMT format [53], and non-linear pattern matching.

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

A new version of Coqine has been developed by Ali Assaf. This version is designed using a Coq plugin architecture, which allows for a smoother integration with Coq's code base and alleviates problems of maintainability that affected the previous version.

The implementation of Holide has been improved, by Ali Assaf. This improved version incorporates sharing at the level of terms and types. This optimization allows to reduce the type-checking time of the OpenTheory standard library from more than 30 minutes to less than 1 minute.

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 called Focalide, through a compilation to Dedukti. This translation can benefit from the flexibility of Dedukti to deal with more dynamic object-oriented languages than FoCaLiZe; they are currently working on generalizing this translation using ζ-calculus as a theoretical foundation for objects.

Resolution and superposition are proof-search methods that are used in state-of-the-art first-order automated theorem provers such as iProver, Vampire, E or SPASS. A shallow embedding of resolution and superposition proofs in the $\lambda\Pi$-calculus modulo has been proposed by Guillaume Burel, thus offering a way to check these proofs in a trusted setting, and to combine them with other proofs. This embedding has been implemented in particular as a backend of iProver Modulo, therefore allowing to check proofs found by iProver Modulo using Dedukti [20].

A shallow embedding in Dedukti of the tableaux proofs generated by Zenon modulo has been designed and implemented by Frédéric Gilbert [22], [23]. The embedding is based on a refined version of previous double-negation translations, introducing as less as possible double negations. This optimization has shown that more than half of the proofs found by Zenon modulo are not using the excluded-middle law, therefore being purely intuitionistic.

6.3. Automated Theorem Proving

Mélanie Jacquel (*Siemens*) and David Delahaye developed *Super Zenon* [5], a generalization of the extension of *Zenon* to superdeduction to handle any first order theory. To do so, they designed heuristics able to automatically transform axioms of a theory into rewrite rules. This new tool has been tested over the first order problems of the TPTP library and a significant increase has been observed. A first distribution of this tool (under GPL licence) is planned in the first months of 2014. In addition, an integration to the *Rodin* platform is also planned with the help of Laurent Voisin (*Systerel*). This integration should allow us to apply this tool in the context of *Event-B*.

Pierre Halmagrand, David Delahaye, Damien Doligez, and Olivier Hermant developed *Zenon Modulo* [22], [23], an extension of *Zenon* to Deduction modulo. Like *Super Zenon*, this new tool is able to deal with any first order theory and relies on an heuristic able to automatically transform axioms of a theory into rewrite rules. This tool has also been tested over the first order problems of the TPTP library and a similar increase of performance (compared to *Super Zenon*) has been observed. Frédéric Gilbert has developed a *Dedukti* backend for this extension, which is based on a double-negation transformation that allows us to transform classical proofs produced by *Zenon Modulo* into intuitionistic proofs in *Dedukti*. This tool is intended to be applied in the framework of the *BWare* project in order to automatically verify proof obligations coming from the modeling of industrial applications. To do so, the idea is to manually transform the *B* set theory into a theory modulo and provide it to *Zenon Modulo* in order to verify the proof obligations of the *BWare* project.

Guillaume Burel and Simon Cruanes have designed a method to scan sets of first-order clauses in order to detect the presence of instances of axiomatic theories (group structures, total orderings, etc.), even during a saturation process (so that theories that only become apparent during the proof search can be detected) [21]. To this end, they introduced the concept of *meta-prover*, a Datalog system that reasons over properties of the problem, and communicates with the saturation prover. This technique made some applications possible, such as the use of generic lemma and an equational redundancy criterion for some theories, and was implemented in Zipperposition.

Simon Cruanes has been working on superposition modulo linear arithmetic, using Zipperposition as a test bed. The focus is on problems with rational or integer arithmetic mixed with first-order reasoning, an area in which SMT solvers struggle. The work is still preliminary, but shows promising results.

Depending on the logic for finite structures, which is defined by Gilles Dowek and Ying Jiang (Beijing), Kailiang Ji has extended the use of proof search algorithms in Deduction modulo to automatically prove some graph properties, such as (un)reachability, which can be described by CTL formulas. A technical report about this has been given on Locali 2013 in Beijing.

Together with Tayssir Touili (University Paris Diderot) Hugo Macedo has shown how to advance the performance of the application of model checking techniques in the domain of malicious software detection. The work consisted in leveraging the reachability analysis used in the model checking of pushdown systems to infer malicious behavior patterns from known malware. From such new application a malware detection tool was prototyped and put to the test with instances of “in the wild” (real world) malicious software. This work was published in a large security venue and the details about the technique follow in [29].

Kim-Quyen Ly extended her formally-proved (in Coq) automated termination-certificate (for first-order rewrite systems) verifier Rainbow for dealing with certificates using arguments filtering [39] and other termination techniques.

6.4. Proof theory

The conservativity of the embedding of pure type systems in the $\lambda\Pi$-calculus modulo was proved by Ali Assaf. This result extends those of Cousineau and Dowek [46] and further justifies the use of the $\lambda\Pi$-calculus modulo as a logical framework. This embedding is the basis for the automated translation tools Holide and Coqine.

Frédéric Blanqui, Jean-Pierre Jouannaud (Univ. Paris 11) and Albert Rubio (Technical University of Catalonia) have developed a method aiming at carrying out termination proofs for higher-order calculi. CPO appears to be the ultimate improvement of the higher-order recursive path ordering (HORPO) [45] in the sense that this definition captures the essence of computability arguments *à la* Tait and Girard, therefore explaining the name of the improved ordering. It has been shown that CPO allows to consider higher-order rewrite rules in a simple type discipline with inductive types, that most of the guards present in the recursive calls of its core definition cannot be relaxed in any natural way without losing well-foundedness, and that the precedence on function symbols cannot be made more liberal anymore.

Frédéric Blanqui worked on the formalization in the Coq proof assistant of various definitions of the notion of α-equivalence on pure λ-terms. In particular, he formalized and formally proved equivalent the definitions of Church (1932), Curry and Feys (1958), Krivine (1993), and Gabbay and Pitts (1999). This work is freely available from the CoLoR library released on December 13th.

Alejandro Díaz-Caro and Gilles Dowek have introduced an extension of λ-calculus with pairs where isomorphic types are equated. Identifying some types requires to also identify some terms via an equivalence relation on terms, leading to an interesting calculus, which is related to several known non-deterministic and probabilistic calculi. A preliminary version of this work has been published on [24]. A complete version in simple types, with its proof of normalisation, is currently under review.

Together with Ying Jiang, Gilles Dowek has started to investigate the links between model-checking and proof-checking. This has materialized by an encoding of CTL for a finite model in predicate logic and by the definition of a proof-system for CTL.

Olivier Hermant has studied optimized versions of double-negation translations, that allow to switch between classical and intuitionistic logics. Such an algorithm has been implemented in Zenon's backend to Dedukti by Frédéric Gilbert. Gilles Dowek has given new version of Gödel's translation of classical logic into constructive logic. This translation is homomorphic, hence it can be seen as a mere definition of the classical connectives from the constructive ones.

6.5. Safety of aerospace systems

Pierre Néron has designed a method to transform straight line programs, such as those used in some aerospace systems into others that do not use some operations such as, square roots and divisions that cannot be performed exactly on decimal numbers. To this end he has defined a new notion of anti-unification, called *constrained anti-unification*, and a new anti-unification algorithm.

6.6. Models of Computation

Alejandro Díaz-Caro and Gilles Dowek have shown how to provide a structure of probability space to the set of execution traces on a non-confluent abstract rewrite system, by defining a variant of a Lebesgue measure on the space of traces. Then, they showed how to use this probability space to transform a non-deterministic calculus into a probabilistic one. As an example, they applied this technique to the previously introduced non-deterministic calculus. [25]

Ali Assaf and Alejandro Díaz-Caro, together with Simon Perdrix (Nancy), Christine Tasson (PPS) and Benoît Valiron (PPS) have determined the relationship between the algebraic λ-calculus, a fragment of the differential λ-calculus and the linear-algebraic λ-calculus, a candidate λ-calculus for quantum computation. Both calculi are algebraic: each one is equipped with an additive and a scalar-multiplicative structure, and their set of terms is closed under linear combinations. However, the two languages were built using different approaches: the former is a call-by-name language whereas the latter is call-by-value; the former considers algebraic equalities whereas the latter approaches them through rewrite rules. They have analysed how these different approaches relate to one another, proposing four canonical languages based on each of the possible choices: call-by-name versus call-by-value, algebraic equality versus algebraic rewriting. They have shown that the various languages simulate one another. Preliminary versions of this work where published in [47] and [41]. Now they are working on a journal version filling the gaps between these previous works.

Together with Pablo Arrighi (Grenoble) and Benoît Valiron (PPS), Alejandro Díaz-Caro has described a type system for the linear-algebraic lambda-calculus. The type system accounts for the linear-algebraic aspects of this extension of lambda-calculus: It is able to statically describe the linear combinations of terms that will be obtained when reducing the programs. This gives rise to an original type theory where types, in the same way as terms, can be superposed into linear combinations. They have proven that the resulting typed lambda-calculus is strongly normalising and features a weak subject reduction. In addition, they have shown how to naturally encode matrices and vectors in this typed calculus [34].

Gilles Dowek has investigated a new definition of the notion of a chaotic system that can be applied to discrete systems and that is compatible with the principle of a finite density of information.

The paper Call-by-value non-determinism in a linear logic type discipline by Alejandro Díaz-Caro, Giulio Manzonetto and Michele Pagani has been published [26].

The paper Universality in two dimensions of Gilles Dowek and Nachum Dershowitz has been published.

The paper Linear-algebraic lambda-calculus: higher-order, encodings and confluence of Pablo Arrighi and Gilles Dowek has been published.

The book *Lambda Calculus with Types* , written by Henk Barendregt, Wil Dekkers, Richard Statman, and 11 contributors, including Gilles Dowek, has been published.

6.7. Constraint solving

Catherine Dubois has extended the formally verified constraint solver (on finite domains) she has developed with Matthieu Carlier and Arnaud Gotlieb with a new local consistency property (bound-consistency).

DEFI Project-Team

6. New Results

6.1. Qualitative methods for inverse scattering problems

6.1.1. A generalized formulation of the Linear Sampling Method

Participants: Lorenzo Audibert, Houssem Haddar.

We proposed and analyzed a new formulation of the Linear Sampling Method that uses an exact characterization of the targets shape in terms of the so-called farfield operator (at a fixed frequency). This characterization is based on constructing nearby solutions of the farfield equation using minimizing sequences of a least squares cost functional with an appropriate penalty term. We first provided a general framework for the theoretical foundation of the method in the case of noise-free and noisy measurements operator. We then explicited applications for the case of inhomogeneous inclusions and indicate possible straightforward generalizations. We finally validated the method through some numerical tests and compare the performances with classical LSM and the factorization methods.

6.1.2. Inverse problems for periodic penetrable media

Participant: Dinh Liem Nguyen.

Imaging periodic penetrable scattering objects is of interest for non-destructive testing of photonic devices. The problem is motivated by the decreasing size of periodic structures in photonic devices, together with an increasing demand in fast non-destructive testing. In this project, we considered the problem of imaging a periodic penetrable structure from measurements of scattered electromagnetic waves. As a continuation of earlier work jointly with A. Lechleiter we considered an electromagnetic problem for transverse magnetic waves (previous work treats transverse electric fields), and also the full Maxwell equations. In both cases, we treat the direct problem by a volumetric integral equation approach and construct a Factorization method.

6.1.3. Transmission Eigenvalues and their application to the identification problem

Participant: Houssem Haddar.

The so-called interior transmission problem plays an important role in the study of inverse scattering problems from (anisotropic) inhomogeneities. Solutions to this problem associated with singular sources can be used for instance to establish uniqueness for the imaging of anisotropic inclusions from muti-static data at a fixed frequency. It is also well known that the injectivity of the far field operator used in sampling methods is related to the uniqueness of solutions to this problem. The frequencies for which this uniqueness fails are called transmission eigenvalues. We are currently developing approaches where these frequencies can be used in identifying (qualitative informations on) the medium properties. Our research on this topic is mainly done in the framework of the associate team ISIP http://www.cmap.polytechnique.fr/~defi/ISIP/isip.html with the University of Delaware. A review article on the state of art concerning the transmission eigenvalue problem has been written in collaboration with F. Cakoni. We also edited a spacial issue of the journal Inverse Problems dedicated to the use of these transmission eigenvalues in inverse problems http://iopscience.iop.org/0266-5611/29/10/100201/. Our recent contributions are the following:

- Together with A. Cossonnière we analyzed the Fredholm properties of the interior transmission problem for the cases where the index contrast changes sign outside the boundary by using a surface integral equation approach.
- With F. Cakoni and N. Chaulet we investigated the asymptotic behaviour of the first transmission eigenvalue of a thin coating with respect to the coating thickness.

6.1.4. The factorization method for inverse scattering problems

6.1.4.1. The factorization method for cracks with impedance boundary conditions

Participant: Houssem Haddar.

With Y. Boukari we used the Factorization method to retrieve the shape of cracks with impedance boundary conditions from farfields associated with incident plane waves at a fixed fre- quency. This work is an extension of the study initiated by Kirsch and Ritter [Inverse Problems, 16, pp. 89-105, 2000] where the case of sound soft cracks is considered. We address here the scalar problem and provide theoretical validation of the method when the impedance boundary conditions hold on both sides of the crack. We then deduce an inversion algorithm and present some validating numerical results in the case of simply and multiply connected cracks [5].

6.1.4.2. The factorization method for EIT with uncertain background

Participants: Giovanni Migliorati, Houssem Haddar.

We extended the Factorization Method for Electrical Impedance Tomography to the case of background featuring uncertainty. This work is based on our earlier algorithm for known but inhomogeneous backgrounds. We developed three methodologies to apply the Factorization Method to the more difficult case of piece-wise constant but uncertain background. The first one is based on a recovery of the background through an optimization scheme and is well adapted to relatively low dimensional random variables describing the background. The second one is based on a weighted combination of the indicator functions provided by the Factorization Method for different realiza- tions of the random variables describing the uncertain background. We show through numerical experiments that this procedure is well suited to the case where many realizations of the measurement operators are available. The third strategy is a variant of the previous one when measurements for the inclusion-free background are available. In that case, a single pair of measurements is sufficient to achieve comparable accuracy to the deterministic case [15].

6.1.4.3. The factorization method for GIBC

Participants: Mathieu Chamaillard, Houssem Haddar.

With N. Chaulet, we studied the identification of some obstacle and some Generalized Impedance Boundary Conditions (GIBC) on the boundary of such obstacle from far field measurements generated by the scattering of harmonic incident waves. The GIBCs are approximate models for thin coatings, corrugated surfaces, rough surfaces or imperfectly conducting media.

We justified the use of the Factorization method to solve the inverse obstacle problem in the presence of GIBCs. This method gives a uniqueness proof as well as a fast algorithm to reconstruct the obstacle from the knowledge of the far field produced by incident plane waves for all the directions of incidence at a given frequency. We also provided some numerical reconstructions of obstacles for several impedance operators.

6.2. Iterative Methods for Non-linear Inverse Problems

6.2.1. Inverse medium problem for axisymmetric eddy current models

Participants: Houssem Haddar, Zixian Jiang, Kamel Riahi.

We continued our developments of shape optimization methods for inclusion detection in an axisymmetric eddy current model. This problem is motivated by non-destructive testing methodologies for steam generators. We finalized a joint work with A. Lechleiter on numerical methods for the solution of the direct problem in weighted Sobolev spaces using approriate Dirichlet-to-Neumann mappings to bound the computational domain. We are also finalized jointly with M. El Guedri the work on inverse solver using a regularized steepest descent method for the problem of identifying a magnetite deposits using axial eddy current probe. We then addressed two new issues:

- We developed asymptotic models to identify thin highly conducting deposits. We derived three possible asymptotic models that can be exploited in the inverse problem. The numerical validation is under study.

- We extended the inverse scheeme to 3D configurations with axisymmetry at infinity: this includes exact charactherization of the shape derivative for a mixed formulation of eddy current problems and a parametric inversion scheme based on a pre-defined discrete grid for deposit location.

6.2.2. The conformal mapping method and inverse scattering at low frequencies

Participant: Houssem Haddar.

Together with R. Kress we employed a conformal mapping technique for the inverse problem to reconstruct a perfectly conducting inclusion in a homogeneous background medium from Cauchy data for electrostatic imaging, that is, for solving an inverse boundary value problem for the Laplace equation. In a recent work [13] we proposed an extension of this approach to inverse obstacle scattering for time-harmonic waves, that is, to the solution of an inverse boundary value problem for the Helmholtz equation. The main idea is to use the conformal mapping algorithm in an iterative procedure to obtain Cauchy data for a Laplace problem from the given Cauchy data for the Helmholtz problem. We presented the foundations of the method together with a convergence result and exhibit the feasibility of the method via numerical examples.

6.2.3. A steepest descent method for inverse electromagnetic scattering problems

Participant: Houssem Haddar.

Together with N. Chaulet, we proposed the application of a non linear optimization techniques to solve the inverse scattering problems for the $3D$ Maxwell's equations with generalized impedance boundary conditions. We characterized the shape derivative in the case where the GIBC is defined by a second order surface operator. We then applied a boundary variation method based on a regularized steepest descent to solve the 3-D inverse problem with partial farfield data. The obtained numerical results demonstrated the possibility of identifying the shape of coated objects as well as the parameters of the coating in the $3D$ Maxwell case.

6.3. Shape and topology optimization

6.3.1. Geometric shape optimization

Participant: Grégoire Allaire.

With Ch. Dapogny and P. Frey, we propose a new approach for geometry and topology optimization of structures which benefits from an accurate description of shapes at each stage of the iterative process (by means of a mesh amenable for mechanical analyses) while retaining the whole versatility of the level set method when it comes to accounting for their evolution. The key ingredients of this method are two operators for switching from a meshed representation of a domain to an implicit one, and conversely; this notably brings into play an algorithm for generating the signed distance function to an arbitrary discrete domain, and a mesh generation algorithm for implicitly-defined geometries.

6.3.2. Worst-case design shape optimization

Participant: Grégoire Allaire.

with Ch. Dapogny, we propose a deterministic method for optimizing a structure with respect to its worst possible behavior when a "small" uncertainty exists over some of its features. The main idea is to linearize the considered cost function with respect to the uncertain parameters, then to consider the supremum function of the obtained linear approximation, which can be rewritten as a more classical function of the design, owing to standard adjoint techniques from optimal control theory. The resulting linearized worst-case objective function turns out to be the sum of the initial cost function and of a norm of an adjoint state function, which is dual with respect to the considered norm over perturbations.

6.3.3. Multi-phase structural optimization

Participant: Grégoire Allaire.

With Ch. Dapogny, G. Delgado and G. Michailidis, we consider the optimal distribution of several elastic materials in a fixed working domain. In order to optimize both the geometry and topology of the mixture we rely on the level set method for the description of the interfaces between the different phases. We discuss various approaches, based on Hadamard method of boundary variations, for computing shape derivatives which are the key ingredients for a steepest descent algorithm. The shape gradient obtained for a sharp interface involves jump of discontinuous quantities at the interface which are difficult to numerically evaluate. Therefore we suggest an alternative smoothed interface approach which yields more convenient shape derivatives. We rely on the signed distance function and we enforce a fixed width of the transition layer around the interface (a crucial property in order to avoid increasing "grey" regions of fictitious materials). It turns out that the optimization of a diffuse interface has its own interest in material science, for example to optimize functionally graded materials.

6.3.4. Level-Set Method

Participant: Olivier Pantz.

We have begin to work, with Gabriel Delagado, on a new level-set optimization method, based on a gradient method. The key idea consists in computing directly the derivative of the discretized cost functions. The main advantage is that it is usually more simple to implement than the standard approach (consisting in using a discretized version of the gradient of the cost function). Moreover, the results obtained are as good or even better than the one obtained in previous works. Nevertheless, this method has its drawbacks, since the cost function is only derivable almost everywhere (the zero level-set has to be transverse to the triangulation of the mesh). It follows that convergence toward the minimum by the gradient method is not granted. To overcome this problem, we intend to use a mix-formulation for the state function. Unfortunately, such a formulation, in the case of linear elasticity is quite difficult to obtain. We thus intend to begin with the simplest scalar case, for which a lot more hybrid formulations are available.

6.3.5. Optimization of a sodium fast reactor core

Participants: Grégoire Allaire, Olivier Pantz.

In collaboration with D. Schmidt, G. Allaire and E. Dombre, we apply the geometrical shape optimization method for the design of a SFR (Sodium Fast reactor) core in order to minimize a thermal counter-reaction known as the sodium void effect. In this kind of reactor, by increasing the temperature, the core may become liable to a strong increase of reactivity ρ, a key-parameter governing the chain-reaction at quasi-static states. We first use the 1 group energy diffusion model and give the generalization to the 2 groups energy equation. We then give some numerical results in the case of the 1 group energy equation. Note that the application of our method leads to some designs whose interfaces can be parametrized by very smooth curves which can stand very far from realistic designs. We don't explain here the method that it would be possible to use for recovering an operational design but there exists several penalization methods that could be employed to this end. This work was partially sponsored by EDF. Our results will be published in the proceedings of the CEMRACS'11, during which part of the results have been obtained.

6.4. Asymptotic Analysis

6.4.1. Effective boundary conditions for thin periodic coatings

Participants: Mathieu Chamaillard, Houssem Haddar.

This topic is the object of a collaboration with Patrick Joly and is a continuation of our earlier work on interface conditions done in the framework of the PhD thesis of Berangère Delourme. Th goal here is to derive effective conditions that model scattering from thin periodic coatings where the thickness and the periodicity are of the same length but very small compared to the wavelength. The originality of our work, compared to abundant literature is to consider the case of arbitrary geometry (2-D or 3-D) and to consider higher order approximate models. We formally derived third order effective conditions after exhibiting the full asymptotic expansion of the solution in terms of the periodicity length.

6.4.2. Homogenization of electrokinetic models in porous media

Participant: Grégoire Allaire.

With R. Brizzi, J.-F. Dufrêche, A. Mikelic and A. Piatnitski, we are interested in the homogenization (or upscaling) of a system of partial differential equations describing the non-ideal transport of a N-component electrolyte in a dilute Newtonian solvent through a rigid porous medium. Our work can be divided in two different contributions. First, in the case of an ideal model (for which the homogenized system was already known) we consider the various limits which can be obtained in the effective parameters when the ratio between the characteristic pore length and the Debye length is either small or large. Second, we studied the homogenization process in the non-ideal case, namely when considering the so-called mean spherical approximation (MSA) model which takes into account finite size ions and screening effects.

6.4.3. A new shell modeling modeling

Participant: Olivier Pantz.

Using a formal asymptotic expansion, we have proved with K. Trabelsi, that non-isotropic thin-structure could behave (when the thickness is small) like a shell combining both membrane and bending effects. It is the first time to our knowledge that such a model is derived. An article on this is currently under review.

6.4.4. A new Liouville type Rigidity Theorem

Participant: Olivier Pantz.

We have recently developed a new Liouville type Rigidity Theorem. Considering a cylindrical shaped solid, we prove that if the local area of the cross sections is preserved together with the length of the fibers, then the deformation is a combination of a planar deformation and a rigid motion. The results currently obtained are limited to regular deformations and we are currently working with B. Merlet to extend them. Nevertheless, we mainly focus on the case where the conditions imposed to the local area of the cross sections and the length of the fibers are only "almost" fulfilled. This will enable us to derive rigorously new non linear shell models combining both membrane and flexural effects that we have obtained using a formal approach.

6.5. Diffusion MRI

Participants: Jing-Rebecca Li, Houssem Haddar, Dang Van Nguyen, Hang Tuan Nguyen.

Diffusion Magnetic Resonance Imaging (DMRI) is a promising tool to obtain useful information on microscopic structure and has been extensively applied to biological tissues. In particular, we would like to focus on two applications:

- inferring from DMRI measurements changes in the cellular volume fraction occurring upon various physiological or pathological conditions.

  This application is one of the first to show the promise of DMRI because it can detect acute cerebral ischemia (cell swelling) on the basis of lower than normal apparent diffusion coefficient a few minutes after stroke.
- estimating the average cell size in the case of tumor imaging

  This application is useful as a diagnostic tool as well as a tool for the evaluation of tumor treatments.

For both of the above applications we approach the problem via the following steps:

- Construct reduced models of the multiple-compartment Bloch-Torrey partial differential equation (PDE) using homogenization methods.
- Invert the resulting reduced models for the biological parameters of interest: the cellular volume fraction in the first case, and the average distance between neighboring cells in the second case.

../../../../projets/defi/IMG/cylindersandspheres.jpg

Figure 1. Computational domain for simulating diffusion in cerebral gray matter.

../../../../projets/defi/IMG/spheresqueezecells3d.jpg

Figure 2. Computational domain for simulating tumor cells.

We obtained the following results.

- We generated fairly complicated meshes that can be used to simulate diffusion in cerebral gray matter. In the Finite Elements code, this required using the mesh generation software Salome, developed at the CEA Saclay. We are working on the problem of increasing the cellular volume fraction to a physically realistic level, which is difficult for the mesh generator because of the very small distances between the neurons. An article describing the Finite Elements code has been accepted by the Journal of Computational Physics, to be published in 2014. An article on a version of the code using Finite Volume discretization has been accepted by Physics in Medicine and Biology, published in 2013.
- We developed a reduce model of the DMRI signal using homogenization methods. Two articles on this topic have been submitted.

DEMAR Project-Team

6. New Results

6.1. Modelling and Identification

6.1.1. Emergence of Motor Synergy in Reaching Task via Tacit Learning -computational motor control

Participants: Mitsuhiro Hayashibe, Shingo Shimoda [RIKEN,Japan].

The dynamics of multijoint limbs often causes complex dynamic interaction torques which are the inertial effect of other joints motion. It is known that Cerebellum takes important role in a motor learning by developing the internal model. We propose a novel computational control paradigm in vertical reaching task which involves the management of interaction torques and gravitational effect. The obtained results demonstrate that the proposed method is valid for acquiring motor synergy in the system with actuation redundancy and resulted in the energy efficient solutions. It is highlighted that the tacit learning in vertical reaching task can bring computational adaptability and optimality with model-free and cost-function-free approach differently from previous studies.

6.1.2. Anatomy Transfer

Participants: Dicko Ali-Hamadi, Tiantian Liu, Benjamin Gilles, Ladislav Kavan, Francois Faure, Olivier Palombi, Marie-Paule Cani.

Characters with precise internal anatomy are important in film and visual effects, as well as in medical applications. We propose the first semi-automatic method for creating anatomical structures, such as bones, muscles, viscera and fat tissues. This is done by transferring a reference anatomical model from an input template to an arbitrary target character, only defined by its boundary representation (skin). The fat distribution of the target character needs to be specified. We can either infer this information from MRI data, or allow the users to express their creative intent through a new editing tool. The rest of our method runs automatically: it first transfers the bones to the target character, while maintaining their structure as much as possible. The bone layer, along with the target skin eroded using the fat thickness information, are then used to define a volume where we map the internal anatomy of the source model using harmonic (Laplacian) deformation. This way, we are able to quickly generate anatomical models for a large range of target characters, while maintaining anatomical constraints.

6.1.3. Center of Mass Estimation in Multicontact Situations: Simulation

Participants: Alejandro González, Mitsuhiro Hayashibe, Emel Demircan [Stanford Univ.], Philippe Fraisse.

Center of mass (CoM) estimation can be used to evaluate human stability during rehabilitation. A personalized estimation can be obtained using the serial equivalent static chain (SESC) method, calibrated using a series of static postures. The estimation accuracy is dependent on the number and quality of poses used during calibration. Currently, this limits the method's application to unimpaired individuals. We present a preliminary study of a SESC identified in a multi-contact scenario during a Sit-to-Stand task. Stanford's SAI (Simulation and Active Interface) platform was used to emulate human motion and predict relevant reaction forces. The CoM estimation obtained is valid for motions similar to those used during identification. Since the SAI's human model is fully defined, in terms of mass and limb lenths, its exact center of mass is known. Using a 3-dimensional model, the estimated mean error was less than 26 mm for a Sit-to-Stand task involving displacements along all axes. As such, personalized CoM estimation can be available for patients with a limited range of whole body motion.

../../../../projets/demar/IMG/VerticalReachingInria.png

Figure 1. (left)Schematic representation of vertical reaching task. (right)End point transition. (a) only with PD feedback control (b) with tacit learning in addition to the PD control.

../../../../projets/demar/IMG/anatomytransferTeaser.png

Figure 2. A reference anatomy (left) is automatically transferred to arbitrary humanoid characters. This is achieved by combining interpolated skin correspondences with anatomical rules.

../../../../projets/demar/IMG/StanBot.png

Figure 3. After identifying the SESC parameter of the humanoid (Stanbot), it is possible to estimate the position of its CoM. SESC identification was performed with a reduced number of postures, to mimic a patient in need of additional support to maintain a standing pose.

Fig. 3 shows the identification results [27]. Even with a limited motion, it was possible to estimate the position of the simulated robot's CoM projected to onto the ground. This estimation errors are likely due to the lack of an exciting trajectory for identification. Nonetheless, if the postures used during identification describe a patient's range of motion, the CoM estimation can still be valuable.

6.1.4. Interface for identification of the Statically Equivalent Serial Chain's parameters and Balance Assessment

Participants: Alejandro González, Mitsuhiro Hayashibe, Philippe Fraisse.

CoM trajectory can be used to improve the current rehabilitation standards. After an identification phase, a personalized CoM estimate can be obtained using a SESC. Furthermore, using low-cost sensors (Kinect and Wii balance board), make the personalized estimate feasible inside a patient's home. This work focuses on the effect that a visual adaptive interface can have on the SESC identification phase. Specifically on improving its speed and quality. A study conducted on 6 subjects showed a faster convergence and a lower root mean square error (rmse) when the adaptive interface is applied. We find that for the same error (30 mm), the identification with the interface was performed in half the time (86 s) than the one without it (163 s). Similarly, for the same session length (120 s), rmse was of 24.5 mm using the interface and of 34.5 mm without it.

Additionally CoM dynamics may be used to determine stability. Fig. 4 shows an example of this during a squat task. The zero rate of angular momentum (ZRAM) can be used to determine the dynamic stability of a humanoid robot. It can be used to determine the position foot placement to avoid falls.

6.1.5. Forward Estimation of Joint Torque from EMG Signal through Muscle Synergy Combinations

Participants: Zhan Li, Mitsuhiro Hayashibe, David Guiraud.

We investigate the approaches of estimating the ankle joint torque from EMG/activations of associated muscle groups. The approaches discussed fall into two main categories: i) full utilization of both of extension and flexion EMG/activations for estimating the joint torque; ii) exploitation of muscle synergy extraction of EMG/actvations and consequent usage of extracted components in reduced space for estimating the joint torque. Comparison is made between the two methods with experimental data of five able-bodied subjects. From the results we conclude that, method ii) with muscle synergy extraction may not degrade the performance of method i) but meanwhile show the muscle synergic ratios for generating the joint torque, and involvement of joint position and velocity information can improve the estimation for both methods.

6.1.6. Prediction of hand tremor through EMG-based fatigue tracking

Participants: Sourav Chandra, Mitsuhiro Hayashibe, Thondiyath Asokan [IITMadras, India].

Laparoscopic surgical procedure is a very tiring procedure for a surgeon due to the specialized prolonged arm movement with a modular tool. Prolonged activity of arm muscle in such condition induces muscle fatigue which induces hand tremor. Hand tremor not only drastically affects positional accuracy; it also increases the collateral tissue damage. Nullification of this tremor has been the area of active research topic in surgical robotics for last few decades. Though Surface ElectroMyoGram (SEMG) has been used for modeling hand tremor of microsugeons, a single model for predicting amplitude and frequency of such tremor has not been investigated for laparoscopy so far. A model of muscle fatigue induced hand tremor in laparoscopic activity is necessary in order to nullify this hand tremor effect and increase positional accuracy. SEMG is a crucial biopotential in order to get the estimation of the muscle fatigue state. A positive correlation was found among sEMG and hand tremor in frequency domain as shown in Fig. 6 below. In this work, a model based prediction of fatigue induced hand tremor will be investigated with the vicinity of SEMG and other wearable inertial sensor data. The model is intended to have a dynamic structure, which can capture the complexity of the muscle fatigue state to some extent and its effect on the hand tremor amplitude and frequency.

6.1.7. Mobile gait analysis

Participants: Vincent Bonnet [M2H, UM1, Montpellier], Christine Azevedo Coste, Christian Geny [CHU Montpellier], Lionel Lapierre [LIRMM, Montpellier], René Zapata [LIRMM, Montpellier].

../../../../projets/demar/IMG/ZRAM.png

Figure 4. We show the trajectory of the zero rate of angular momentum (ZRAM) for a sqat task. When the ZRAM is found inside the support polygon, the movement can be considered stable. Unstablle movements (C-D) do not determine a fall. The subject may still recover using a balancing strategy; or by taking a step.

Figure 5. Normalized muscle synergy ratios of the five subjects, under isotonic situation with 10Nm plantar load and 5Nm dorsi load

The Video-Kinect-Bot, an affordable mobile platform for pathological gait analysis was developed to assess pathological spatio-temporal parameters. The system, drove by a Kinect sensor, is able to follow a patient at constant distance on his own defined path, and to estimate gait spatio-temporal parameters. Robust Tracking-Learning-Detection algorithm estimates the positions of targets attached to the trunk and heels of the patient. Real-condition experimental validation including corridor, occlusion cases, and illumination change was performed. A gold standard stereophotogrammetric system was also used and shown a good tracking of patient and an accuracy in stride length estimate of 2%.

The Empirical Mode Decomposition (EMD) method was evaluated to estimate the 3D orientation of the lower trunk during walking using the angular velocity signals generated by a wearable inertial measurement unit (IMU) and notably flawed by drift. The IMU was mounted on the lower trunk (L4-L5) with its active axes aligned with the relevant anatomical axes. The proposed method performs an offline analysis but has the advantage of not requiring any parameter tuning. The method was validated in two groups of 15 subjects, one during overground walking, with 180° turnings, and the other during treadmill walking, both for steady-state and transient speeds, using stereophotogrammetric data. Comparative analysis of the results showed that the IMU/EMD method is able to successfully detrend the integrated angular velocities and estimate lateral bending, flexion-extension as well as axial rotations of the lower trunk during walking with RMS errors of 1 deg for straight walking and lower than 2.5 deg for walking with turnings. This work was accepted for publication in Sensors journal for a special issue concerning wearable-sensor for gait analysis in 2014 with the following collaborators V. Bonnet, S. Ramdani, C. Azevedo-Coste, P. Fraisse, C. Mazzà and A. Cappozzo. Data relative to the pitch, roll and yaw angles obtained for one randomly selected treadmill walking trial. The integrated angular velocities (grey line) and the resulting trends (black line) are estimated using EMD (a) during all the trial; zoom over 20 s on the corresponding detrended angles are thereafter estimated (black line) and compared with those obtained using stereophotogrammetry (grey line).

This work is supported by a NOVARTIS funding (see Partnerships and Cooperations section).

../../../../projets/demar/IMG/semg.png

Figure 6. Frequency domain correlation of sEMG and hand tremor during prolonged arm movement

../../../../projets/demar/IMG/VikiBot.png ../../../../projets/demar/IMG/Tracking_snapshot.

*Figure 7. (**Left:**) Detail of the Video-Kinect-Bot system and experimental setup (**Right:**) example of subject tracking in real-environement.*

6.2. Function control and synthesis

6.2.1. Analysis of infection risk in surgery block

Participants: Christine Azevedo Coste, Roger Pissard Gibollet [SED Inria Grenoble Rhône-Alpes], Gabriel Birgand [Bichat Hospital, Paris], Jean-Christophe Lucet [Bichat Hospital, Paris], Gaelle Toupet [Bichat Hospital, Paris].

Despite the increasing implementation of preventive measures, surgical-site infection still induces a substantial burden. Inappropriate staff behaviors can lead to environmental contamination in the operating room and subsequent surgical site infection. The present study focuses on the continued assessment of operating room staff behavior using a motion tracking system, and the evaluation of the impact of this behavior on the surgical-site infection risk during surgical procedures.
A multicenter observational study has been done in 2013, including 10 operating rooms of cardiac and orthopedic surgery in 12 healthcare facilities. A motion tracking system including 8 optical cameras (VICON-Bonita®) recorded movements of reflective markers placed on the surgical caps/hoods of each person entering the room. Different configurations of markers positioning were used to distinguish between staff category. Doors opening were observed as well by means of wireless inertial sensors fixed on the doors and synchronized with the motion tracking system. We have collected information on the operating room staff, surgical procedures and surgical environment characteristics ([2]).
Recorded data will be analyzed and staff behaviors will be assessed by the quantification of displacements within the operating room. Results will aim at bringing a rational to the prevention of airborne microorganism transmission by the description of best behaviors rules in the operating room.
This protocol was approved by the Institutional Review Board of the (IRB) of Paris North Hospitals, Paris 7 University, AP-HP (n° 11-113, April 6 2012). The work is supported by Inria SENSBIO ADT and ARIBO Preqhos project.

6.2.2. Drop-foot correction in post-stroke hemiplegic patients

Participants: Christine Azevedo Coste, Roger Pissard-Gibollet [SED Inria Grenoble Rhône-Alpes], Jérôme Froger [Nîmes Hospital, Le Grau du Roi], Claire Delablachelerie [Nîmes Hospital, Le Grau du Roi].

Electrical stimulation has been proven to have orthotic and carryover effects on individuals with post-stroke hemiplegia with a foot drop syndrome. One of the drawbacks of the technique is the lack of adaptability to changes in gait (speed, type of floor, stairs, dorsiflexion quality etc). But, real-time modification of stimulation patterns is not feasible using gait event detection like proposed in all available stimulators. In the present study we investigate two questions: 1) is it possible to validate on individuals with foot drop an algorithm able to estimate online the continuous gait cycle phase from a unique wireless sensor placed on lower limbs and 2) is it possible to trig a drop foot stimulator based on events extracted from this phase information.

Methods : 20 subjects with post-stroke hemiplegia participated to the study. A wireless inertial measurement unit was placed on the unaffected leg of the subjects and was used to estimate the gait phase on a distant labtop. The subjects performed 3 trials in each of the 3 following conditions: **C1** no stimulation aid, **C2** electrical stimulation assistance triggered by heel switch **C3** electrical stimulation assistance wirelessly triggered based on the proposed algorithm.

Results : 1) the proposed algorithm was able to estimate online the continuous gait cycle phase, 2) events could be extracted from this phase information in order to trig an electrical stimulator using this algorithm instead of heel switch.

Conclusion : the online estimation of continuous gait cycle phase on individuals with stroke is possible. Events can be extracted from the phase information in order to trig a stimulator **C3** instead of using heel switch detection **C2**. The robustness of the proposed solution to gait modifications is intrinsically guaranteed by the use of automatic control theory. These results open promising applications using programmable stimulators which parameters could be modified online based on gait phase observation.

This protocol was approved by Nîmes Ethical Committee, AFSSAPS and CNIL. The work is supported by Inria SENSBIO ADT.

6.2.3. Freezing of Gait detection in Parkinsonian individuals

Participants: Christine Azevedo Coste, Christian Geny [CHU Montpellier], Maud Pasquier [Inria Grenoble Rhône-Alpes], Benoît Sijobert.

Parkinson's disease (PD) is the second most common neurodegenerative disorder. This chronic disease can lead to gait disturbances and falls inducing important reduction of the quality of life. One common symptom is the Freezing of Gait (FOG), an episodic inability to generate effective stepping in the absence of any known cause other than Parkinsonism. It can occur during initiation of the first step, turning, dual task, walking through narrow spaces, reaching destinations or passing through doorways. It is an episodic absence or marked reduction of forward progression of the feet despite the intention to walk. FOG are reported by the patient as a subjective feeling of "the feet being glued to the ground". Clinical evaluation of video recordings of patients by one to three observers is the gold standard to identify FOG events. The evaluation of clinical effects of the treatments would benefit from objective, standardized FOG measures. Moore et al. (2008;2013) have proposed a technique to identify FOG episodes based on the frequency properties of leg vertical accelerations (fig.9). The approach is based on the hypothesis that FOG occurrences are associated to trembling motion, which affect limb acceleration signal. They have introduced the so-called freeze index (FI): the ratio between the signal power in the trembling band (3 Hz - 8 Hz) and the signal power in the locomotor band (0.5 Hz-3 Hz). The FI method was validated using one to 7 accelerometers mounted on patients with satisfactory detection results. If many FOG episodes can be associated to festination (trampling) it is not the case for all of them. Therefore, we claim that all the FOG episodes cannot be detected by the FI method. In the present paper we propose a complementary index in order to take into account not only festination but also other freezing characteristics. Furthermore we intend to propose a solution based on a minimal number of embedded sensors and detection algorithms for future real-time applications.

../../../../projets/demar/IMG/archi.jpg

*Figure 8. **System architecture**. Description of the system architecture used in the study. A sensor node (inertial measurement unit (IMU)) is placed on the unaffected side shank. Data is sent to the sink node of the laptop. Data is processed on the laptop and a gait phase is estimated. Depending on the phase value the stimulator is switched ON through its trigger node. An extra sensor node is also sending data to the sink node and data is saved for offline processing.*

../../../../projets/demar/IMG/freezing.png

Figure 9. Freezing of Gait observation using inertial sensors.

This work is supported by SENSBIO Inria ADT and DEMARPARK AOI (see Partnerships and Cooperations section).

6.2.4. Effects of direct electrical stimulation of the brain during awake surgeries: towards improvements of the functional mapping

Participants: François Bonnetblanc, David Guiraud, Marion Vincent, Mitsuhiro Hayashibe, Hugues Duffau [Neurosurgery department, CHU-Gui de Chauliac], Guillaume Herbet [Neurosurgery department, CHU-Gui de Chauliac], Benedicte Poulin-Charronnat [LEAD, Univ. Dijon].

« Awake surgery » consists in removing some infiltrative and slow-growing brain tumoral tissu in an awake patient. The neurosurgeon performs an anatomo-functional mapping of the brain by electrically stimulating brain areas near the tumor to discriminate functional vs. non functional areas. This stimulation is both made cortically and sub-cortically to preserve the functional connectivity. During the surgery itself, the patients are also involved by performing some tasks. Their recovery remains impressive with respect to the lesional volume. Despite the slow-growth of the lesion is invoked, these observations question our understanding of brain plasticity phenomena. Our multi-disciplinary approach aims to (i) better understand the effects of direct electrical stimulation of the brain to improve the functional mapping and also (ii) to build new functional assessments performed by the patient and based on new technologies applied to Health. By systematically performing these precise assessments before, after and during the surgery we hope to better understand brain functions, plasticity and dynamics in order to improve the surgical planning, functional mapping, rehabilitation procedures and quality of life of the patients.

6.2.5. Translational research and stroke

Participants: Anirban Dutta, David Guiraud.

Stroke is caused when an artery carrying blood from heart to an area in the brain bursts or a clot obstructs the blood flow thereby preventing delivery of oxygen and nutrients. About half of the stroke survivors are left with some degree of disability where the impairment of walking has been mentioned most frequently as the most important disability. There is, therefore, a pressing need to leverage insights from animal and human studies to address the complexity in clinical translation of rational multi-level electrotherapy protocols where the ability to customize such novel electrotherapy protocols has only recently become possible with advanced computational tools. Therefore the challenge is to develop advanced computational modeling tools at Inria, France, to design and customize innovative electrotherapy protocols to patient-specific needs, and then closely integrate them to drive (perhaps the first) individualized non-invasive electrotherapy program for clinical validation. The ongoing steps are i) Develop computational methods to identifying neural circuits related to the recovery from stroke [26], ii) Develop a computational method for online targeting of neural circuits and related pathways with non-invasive electrotherapy [21],[25], iii) Validate individualized multi-level non-invasive electrotherapy program with NIBS as an adjuvant treatment to NMES-assisted gait rehabilitation following stroke.

6.2.6. Projet PERIMED

Participants: Thomas Guiho, Christine Azevedo, Luc Bauchet, Charles Fattal, David Guiraud, Jean-Rodolphe Vignes.

Born in the 70's, Spinal cord stimulation is a general term including both peridural and intradural stimulation. Encouraged by Harkema's clinical result (in one paraplegic patient with step-like EMG activity) [43], several recent studies in rodents elicited locomotor synergies, bladder/bowel improvements and, in certain circumstances, restoration of supraspinal control after spinal cord injury [45]. Based on this previous work, our approach, mainly focused on bladder and bowel functions, aims both at asserting these discoveries in an intermediate model (pigs weighing between 50 and 60 kgs) and at providing further insight in spinal cord circuitries.

6.2.7. Investigation of strategies for selective small nerve fiber stimulation in an animal model

Participants: Paweł Maciejasz, Olivier Rossel, Christine Azevedo Coste, David Andreu, David Guiraud, Hubert Taillades [Institute of Biology, Montpellier].

../../../../projets/demar/IMG/fb.png

Figure 10. Awake brain surgery and functional mapping.

../../../../projets/demar/IMG/pig.png

Figure 11. Supraspinal cord stimulation .

The electrical stimulation of nerve fibers may allow to restore or augment some body functions lost due to disease or injury. However, in typical peripheral nerves there are thousands of nerve fibers with various diameters and functions. When standard rectangular pulses are used for nerve fiber stimulation, the big fibers are activated before smaller ones. However, for many clinical applications it would be beneficial if small fibers could be activated without activation of the big ones.

Already many stimulation techniques have been proposed for fiber type and diameter selective stimulation, e.g. analog block, slowly rising pulses, high frequency block. However, due to limited efficiency of those techniques, they are still not used in clinical practice. Based on the results of the computer simulations and the experiments performed previously in the earthworm model by our team, we have proposed some modification to the existing techniques, that may increase their efficiency.

In order to verify if the proposed modifications allow for increased selectivity of stimulation as compared to the techniques already proposed, a series of experiments in rabbit model has been scheduled. The experiment consists of two phases. The objective of the first one (3-6 rabbits), the preliminary one, is to determine an adequate method to evaluate the effects of stimulation, i.e. to find out a reliable method that would allow for discrimination between various types of fibers being activated by the stimulation. The objective of the second phase (5-15 rabbits), the exploratory one, is to compare and quantify the performance of various strategies for fiber type selective stimulation.

The experiments have been performed by the DEMAR team in rabbits in the Institute of Biology in Montpellier. During this experiment the sciatic nerve of the rabbit has been stimulated using tripolar nerve cuff electrode, whereas ENG and EMG signals, as well as ankle torque have been recorded. The experiments have been started in December 2012 and so far only the first phase of the study has been completed.

The experiment was authorized through the local ethics committee for animal experiment (authorization N° CEEA-LR-12084). The work is supported by INTENSE Project.

6.3. Neuroprostheses and technology

6.3.1. Abstraction and composition for formal design of neuroprotheses

Participants: Hélène Leroux, David Andreu, Karen Godary [LIRMM].

In the framework of specification and implementation of complex digital systems on FPGA, we have developped an approach based on components whose behavior and composition are specified by generalized interpreted T-time Petri nets. One of the inherent difficulties for designer is, on the behavioral part, to account for exceptions. This often leads to a complex modeling and is a source of human errors. Indeed, it is intricate to express all the possible situations (i.e. current state of model). We have defined a way to model exception handling by integrating the well-know concept of macroplace into the formalism. The analysability of the model and the efficiency of the implementation on FPGA (reactivity and surface, ie number of logic blocks) have been preserved. An example of macroplace is given in figure 13; it contains a sub-net (set of places of its refinement) from which exception handling is simply described by a dedicated output transition (transition te on fig. 13), whatever is the current state of the sub-net.

The new formalism has been defined, as well as the model transformation based equivalent PNML generation for using existing analysis tools.

Ongoing work deals with solving state evolution conflicts introducing priorities between transitions, to avoid reaching inconsistent global state while synchronously executing the model.

6.3.2. New FES dedicated digital processor for neurostimulator

Participants: David Andreu, David Guiraud.

We designed (patent pending) and prototyped a new neural FES dedicated processor and its associated (more compact and efficient) set of instructions, as well as an embedded sequencer for accurate timing in sequencing stimulations to be performed (by the stimulator). The new neural stimulator is based on a dedicated ASIC (Application Specific Integrated Circuit), that is able to drive up to 24 channels of stimulation in absolute synchronization, and with a programmable and controlled current level distribution (patent pending). This ASIC also allows for impedance measurement. The functions of the stimulator are currently implemented in two separate chips: an analog stimulation front-end (ASIC) and a field-programmable gate array (FPGA) embedding the logic control. The FPGA embeds the new FES dedicated processor setting the output stage configuration (poles configuration and current ratio between them) and running potentially complex stimulation profiles (with a 1 μs time step and 5 μA current step); example of generated stimulations are shown in Figure 12 . It also embeds the protocol stack allowing for remote programming and online control. Online control relies on advanced and efficient modulation mechanisms, e.g. coefficient based modulation preserving balanced stimulation (Figure 12). And last but not least, it also embeds a monitoring module ensuring the respect of safety constraints stemming both from target tissue protection and electrode integrity preservation; this reference model based monitoring module ensures (configurable) current and quantity of injected charges limits and thus safe stimulation whatever are electrodes to be used (particularly for thin-film micro-electrodes). Safety limits must be defined by users (partners) according to the target and electrodes to be used.

6.3.3. Attitude reconstruction from IMU signals

Participants: Jonathan Peguet [IFMA], Daniel Simon, Christine Azevedo Coste, Roger Pissard-Gibollet [SED Inria Grenoble Rhône-Alpes].

Inertial Measurement Units (IMUs) are currently used by the team for real-time estimation of limbs attitude, e.g. as in section 6.2.2 where the attitude of a leg while walking feeds a agit phase estimator. The IMUs embedded in the FOX nodes (manufactured by HiKob) include 3 gyrometers, 3 accelerometers and 3 magnometers, from which the attitude (e.g. Euler angles) of the node can be computed. The raw measurement signals can be either processed locally in the nodes, or sent on wireless links to be processed on a remote computer.

The raw signals issued from sensors are subject to noise and bias. Additionally, the raw data flow can be corrupted by timing disturbancies induced by communication and computation. Hence, the attitude reconstruction filters must be robust against disturbancies such as noise, bias, jitter and data loss. To evaluate the robustness of attitude reconstruction filters, a software simulation package dedicated to IMUs design and analysis has been customized from the Imusim package (initially developed in Python under GPL at Univ. of Edinburgh, U.K. [44]).

The Imusin modeling features include realistic IMUs models with noise and bias, calibration procedures, radio channels deficiencies and computing timing parameters. Several versions of Extended Kalman Filters and Non-Linear Observers, in particular those previously developed at Inria Grenoble Rhône-Alpes, have been integrated and succesfully tested against measuring noise. The work is supported by Inria SENSBIO ADT.

6.3.4. Fast simulation of hybrid dynamical systems

Participants: Abir Ben Khaled [IFPEN], Daniel Simon, Mongi Ben Gaid [IFPEN].

When dealing with the design of complex systems, simulation is an indisputable step between concept design and prototype validation. Realistic simulations allow for the preliminary evaluation, tuning and possibly redesign of proposed solutions ahead of implementation, thus lowering the risks. However, the simulation of high-fidelity models is time consuming, and reaching real-time constraints is out of the capabilities of monotithic simulations running on single cores.

The aim of the on-going work is to speed up the numerical integration of hybrid dynamical systems, eventually until reaching a real-time execution, while keeping the integration errors inside controlled bounds. The basic approach consists in splitting the system into sub-models, which are integrated in parallel. I has been shown that an efficient partition must minimize the interactions between sub-models, in particular by confining

../../../../projets/demar/IMG/PastedGraphic-1.jpg

Figure 12. Examples of stimulations (observed at the output of the stimulator)

discontinuities processing inside each component. Automatic partitioning, based on some particular incidence matrices of the original system, has been investigated [17]. The method was tested with an automotive engine model, but it is generic and can be applied to other systems of hybrid ODEs/DAEs, as are large sets of muscular fibers.

6.3.5. ENG amplifier front-end

Participants: Mariam Abdallah, Fabien Soulier, Serge Bernard, Guy Cathébras.

Electroneurogram acquisition systems are usually based on tripolar cuff electrodes that are known to decrease noise from external sources, such as muscular fibers (EMG) or stimulation artifacts. Thus, we studied a preamplifier associated with this kind of electrode in a true-tripole configuration. It is designed at the transistor level to lower the number of transistors while still rejecting parasitical signals. This kind of integration reduces the size, power consumption and noise of the preamplifier compared to classical true-tripolar structures.

../../../../projets/demar/IMG/eng_true_tripole.png

Figure 13. Classical structure of a true-tripolar ENG preamplifier.

The true-tripole configuration consists of linear combination of signals coming from the three poles

$$V_{out} = A\left(V_{in1} - \frac{V_{in2} + V_{in3}}{2}\right) \tag{6}$$

This combination is usually realized thanks to several differential amplifiers as shown in the figure 13 , whereas the proposed preamplifier is designes as a differential pair whose negative input transistor is split into only two smaller ones. (fig. 14). The circuit is based on a modified ASIC in AMS CMOS $0.35\,\mu$m technology, with 3.3 V supply. The preamplifier provides three functions which are:

- to combine the input signals as shown in the equation (1),
- to barely amplify the neural signal to an acceptable SNR,
- and to present a differential output to a variable-gain amplifier (not presented here, but integrated into the ASIC).

../../../../projets/demar/IMG/eng_preamplifier.png

Figure 14. Three input preamplifier schematic.

To characterize the three-input preamplifier, we have to define three orthogonal modes, starting with the main mode expected to be amplified

$$V_{in1} = -2V_{in2} = -2V_{in3}, \tag{7}$$

the common mode, and the differential parasitic mode expected to be as low as possible to achieve a good EMG rejection:

$$V_{in1} = V_{in2} = V_{in3}, \tag{8}$$

$$V_{in1} = 0, \qquad V_{in2} = -V_{in3}. \tag{9}$$

DC and AC simulations were performed for these three modes. The results are presented in the figure 15 . The main results of these simulations are:

- more than 150 dB rejection ratio for the common and differential modes compared to main one,
- dynamic range of about 5 mV,
- 200 kHz bandwidth (that is far above the needs for ENG acquisition),
- the estimated flicker noise due to input and load transistors is below the μV on the required bandwidth.

This work has been presented to the 18th IFESS Annual Conference [16].

6.3.6. Characterization of the CAFE12 chip

Participants: Jérémie Salles, Fabien Soulier, Serge Bernard, Guy Cathébras.

../../../../projets/demar/IMG/eng_dc_charact.png ../../../../projets/demar/IMG/eng_ac_charact.png

Figure 15. DC (on the left) and AC (on the right) simulation results for main (b), common (a) and differential modes (c). AC differential mode is too low to be simulated using typical values.

The circuit CAFE12 (Cool Analog Front End, 12 poles) used in StimND, which was designed in 2006, exploit the bases of a DEMAR patent. A characterization of a circuit (1st version, manufactured in 2006-2007) showed limitations to its capabilities. Thus a 2nd version was designed and manufactured in 2012 to improve the circuit linearity and consumption. CAFE12 is an ASIC generating 12 current outputs. This ASIC was developed in high voltage CMOS technology (H35, Austria Mikro Systems). Each output is able to deliver/absorb a current as high as 5mA.

The measurements presented below were carried on 3 CAFE12_V2 prototypes (C8002 & C8003). Some comparisons with 2 CAFE12_V1 (T1201 & nD09) prototypes are also shown.

- Integral non linearity (INL) Figure 16 highlights the INL improvement of the anodic generators. This significant decrease is due to a wiring modification on a specific operational amplifier (OPA). No improvement on the cathodic side was expected.

- Differential non linearity (DNL) The noticed improvement on the INL is also noticeable on the differential non linearity on both anodic and cathodic sides (Figure 17 and Figure 18).

The following table 1 sums up the characterization results. The main achievements are better anodic generators (linearity and gain) and a reduced static consumption.

Table 1.

	INL (LSB)		DNL (LSB)		Gain accuracy $(\%)$		Power (mW)
	Anode	Cathode	Anode	Cathode	Anode	Cathode	Total
CAFE12_V1	-26 to 4	±5	±0.02	±0.015	±13.43	±6.19	50
CAFE12_V2	-6 to 2	-4 to 6	±0.01	±0.01	±6.53	±5.90	38

../../../../projets/demar/IMG/INL_anodes_Anciens_Nouveaux.png

Figure 16. INL comparison, anodic generators (CAFE12_V1 & CAFE12_V2)

../../../../projets/demar/IMG/DNL_anodes_Anciens_Nouveaux.png

Figure 17. DNL comparison, anodic side (CAFE12_V1 & CAFE12_V2)

../../../../projets/demar/IMG/DNL_cathodes_Anciens_Nouveaux.png

Figure 18. DNL comparison, cathodic side (CAFE12_V1 & CAFE12_V2)

DIANA Team

6. New Results

6.1. Service Transparency

Participants: Chadi Barakat, Walid Dabbous, Maksym Gabielkov, Young-Hwan Kim, Arnaud Legout, Byungchul Park, Ashwin Rao, Riccardo Ravaioli, Damien Saucez, Thierry Turletti.

The Complete Picture of the Twitter Social Graph

We made an in-depth study of the macroscopic structure of the Twitter social graph unveiling the highways on which tweets propagate, the specific user activity associated with each component of this macroscopic structure, and the evolution of this macroscopic structure with time for the past 6 years. For this study, we crawled Twitter to retrieve all accounts and all social relationships (follow links) among accounts; the crawl completed in July 2012 with 505 million accounts interconnected by 23 billion links. Then, we presented a methodology to unveil the macroscopic structure of the Twitter social graph. This macroscopic structure consists of 8 components defined by their connectivity characteristics. Each component group users with a specific usage of Twitter. For instance, we identified components gathering together spammers, or celebrities. Finally, we introduced a method to approximate the macroscopic structure of the Twitter social graph in the past, validate this method using old datasets, and discuss the evolution of the macroscopic structure of the Twitter social graph during the past 6 years. This work is accepted in Sigmetrics'14 [23].

Meddle: Middleboxes for Increased Transparency and Control of Mobile Traffic

Meddle is a platform that relies on traffic indirection to diagnose mobile Internet traffic. Meddle is motivated by the absence of built-in support from ISPs and mobile OSes to freely monitor and control mobile Internet traffic; the restrictions imposed by mobile OSes and ISPs also make existing approaches impractical. Meddle overcomes these hurdles by relying on the native support for traffic indirection by mobile OSes. Specifically, Meddle proxies mobile Internet traffic through a software defined middleboxes configured for mobile traffic diagnosis. We use Meddle to tests the limits of the network perspective of mobile Internet traffic offered by traffic indirection. We use this perspective to characterize and control the behavior of mobile applications and provide a first look at ISP interference on mobile Internet traffic. We then performed controlled experiments on 100 popular iOS and Android applications to show how Meddle can be used to identify misbehavior and to block traffic causing this misbehavior. Unlike existing solutions, this activity can be performed without warranty voiding the device and activated on the fly on-demand. This work is done in the context of Aswhin Rao's PhD thesis [11] in collaboration with Northeastern University and Berkeley.

Understanding of modern web traffic

This recent years and with the advent of mobile devices, web traffic has changed and moved from static to dynamic generation. Interestingly, while it is well known that network protocols are intertwined in such a way the characteristics of a layer are affected by those of other layers, most of the measurement work done so far does not pay enough attention to this aspect. We then conducted a cross-layer measurement analysis that confronts all the layers from the very deep technological details to the very high level of users behaviors to shed new light on this issue. To support our study, we analysed an Internet packet traffic trace and showed how this cross-layer analysis approach can explain why TCP flows in mobile traffic are larger than usual. We are currently refining our study to characterises the discrepancies between the different network stack protocol implementations based on the mobile/non-mobile nature of the devices but also their operating system and version. This work is currently under submission.

Checking Traffic Differentiation at the Internet Access

In the last few years, ISPs have been reported to discriminate against specific user traffic, especially if generated by bandwidth-hungry applications. The so-called network neutrality, advocating that an ISP should treat all incoming packets equally, has been a hot topic ever since. We propose Chkdiff, a novel method to detect network neutrality violations that takes a radically different approach from existing work: it aims at both application and differentiation technique agnosticism. We achieve this in three steps. Firstly, we perform measurements with the user's real traffic instead of using specific application traces. Secondly, we do assume that discrimination can take place on any particular packet field, which requires us to preserve the integrity of all the traffic we intend to test. Thirdly, we detect differentiation by comparing the performance of a traffic flow against that of all other traffic flows from the same user, considered as a whole. Chkdiff performance strongly depends on the way routers reply to probe packets. We carried out large scale experiments to understand the way routers reply to our probes and we calibrated models to these replies. The next step will be to evaluate the performance of Chkdiff under these models, before making the tool public and available to the community. Chkdiff is currently the subject of a collaboration with I3S around the PhD thesis of Riccardo Ravaioli (funded by the Labex UCN@Sophia). The work is ongoing and will be submitted soon.

Lightweight Enhanced Monitoring for High-Speed Networks

Within the collaboration with Politecnico di Bari, we worked on LEMON, a lightweight enhanced monitoring algorithm based on packet sampling. This solution targets a pre-assigned accuracy on bitrate estimates, for each monitored flow at a router interface. To this end, LEMON takes into account some basic properties of the flows, which can be easily inferred from a sampled stream, and exploits them to dynamically adapt the monitoring time-window on a per-flow basis. Its effectiveness is tested using real packet traces. Experimental results show that LEMON is able to finely tune, in real-time, the monitoring window associated to each flow and its communication overhead can be kept low enough by choosing an appropriate aggregation policy in message exporting. Moreover, compared to a classic fixed-scale monitoring approach, it is able to better satisfy the accuracy requirements of bitrate estimates. Finally, LEMON incurs a low processing overhead, which can be easily sustained by currently deployed routers, such as a CISCO 12000 device. This work has been published in [18].

Packet Extraction Tool for Large Volume Network Traces

Network packet tracing has been used for many different purposes during the last few decades, such as network software debugging, networking performance analysis, forensic investigation, and so on. Meanwhile, the size of packet traces becomes larger, as the speed of network rapidly increases. Thus, to handle huge amounts of traces, we need not only more hardware resources, but also efficient software tools. However, traditional tools are inefficient at dealing with such big packet traces. We proposed pcapWT, an efficient packet extraction tool for large traces. PcapWT provides fast packet lookup by indexing an original trace using a Wavelet Tree structure. In addition, pcapWT supports multi-threading for avoiding synchronous I/O and blocking system calls used for file processing, and is particularly efficient on machines with SSD. PcapWT shows remarkable performance enhancements in comparison with traditional tools such as tcpdump and most recent tools such as pcapIndex in terms of index data size and packet extraction time. Our benchmark using large and complex traces shows that pcapWT reduces the index data size down below 1% of the volume of the original traces. Moreover, packet extraction performance is 20% better than with pcapIndex. Furthermore, when a small amount of packets are retrieved, pcapWT is hundreds of times faster than tcpdump. These results, done in collaboration within the CIRIC, have just been submitted to Computer Networks[34].

Impact of new transport protocols on BitTorrent performance

In the paper [27], we address the trade-off between the data plane efficiency and the control plane timeliness for the BitTorrent performance. We argue that loss-based congestion control protocols can fill large buffers, leading to a higher end-to-end delay, unlike low-priority or delay-based congestion control protocols. We perform experiments for both the uTorrent and mainline BitTorrent clients, and we study the impact of uTP (a novel transport protocol proposed by BitTorrent) and several TCP congestion control algorithms (Cubic, New Reno, LP, Vegas and Nice) on the download completion time. Briefly, in case peers in the swarm all use the same congestion control algorithm, we observe that the specific algorithm has only a limited impact on the swarm performance. Conversely, when a mix of TCP congestion control algorithms coexists, peers employing a delay-based low-priority algorithm exhibit shorter completion time.

6.2. Open Network Architecture

Participants: Bruno Astuto Arouche Nunes, Chadi Barakat, Daniel Camara, Walid Dabbous, Lucia Guevgeozian Odizzio, Young-Hwan Kim, Mohamed Amine Larabi, Arnaud Legout, Emilio Mancini, Xuan-Nam Nguyen, Thierry Parmentelat, Alina Quereilhac, Damien Saucez, Julien Tribino, Thierry Turletti, Fréderic Urbani.

Delay Tolerant Networks

Delay Tolerant Networks (DTNs) stand for wireless networks where disconnections may occur frequently. In order to achieve data delivery in such challenging environments, researchers have proposed the use of store-carry-and-forward protocols: there, a node may store a message in its buffer and carry it along for long periods of time, until an appropriate forwarding opportunity arises. Multiple message replicas are often propagated to increase delivery probability. This combination of long-term storage and replication imposes a high storage and bandwidth overhead. Thus, efficient scheduling and drop policies are necessary to: *(i)* decide on the order by which messages should be replicated when contact durations are limited, and *(ii)* which messages should be discarded when nodes' buffers operate close to their capacity. We worked on a content-centric dessemination algorithm for delay-tolerant networks, called for short CEDO, that distributes content to multiple receivers over a DTN. CEDO assigns a utility to each content item published in the network; this value gauges the contribution of a single content replica to the network's overall delivery-rate. CEDO performs buffer management by first calculating the delivery-rate utility of each cached content-replica and then discarding the least-useful item. When an application requests content, the node supporting the application will look for the content in its cache. It will immediately deliver it to the application if the content is stored in memory. In case the request cannot be satisfied immediately, the node will store the pending request in a table. When the node meets another device, it will send the list of all pending requests to its peer; the peer device will try to satisfy this list by sending the requester all the matching content stored in its own buffer. A meeting between a pair of devices might not last long enough for all requested content to be sent. We address this problem by sequencing transmissions of data in order of decreasing delivery-rate utility. A content item with few replicas in the network has a high delivery rate utility; these items must be transmitted first to avoid degrading the content delivery-rate metric. The node delivers the requested content to the application as soon as it receives it in its buffer. We implemented CEDO over the CCNx protocol, which provides the basic tools for requesting, storing, and forwarding content. Detailed information on CEDO and the implementation work carried out herein can be found in this publication [22] and at the following web page: http://planete.inria.fr/Software/CEDO/.

Predicting nodes spatial node density in mobile ad-hoc networks

User mobility is of critical importance when designing mobile networks. In particular, "waypoint" mobility has been widely used as a simple way to describe how humans move. This paper introduces the first modeling framework to model waypoint-based mobility. The proposed framework is simple, yet general enough to model any waypoint-based mobility regimes. It employs first order ordinary differential equations to model the spatial density of participating nodes as a function of (1) the probability of moving between two locations within the geographic region under consideration, and (2) the rate at which nodes leave their current location. We validate our models against real user mobility recorded in GPS traces collected in three different scenarios. Moreover, we show that our modeling framework can be used to analyze the steady-state behavior of spatial node density resulting from a number of synthetic waypoint-based mobility regimes, including the widely used

Random Waypoint (RWP) model. Another contribution of the proposed framework is to show that using the well-known preferential attachment principle to model human mobility exhibits behavior similar to random mobility, where the original spatial node density distribution is not preserved. Finally, as an example application of our framework, we discuss using it to generate steady-state node density distributions to prime mobile network simulations. This work was done in collaboration with Dr. Katia Obraczka, from UC Santa Cruz, and was published in WINET [12].

Software Defined Networking in Heterogeneous Networked Environments

We worked on the exploration of the software defined networking paradigm to facilitate the implementation and large scale deployment of new network protocols and services in heterogeneous networked environments. Our activities related to this research thrust are described hereafter. We wrote a survey of the emerging field of Software-Defined Networking (SDN). SDN is currently attracting significant attention from both academia and industry. Its field is quite recent, yet growing at a very fast pace. Still, there are important research challenges to be addressed. We look at the history of programmable networks, from early ideas until recent developments. In particular we described the SDN architecture in detail as well as the OpenFlow standard. We presented current SDN implementations and testing platforms and examined network services and applications that have been developed based on the SDN paradigm. We concluded with a discussion of future directions enabled by SDN ranging from support for heterogeneous networks to Information Centric Networking (ICN). The survey will be published in 2014 in the IEEE Surveys and Tutorials journal [32].

We have also specified a number of use cases motivating the need for extending the SDN model to heterogeneous networked environments. Such environments consist of infrastructure-based and infrastructure-less networks. These specifications and use cases were summarized in a recent publication [19].

We have also implemented a Capacity Sharing platform by leveraging SDN in hybrid networked environments, i.e., environments that consist of infrastructure-based as well as infrastructureless networks. The proposed SDN-based framework provides flexible, efficient, and secure capacity sharing solutions in a variety of hybrid network scenarios. In the paper published at the Capacity Sharing Workshop CSWS 2013 [40], we identify the challenges raised by capacity sharing in hybrid networks, describe our framework in detail and how it addresses these challenges, and discuss implementation issues.

The aforementioned capacity sharing work is just one application and a preliminary of our longer term effort. We have also started to specify the H-SDN protocols based on the use cases mentioned above, including the capacity sharing use case. These efforts are part of a broader work where we propose a framework to enable the implementation and deployment of more generic H-SDN networks and applications. This framework contemplates important issues regarding H-SDN deployment, such as: security, increased scalability and performance by distribution of SDN control and seamless handover of mobile stations, to name a few. We have targeted Mobisys2014 as a venue for publishing our proposal and results regarding this topic [39].

Rule Placement in Software-Defined Networking

OpenFlow is a new communication standard that decouples control and data planes to simplify traffic management. More precisely, OpenFlow switches populate their forwarding tables by opportunistically querying a centralized controller for flows whose rules (i.e., forwarding actions) are not yet installed. However, the flexibility offered by this new paradigm comes at the expense of extra signaling overhead as, in practice, switches might not be able to store all rules in their local forwarding tables. The question of which rules to install then becomes essential. In our research, we leverage the fact that some flows are more important to manage than others, and thus construct an optimal placement problem of rules in OpenFlow switches that ensures the most valuable traffic is matched by its appropriate rules while respecting switches and links capacity constraints. The rest of the traffic with no installed rules follows a default, yet less appropriate, path within the network. We have formulated and solved this optimisation problem in the case of realistic operational needs, and prove that the optimal placement of rules is NP-hard. The intrinsic complexity of the problem led us to design a greedy heuristic that we evaluated with two representative use cases: BGP multihoming and Access Control Lists. On one hand, the evaluation shows the versatility and the generality of the optimization problem, and on another hand, it demonstrates that heuristics with apparent simplicity are still efficient. We are now extending this work to support traffic dynamics and mobility. This work is currently under submission.

Information-Centric Networking and economical aspects

With the explosion of broadband Over-The-Top (OTT) services all around the world, the Internet is autonomously migrating toward overlay and incrementally deployable content distribution infrastructures. Information-Centric Networking (ICN) technologies are the natural candidates to more efficiently bind and distribute popular contents to users. However, the strategic incentives in exploiting ICN, for both users and ISPs, are much less understood to date. In this work, we shed light on how OTTs shall shape prices and discounts to motivate ICN usage, depending on their awareness over content distribution costs. Actually, the Internet ecosystem is fast and dynamic and new ideas can rapidly reach millions of users spread worldwide without having to rely on special involvement of intermediate transit networks. In this context, Over-The-Top broadband content providers can leverage their customer resources to allowing, from one hand, to improve access performance, and, from the other hand, to reduce operational costs the OTT provider would incur on by directly serving the customers. In this context, Information-Centric Networking appears as an adequate offloading technique, if incrementally deployed as an overlay network. This paper analyses the incentive compatibility in the adoption of a ICN overlay for OTT services and is, as of our knowledge, we are the first in addressing the topic by following a non-cooperative game theory reasoning, we believe adequate in its non-cooperative nature due to independency between the involved ICN stakeholders. Our analysis allows us to assess that the business model currently standing for legacy CDNs does not make strategic sense for ICN overlays and that, hover, it exists incentives for OTT customers to get involved in the distributions process via an ICN overlay reducing so server load. These unique specifications for the design of an ICN overlay for OTT content distribution do also have relevant implications for ICN protocol design. The OTT provider would need a form of control over the ICN overlay operations. We identify the usage of a OTT- set policy metric for ICN routing as the most appropriate way to ensure ICN users follow the equilibrium strategy suggested by our incentive compatibility framework. We highlight moreover the need of a scalable way of building and controlling ICN overlays over the legacy TCP/IP Internet to support related signaling, forwarding rule registration, and positive strategic behaviour.

Information-Centric Networking and rate control implications

Information-centric networking (ICN) leverages content demand redundancy and proposes in-network caching to reduce network and servers load and to improve quality of experience. We have studied the interaction between in-network caching of ICN and Additive Increase Multiplicative Decrease (AIMD) end-to-end congestion control with a focus on how bandwidth is shared, as a function of content popularity and caches provisioning. As caching shortens AIMD feedback loop, the download rate of AIMD is impacted. We earlier shed light on the potential negative impact of in-network caching on instantaneous throughput fairness. The work accomplished in 2013 precisely quantify the issue thanks to an analytic model based on Discriminatory Processor Sharing and real experiments, we observe that popular contents benefit from caching and realize shorter download times at the expense of unpopular contents which see their download times inflated by a factor bounded by $\frac{1}{1-\rho}$, where ρ is the network load. This bias can be removed by redefining congestion control to be delay independent or by over-provisioning link capacity at the edge so that to compensate for the greediness of popular contents. The experimentation study has been supported by the work of Ilaria Cianci internship on the CCN-Jocker emulator. This work is currently under submission.

Routing in Information-Centric Network

The idea behind Information-Centric Networking (ICN) is to omit the notion of host and location and use contents as direct routing and forwarding primitives, instead of IP addresses. This shift of paradigm allow ICN to natively offer in-network caching, i.e., to cache content on the path from content providers to requesters. Actually our studies shows a large spatial and temporal locality of contents amongst users in the same network which proves that in-network caching can achieve good overall performance. However, caching contents strictly on their paths is far from being optimal when paths are not shared among content consumers as contents may be replicated on routers so reducing the total volume of contents that can be cached. To overcome this limitation, we introduced the notion of off-path caching in [21] where we allocate content to well defined off-path caches within the network and deflect the traffic off the optimal path toward these caches that are spread across the network. Off-path caching improves the global hit ratio by efficiently utilizing the network-wide available caching capacity and permits to reduce egress links bandwidth usage.

Locator/Identifier Separation Protocol (LISP)

The future Internet has been a hot topic during the past decade and many approaches proposed towards this future Internet, ranging from incremental evolution to complete clean state ones, have been proposed. One of the proposition, LISP, advocates for the separation of the identifier and the locator roles of IP addresses to reduce BGP churn and BGP table size. Up to now, however, most studies concerning LISP have been theoretical and, in fact, little is known about the actual LISP deployment performance. We filled this gap through measurement campaigns carried out on the LISP Beta Network. More precisely, we evaluated the performance of the two key components of the infrastructure: the control plane (i.e., the mapping system) and the interworking (i.e., communication between LISP and non-LISP sites). Our measurements highlight that performance offered by the LISP interworking infrastructure is strongly dependent on BGP routing policies. If we exclude misconfigured nodes, the mapping system typically provides reliable performance and relatively low median mapping resolution delays. Although the bias is not very important, control plane performance favours USA sites as a result of its larger LISP user base but also because European infrastructure is unreliable. Finally, the LISP Map-versioning RFC mentioned in the last year activity report was published this year [33]. All details are reported in [17], [29].

Running Live CCNx Experiments on Wireless and Wired Testbeds with NEPI

CCNx has long left the early development stage where simulation and emulation frameworks, like ccnSim and mininet, were enough to validate new approaches and improvements. It has now reached a level of maturity which calls for evaluation in more realistic environments. If it is to be deployed in the wild Internet or even in private network settings, a framework that provides proper validation in comparable environments is required. For this purpose we demonstrate the capabilities of the NEPI framework to run CCNx experiments in realistic environments. NEPI can run CCNx experiments directly on Internet settings as well as wireless or wired private network environments. This framework allows to automate host con guration, software installation, result collection and to define execution sequence between applications. Furthermore, it provides the ability to conduct interactive experiments where researchers are free to modify the experiment scenario on the fly. These results were demonstrated at CCNxCon'2013 [38].

Evaluating costs of CCN overlays

We are currently involved in a collaboration with PARC (Palo Alto research center) regarding the evaluation of the CCN (Control Centric Networking) technology. Early results of this work were presented in the poster session at the CCNxConf 2013 meeting. In this work we present a set of scenarios to evaluate the performance of CCN overlays on top of the Internet, for worse case conditions. We used the NEPI experiment API to construct different overlay topologies on PlanetLab, for which we varied the topology configuration (e.g. number and degree of nodes), the CCN parameters (e.g. pipeline, cache usage, prefix routes) and the traffic patterns (e.g. single stream, prefix independent chunks). The objective of this study is to find correlations between these variables and the time to deliver content and the overlay network utilization. Our contribution is twofold. In one hand we provide a benchmark which can be used as reference for comparison of new CCNx versions and for other ICN solutions, and as input traces for CCN simulations. In the other hand, we provide results that can be used to improve the CCNx implementation and that can help Internet providers or end users to better design CCN overlays to satisfy their needs. The work is still ongoing and will be submitted soon.

Enabling Iterative Development and Reproducible Evaluation of Network Protocols

Over the last two decades several efforts have been made to provide adequate experimental environments, aiming to ease the development of new network protocols and applications. These environments range from network simulators providing highly controllable evaluation conditions, to live testbeds providing realistic evaluation environment. While these different approaches foster network development in different ways, there is no simple way to gradually transit from one to another, or to combine their strengths to suit particular evaluation needs. We believe that enabling a gradual transition from a pure simulated environment to a pure realistic one, where the researcher can decide which aspects of the environment are realistic and which are controllable, allows improving network solutions by simplifying the problem analysis and resolution. We have designed a new network experimentation framework, called IDEV, where simulated and real components can be arbitrarily combined to build custom test environments, allowing refining and improving new protocols and applications implementations by gradually increasing the level of realism of the evaluation environment. Moreover, we proposed a testbed architecture specifically adapted to support the proposed concept, and discuss the design choices we made based on our previous experience in the area of network testbeds. These choices address key issues in network testbed development, such as ease of experimentation, experiment reproducibility, and testbed federation, to enable scaling the size of experiments beyond what a single testbed would allow. This work has been described in a paper that will be published in the Computer Networks journal in 2014, see [15].

Direct Code Execution: Revisiting Library OS Architecture for Reproducible Network Experiments

We proposed Direct Code Execution (DCE), a framework that dramatically increases the number of available protocol models and realism available for ns-3 simulations. DCE meets the goals recently proposed for fully reproducible networking research and runnable papers, with the added benefits of 1) the ability of completely deterministic reproducibility, 2) the scalability that simulation time dilation offers, 3) capabilities supporting automated code coverage analysis, and 4) improved debuggability via execution within a single address space. We reported on packet processing benchmark and showcased key features of the framework with different use cases. Then, we reproduced a previously published Multipath TCP (MPTCP) experiment and highlight how code coverage testing can be automated by showing results achieving 55-86% coverage of the MPTCP implementation. We also demonstrated how network stack debugging can be easily performed and reproduced across a distributed system. Our first benchmarks are promising and we believe this framework can benefit the network community by enabling realistic, reproducible experiments and runnable papers. This work has been published in the ACM CoNext conference 2013 [25], in Santa Barbara, CA, USA and will be published in IEEE Communication Magazine in 2014 [14]. DCE has been demonstrated at the ACM MSWiM conference at Barcelona, Spain in November 2013 [42] In the same context, we designed DCE Cradle, a framework that allows to use any features of the Linux kernel network stack with existing ns-3 applications. DCE Cradle uses DCE to address the brittleness of Network Simulation Cradle (NSC). We carefully designed DCE Cradle without breaking the existing functionality of DCE and ns-3 socket architecture by considering the gaps between the asynchronous ns-3 socket API and the general POSIX socket API. We validated the implementation of DCE Cradle with the behavior of TCP implementation in congested links, and then studied its performance by focusing on the simulation time and network scale. We showed that DCE Cradle is at most 1.3 times faster than NSC, while it is about 2.2 times slower than the ns-3 native stack. Then we showcased an actual implementation of the DCCP transport protocol to verify how easy it is to simulate a real implementation using DCE Cradle. We believe that this tool can highly benefit the network community by enabling more realistic evaluation of network protocols. This work has been published in the ns-3 workshop in 2013 in Cannes and got the best paper award [26].

The ns-3 Consortium

We have founded in 2012 a consortium between Inria and University of Washington. The goals of this consortium are to (1) provide a point of contact between industrial members and the ns-3 project, to enable them to provide suggestions and feedback about technical aspects, (2) guarantee maintenance of ns-3's core, organize public events in relation to ns-3, such as users' day and workshops and (3) provide a public face that is not directly a part of Inria or NSF by managing the http://www.nsnam.org web site. The Consoritum started his activities in March 2013. Two European institutions (Centre Tecnològic de Telecomunicacions de Catalunya - CTTC and INESC Porto)) and two Amercian universities (Georgia Tech and Bucknell) joined the consortium as Executive members in 2013. For more details see the consortium web page https://www.nsnam.org/consortium/.

Contiki over ns-3

This year we worked on the adaptation of Contiki OS over ns-3. Contiki is a popular, and highly optimized, operating system for sensor nodes. We developed a proof of concept adaptation layer that, even though simple and limited, was able to show that such interaction is indeed possible. The adaptation layer was capable of transferring data from different sensors using ns-3 to interconnect them. Sensor nodes were controlled by the ns-3 scheduler, respecting the ns-3 clock and executing over simulated time. In fact, the sensors were not even aware they were placed over a simulated network.

Federation of experimental testbeds

We are involved in the F-Lab (French ANR) project, the FED4FIRE (E.U. IP) project and have the lead of the "Control Plane Extensions" WorkPackage of OpenLab (E.U. IP) project. Within these frameworks, as part of the co-development agreement between the DIANA team and Princeton University, we kept contributing into one of the most visible and renown implementations of the Testbed-Federation architecture known as SFA for Slice-based Federation Architecture. As a sequel of former activities we also keep a low-noise maintenance activity of the PlanetLab software, which has been running in particular on the PlanetLab global testbed since 2004, with an ad-hoc federated model in place between PlanetLab Central (hosted by Princeton University) and PlanetLab Europe (hosted at Inria) since 2007. During 2013, as a step forward to our contribution to the specification of the Aggregate Manager (AM) API v3, which is the control plane interface through which experimenters discover and reserve resources at testbeds, we have focused on coming up with a separate implementation of SFAWrap that supports AM API v3 and brings a more elaborate lifecycle for slices provisioning. Secondly, we implemented a AM API v2 to AM API v3 adapter, which represents the glue between the already existing AM API v2 compliant testbed drivers and the AM API v3 compliant interfaces of SFAWrap. The v2 to v3 adapter provides AM API v3 compatibility to already existing AM API v2-based testbed drivers until their authors find the time to adapt their driver for a native support of AM API v3 if they want to take full advantage of the new lifecycle. Thirdly, within the contexts of the formerly listed projects, and as a consequence of the growing need for testbeds federation, the providers of testbeds such as: BoneFire, SmartSantander decided to adopt SFAWrap in order to join the global federation of testbeds by exposing their testbeds through SFA. Thus, we had to provide to those partners a close support to achieve this goal. Finally, as for any kind of software development project, and due to the growing usage of SFAWrap, we had to be active on both operational and maintanace tasks. See [37] and [41] for more details. We also contributed, in the context of the Fed4FIRE project, to the definition and early implementation of an architecture for heterogeneous federation of future internet experimental facilities. The results of this work were presented at the FutureNetworkSummit 2013 conference. In this work, requirements involving different aspects of the federation of heterogeneous facilities where collected and analysed, and a multilayer architecture was proposed to address them. Our contribution mainly focuses on the experiment control plane of the federation architecture [28]. The experiment control plane involves the interface between the experimenter and the facilities, and it covers tasks such as federation of the resource discovery, provisioning, reservation, configuration and deployment. The proposed architecture combines the use of SFA (Slice Federation Architecture) and OMF (cOntrol and Management Framework) into a common middle-ware that allows to federate resource control within an experiment across facilities.

DICE Team

6. New Results

6.1. Economy of the world data flows

We have attempted to measure data flows in the world to estimate the concentration of the data industry. It is well known that the main plateforms of the Web, Google, Facebook, Amazon, etc. are concentrated in a few countries, mostly in the USA. Some countries, mostly asian, such as China, Russia, Korea or Japan have succesfully developed their own Web 2.0 industry, while others, such as European countries, hava failed to do so. We have explored in [6], the strategy of China, which has the largest Web industry behind the US and has made a priority of keeping its data at home, with systems in all activity sectors developed in general only one or two years after their main americain counterparts. The innovation strategy of China aims in all fields to achieve technological independence, with at most 30% of foreign IP as we have shown in [2].

A tentative measure of the flows of personal data from different regions is proposed in [3], based on the traffic on the largest platform at the international level. We show in particular that personal data captured in Europe on Web platforms mostly go to the US industry. In [4], we investigate the invisible part of the Web constituted by the trackers that are hidden on Web pages and transfer data to third parties, and show that the domination of the US is even stronger on trackers than it is on the visible Web.

6.2. Flow systems

We are currently working with Bull SA, Manuel Selva (PhD) and Lionel Morel from the Socrates team to build a monitoring framework for dynamic data-flow system in many-core architectures. Data-flow computing models computation as a pipeline of computation units absorbing a continuous stream of data. This computing model suits application development for embedded devices such as MPEG-4 video encoders. The incoming data flow is sliced into small size token (e.g. video frames). Each time, all computational units take some tokens from their inputs and produce some tokens on their outputs. We focus [7], [8] on a management layer for handling dynamic dataflow programs in many-core architectures, where computation units may be relocated at runtime from one core to another. The questions raised by Twitter Storm, Google Millwheel or Yahoo S4, are in essence very similar. Can our current architectures hold the information dataflow produced by users in terms of computing power and memory usage? We are currently extending these embedded results to study dataflow architectures with ATOS on flow computing inside Web browsers.

François Goichon will defend his PhD on resource access equity into best-effort operating systems such as Linux. Linux is built over a layered architecture, where each layer owns a local policy that may lead to a global policy being far from best-effort. With Guillaume Salagnac from Socrates team, we show [5], [9] that we can develop malware user space applications exploiting embedded linux firmware and device drivers differential policy that can block other concurrent applications from accessing CPU time. When this kind of applications are installed in multi-tenant architectures as found in cloud shared space, it can slowdown the entire system. These results are interesting for Dice when considering access time in web browser. Current in-browser applications are developed in Javascript, which imposes a single threaded executed model to the developer, yet operated on a multi-core architecture. Best-effort operating systems are not the best approaches to handle flow based applications that become the norm, and we think that some small, low-level shifts, should be considered.

DIONYSOS Project-Team

6. New Results

6.1. Quality of Experience

Participants: Yassine Hadjadj-Aoul, Adlen Ksentini, Gerardo Rubino, César Viho, Pantelis Frangoudis, Hyunhee Park, Kandaraj Piamrat.

We continue the development of the PSQA technology (Pseudo-Subjective Quality Assessment) in the area of Quality of Experience (QoE). PSQA is today a stable technology allowing to build measuring modules capable of quantifying the quality of a video or an audio sequence, as perceived by the user, when received through an IP network. It provides an accurate and efficiently computed evaluation of quality. Accuracy means that PSQA gives values close to those that can be obtained from a panel of human observers, under a controlled subjective testing experiment, following an appropriate standard (which depends on the type of sequence or application). Efficiency means that our measuring tool can work in real time, if necessary. Observe that perceived quality is, in general, the main component of QoE when the application or service involves video and audio, or voice. PSQA works by analyzing the networking environment of the communication and some the technical characteristics of the latter. It works without any need to the original sequence (as such, it belongs to the family of *no-reference* techniques).

It must be pointed out that a PSQA measuring or monitoring module is network-dependent and application-dependent. Basically, for each specific networking technology, application, service, the module must be built from scratch. But once built, it works automatically and efficiently, allowing if necessary its use in real time, typically for controlling purposes.

Learning tools. At the heart of the PSQA approach there is the statistical learning process necessary to develop measuring modules. So far we have been using Random Neural Networks (RNNs) for that purpose (see [74] for a general description), but recently, we started to explore other approaches. For instance, in the last ten years a new computational paradigm was presented under the name of *Reservoir Computing* (RC) [71] with the goal of attacking the main limitations in training time for recurrent neural networks while introducing no significant disadvantages. Two RC models have been proposed independently and simultaneously under the name of *Liquid State Machine* (LSM) [73] and *Echo State Networks* (ESN) [71]. They constitute today one of the basic paradigms for Recurrent Neural Networks modeling [72]. The main characteristic of the RC model is that it separates two parts: a static sub-structure called *reservoir* which involves the use of cycles in order to provide dynamic memory in the network, and a parametric part composed of a function such as a multiple linear regression or a classical single layer network. The reservoir can be seen as a high-dimensional dynamical system that expand the input stream in a space of states. The learning part of the model is the parametric one. In [41] we propose a new learning tool which merges the capabilities of Random Neural Networks (RNNs) with those of RC models. We keep some of the nice features of RNNs with the ability of RC models in predicting time series values. Our tool is called Echo State Queueing Network. In the paper, we illustrate its performances in predicting, in particular, Internet traffic. In [63], more results about the good behavior of our new tool are presented.

QoE for SVC. A recent video encoding scheme called Scalable Video Coding (SVC) provides the flexibility and the capability to adapt the video quality to varying network conditions and heterogeneous users. Last year, we started to look at the relations between the way SVC is used and the obtained perceived quality. This year we continued these efforts, together with exploring the use of QoE estimation tools for SVC video coding in network control. In [46] we evaluate different configurations for SVC-based adaptive streaming in terms of user QoE. The aim is to provide recommendations about the different rates to be used in order to create the video representation configuration. These results are part of the PhD [11]. In [25], we extended our previous work on SVC in DVB-T2, by proposing an analytical model to evaluate the performance of associating SVC with DVB-T2 and QoE. To do this, we developed a discrete time Markov Chain model which captures the system evolution in terms of number of SVC layers that need to be decoded in order to increase user QoE. In [45], we introduced a new solution to be used by a DASH client for selecting the video representation. Our proposal relies on using the PTP synchronization protocol in order to estimate the end-to-end delays between the client and the server, and hence to correlate this information with network load. The correlation between delays and load was based on a fitting function.

In [54], we focus on SVC multicast over IEEE 802.11 networks. Traditionally, multicast uses the lowest modulation resulting in a video with only base quality even for users with good channel conditions. To optimize QoE, we propose to use multiple multicast sessions with different transmission rates for different SVC layers. The goal is to provide at least the multicast session with acceptable quality to users with bad channel conditions and to provide additional multicast sessions having SVC enhancement layers to users with better channel conditions. The selection of modulation rate for each SVC layer and for each multicast session is achieved with binary integer linear programming depending on network conditions with a goal to maximize global QoE. The results show that our algorithm maximizes global QoE by providing highest quality videos to users with good channel conditions and by guaranteeing at least acceptable QoE for all users.

VoIP. We continued to work on the perceptual quality of voice-based applications and services. In [17], we consider a well-known and widely used *full-reference* technique for measuring speech quality called PESQ, and we propose a learning-based tool for approximating PESQ output without any need for the original signal, following the same black-box parametric PSQA approach. The procedure uses the Echo State Networks previously mentioned.

In [48], we propose a new packet loss model that differentiates loss instances depending on their perceptual impact. In particular, the model captures the differences between short and long interruptions from the perceptual quality viewpoint. In some cases, the delays and their variation have a strong impact on the perceived quality. In [49] we explore the variability of packet delays on MANETs. For that purpose, a wide range of representative scenarios are defined and simulated. The gathered traces are then inspected from qualitative and quantitative perspectives. In [50], a Markovian model is proposed to capture these and other features of delays in the same class of mobile networks.

6.2. Network Economics

Participant: Bruno Tuffin.

The general field of network economics, analyzing the relationships between all actors of the digital economy, has been an important subject for years in the team.

A new book on the subject. We have published a book on this broad topic [61]. Presenting a balance of theory and practice, this up-to-date guide provides a comprehensive overview of the key issues in telecommunication network economics, as well as the mathematical models behind the solutions. These mathematical foundations enable the reader to understand the economic issues arising at this pivotal time in network economics, from business, research, and political perspectives. This is followed by a unique practical guide to current topics, including app stores, volume-based pricing, auctions for advertisements, search engine business models, the network neutrality debate, the relationship between mobile network operators and mobile virtual network operators, and the economics of security. The guide discusses all types of players in telecommunications, from users, to access and transit network providers; to service providers (including search engines, cloud providers or content delivery networks); to content providers, and regulatory bodies. The book is designed for graduate students, researchers, and industry practitioners working in telecommunications.

Research contributions in network economics during 2013 can be decomposed into the application of auction theory, cognitive networks, and network/search neutrality analysis.

Auction theory. In the next generation Internet, we have seen the convergence of multimedia services and Internet with the mobility of users. Vertical handover decision (VHD) algorithms are essential components of the mobility management architecture in mobile wireless networks. VHD algorithms help mobile users to choose the best mobile network to connect among available candidates. It also can help the network manager to optimize easily the limited resources shared among the network providers and the users. In [26], we formulate

VHD algorithm as a resource allocation problem for down-link transmission power in multiple W-CDMA networks and show how combinatorial double-sided auctions can be applied to this specific problem. The proposed pricing schemes make use of the signal interference to noise ratio (SINR), achievable data rates, power allocation at mobile networks, and monetary cost as decision criteria, and our model differentiates new calls and on-going communications to take into account that the last category has somewhat more importance. Several combinatorial double-sided auction are proposed to maximize the social welfare and /or to provide incentives for mobile users and mobile operators to be truth-telling in terms of valuation or cost. Finally, the economic properties of the different proposed pricing schemes are also studied by means of simulations.

Cognitive networks. Cognitive radio technologies for spectrum sharing have received an enormous interest from the research community for the last decade, and more recently from regulators and mobile operators. We have studied a cognitive radio network in [47] where primary operator and an entrant secondary operator compete for users. The system is modeled using queueing and game theories. The economic viability of supporting the secondary operator service using an opportunistic access to the spectrum owned by the primary operator is assessed. Against the benchmark of the primary operator operating as a monopolist, we show that the entry of the secondary operator is desirable from an efficiency perspective, since the carried traffic increases. For a range of parameter values, a lump sum payment can be designed so that the incumbent operator has an incentive to let the secondary operator enter. Additionally, the opportunistic access setting has been compared against a leasing-based alternative, and we have concluded that the former outperforms the latter in terms of efficiency and incentive.

Network/search neutrality analysis. Network neutrality is the topic of a vivid and very sensitive debate, in both the telecommunication and political worlds, because of its potential impact in everyday life. That debate has been raised by Internet Service Providers (ISPs), complaining that content providers (CPs) congest the network with insufficient monetary compensation, and threatening to impose side payments to CPs in order to support their infrastructure costs. While there have been many studies discussing the advantages and drawbacks of neutrality, there is no game-theoretical work dealing with the observable situation of competitive ISPs in front of a (quasi-)monopolistic CP. However, this is a typical situation that is condemned by ISPs, and, according to them, another reason of the non-neutrality need. We develop and analyze in [23] a model describing the relations between two competitive ISPs and a single CP, played as a three-level game corresponding to three different time scales. At the largest time scale, side payments (if any) are determined. At a smaller time scale, ISPs decide their (flat-rate) subscription fee (toward users), then the CP chooses the (flat-rate) price to charge users. Users finally select their ISP (if any) using a price-based discrete choice model, and decide whether to also subscribe to the CP service. The game is analyzed by backward induction. As a conclusion, we obtain among other things that non-neutrality may be beneficial to the CP, and not necessarily to ISPs, unless the side payments are decided by ISPs.

The very related recently raised search neutrality debate questions the ranking methods implemented by search engines: when a search is performed, do they (or should they) display the web pages ordered according to the quality-of-experience (relevance) of the content? In [68], we analyze that question in a setting when content is offered for free, content providers making revenue through advertising. For content providers, determining the amount of advertising to add to their content is a crucial strategic decision. Modeling the trade-off between the revenue per visit and the attractiveness, we investigate the interactions among competing content providers as a non-cooperative game, and consider the equilibrium situations to compare the different ranking policies. Our results indicate that when the search engine is not involved with any high-quality content provider, then it is in its best interest to implement a neutral ranking, which also maximizes user perceived quality-of-experience and favors innovation. On the other hand, if the search engine controls some high-quality content, then favoring it in its ranking and adding more advertisement yields a larger revenue. This is not necessarily at the expense of user perceived quality, but drastically reduces the advertising revenues of the other content providers, hence reducing their chances to innovate.

6.3. Wireless and Mobile Networks

Participants: Yassine Hadjadj-Aoul, Adlen Ksentini, César Viho, Osama Arouk, Btissam Er-Rahmadi, Hyunhee Park, Kandaraj Piamrat.

We continue our activities around wireless and mobile networks, where we focus particularly on 4G networks as well as on a new mobile architecture known as mobile cloud.

LTE improvements. First part of our works concentrates on emerging applications and their impact on 4G networks. In [58], we proposed a solution to handle social network traffic, which is characterized by its elasticity and intensity in a short period of time. The proposed contribution is based on content detection systems such as Deep Packet Inspection (DPI) to identify traffic belonging to a group of users (sharing the same content) of a social network. Upon detecting the type of traffic, we proposed to control it by creating a multicast group. This would reduce the amount of traffic exchanged by switching from unicast communications to multicast communications. Another solution is to cache, at the geographically nearest base station, the shared content among users. Here we positioned ourselves in the case where the social network traffic comes from the same geographical region. We also investigated network decentralization in conjunction with the selective IP traffic offload approaches to handle such increased data traffic. We first devised different approaches based on a per-destination-domain-name basis, which offer operators a fine-grained control to determine whether a new IP connection should be offloaded or accommodated via the core network. Two of our solutions are based on Network Address Translation (NAT) named simple-NATing and twice-NATing, whereas a third one employs simple tunneling, and a fourth adopts multiple Access Point Names. We also proposed methods enabling user equipment devices to always have efficient packet data network connections [30]. Another aspect, we addressed is the gateway selection process, where in [59] we argue the need for other metrics to improve the gateway selection mechanisms in distributed mobile networks. We therefore proposed to consider the end-to-end connection and the service/application type as two important additional metrics in the selection of data anchor gateways in the context of the Evolved Packet System (EPS).

M2M. In [56], [32] we addressed another type of traffic that appeared these last years, namely Machine to machine communication or Machine Type Communication (MTC). Such traffic is known by its intensity and its impact on increasing congestion in both parts of the 4G networks, the Radio Access Network (RAN) and the core network parts. The main spirit of the proposed solutions is to proactively anticipate system overload by reducing the amount of MTC signaling messages exchanged in normal network operations. The first solution reduces the number of exchanged signaling messages when triggering MTC devices with low mobility. It enables direct triggering of MTC devices with low mobility by MTC-IWF (MTC InterWorking Function), without involving the MME (Mobility Management Entity). Second solution defines a method for controlling and anticipating network overload in case of an event/scenario whereby a mass of messages with some common Information Elements (IE) are to be exchanged on an interface between two nodes. The network overload control is achieved via dynamic creation of a profile characterizing the event/scenario and the common IEs.

Home networks. In-home wireless networks are now wide-spreading as today's home network is composed of at least one wireless network. The dramatic increase of traffic in such networks yields to difficulties in guaranteeing user experience especially for some specific services like IPTV. This is particularly complicated when using UDP at transport layer and traditional MAC protocol at link layer. Therefore, we investigated comparison of different combinations of transport and link layer performances for the delivery of IPTV. For validation, we use NS-3 and a realistic propagation model generated with a real house description. We analyze impact of link layer (with or without coordination) and transport layer (UDP or TCP). Then, we propose a combined solution using TCP over a coordinated MAC protocol (see [52]). The proposed solution can be easily deployed in real products and is compatible with existing devices.

Another part of our activities in wireless network are related to energy saving. Indeed, one of the biggest problem today in the wireless world is that wireless devices are battery driven, which reduce their operating lifetime. The experimental measurements we have achieved in [18], [42] revealed that operating system overhead causes a drop in performance and energy consumption properties as compared to the GPP in case of certain low video qualities. We propose, thus, a new approach for energy-aware processor switching (GPP or DSP) which takes into consideration the video quality. We show the pertinence of our solution in the context of adaptive video decoding and implement it on an embedded Linux operating system.

6.4. Future Networks

Participants: Yassine Hadjadj-Aoul, Adlen Ksentini, Leila Ghazzai, Jean-Michel Sanner.

Mobile cloud. One of the 5G-architecture visions considers the usage of cloud to build mobile networks and help in decentralizing mobile networks on demand, elastically, and in the most cost-efficient way. This concept of carrier cloud becomes of vital importance knowing that several cloud providers are distributing their cloud/network, globally deploying more regional data centers, to meet their ever-increasing business demands. As an important enabler of the carrier cloud concept, network function virtualization (NFV) is gaining great momentum among industries. NFV aims for decoupling the software part from the hardware part of a carrier network node, traditionally referring to a dedicated hardware, single service and single-tenant box, and that is using virtual hardware abstraction. Network functions become thus a mere code, runnable on a particular, preferably any, operating system and on top of a dedicated hardware platform. The ultimate objective is to run network functions as software in standard virtual machines (VMs) on top of a virtualization platform in a general-purpose multi-service multi-tenant node (e.g., Carrier Grade Blade Server) put into the cloud. In [31], we presented and detailed the Follow Me Cloud (FMC) concept, whereby mobile services hosted in federated clouds follow mobile users as they move and according to their needs. We then provided in [55] a detailed analytical model based on continuous time Markov chain which considers to evaluate the performance of FMC in terms of service migration cost and QoS gain for user. An efficient mobile cloud cannot be built without efficient algorithms for the placement of NFV over this federated cloud. In this vein, in [57] we argued the need for avoiding or minimizing the frequency of mobility gateway (S-GW) relocations and discussed how this gateway relocation avoidance can be reflected in an efficient network function placement algorithm for the realization of mobile cloud. The problem was modeled by an Integer Linear Problem and proved to be NP hard. Therefore, two heuristics were proposed for the creation of a NFV S-GW instance in the cloud.

SDN. We started an activity on Software Defined Networking (SDN), a recent idea proposed to handle network management problems. SDN are becoming an important issue with the ever-increasing network complexity. They are proposed as an alternative to the current architecture of the Internet, which cannot meet the supported services requirements such as Quality of Service/Experience (Qos/QoE), security and energy consumption. We particularly address the scalability issue by proposing a hierarchical controller-based architecture handling the whole control chain.

6.5. Interoperability assessment and improvement

Participants: César Viho, Anthony Baire, Nanxing Chen.

The Internet of Things (IoT) brings new challenges to interoperability assessment by introducing the necessity to deal with non reliable environments connecting plenty billions of objects widely distributed. Therefore, in the recent period, we propose an interoperability testing methodology using a *passive* approach. It appeared more suitable for this distributed, unrliable and constrained environment brought by IoT. We have also developed a tool that implements this passive method. It has been used successfully to test CoAP implementations during the two CoAP Plugtest interoperability sessions on IoT protocols (CoAP and 6LoWPAN) organized by ETSI and IPSO Alliance. These contributions are published in [10].

6.6. Performance Evaluation of Distributed Systems

Participants: Bruno Sericola, Romaric Ludinard.

Network Monitoring and Fault Detection. Monitoring a system is the ability of collecting and analyzing relevant information provided by the monitored devices so as to be continuously aware of the system state. However, the ever growing complexity and scale of systems makes both real time monitoring and fault detection a quite tedious task. Thus the usually adopted option is to focus solely on a subset of information states, so as to provide coarse-grained indicators. As a consequence, detecting isolated failures or anomalies is a quite challenging issue. We propose in [38] and [60] to address this issue by pushing the monitoring task at the edge of the network. We present a peer-to-peer based architecture, which enables nodes to adaptively and efficiently self-organize according to their ”health” indicators. By exploiting both temporal and spatial correlations that exist between a device and its vicinity, our approach guarantees that only isolated anomalies (an anomaly is isolated if it impacts solely a monitored device) are reported on the fly to the network operator. We show that the end-to-end detection process, *i.e.*, from the local detection to the management operator reporting, requires a logarithmic number of messages in the size of the network. These results also led to the patent [70].

Robustness Analysis of Large Scale Distributed Systems. In the continuation of previous work which proposed an in-depth study of the dynamicity and robustness properties of large-scale distributed systems, in [15] we analyze the behavior of a stochastic system composed of several identically distributed, but non independent, discrete-time absorbing Markov chains competing at each instant for a transition. The competition consists in determining at each instant, using a given probability distribution, the only Markov chain allowed to make a transition. We analyze the first time at which one of the Markov chains reaches its absorbing state. When the number of Markov chains goes to infinity, we analyze the asymptotic behavior of the system for an arbitrary probability mass function governing the competition. We give conditions for the existence of the asymptotic distribution and we show how these results apply to cluster-based distributed systems when the competition between the Markov chains is handled by using a geometric distribution.

Secure Uniform Sampling in Dynamic Systems. In [37], we consider the problem of achieving uniform node sampling in large scale systems in presence of a strong adversary. We first propose an omniscient strategy that processes on the fly an unbounded and arbitrarily biased input stream made of node identifiers exchanged within the system, and outputs a stream that preserves Uniformity and Freshness properties. We show through Markov chains analysis that both properties hold despite any arbitrary bias introduced by the adversary. We then propose a knowledge-free strategy and show through extensive simulations that this strategy accurately approximates the omniscient one. We also evaluate its resilience against a strong adversary by studying two representative attacks (flooding and targeted attacks). We quantify the minimum number of identifiers that the adversary must insert in the input stream to prevent uniformity. To our knowledge, such an analysis has never been proposed before.

6.7. Monte Carlo

Participants: Gerardo Rubino, Bruno Tuffin, Pablo Sartor Del Giudice.

We maintain a research activity in different areas related to dependability, performability and vulnerability analysis of communication systems, using both the Monte Carlo and the Quasi-Monte Carlo approaches to evaluate the relevant metrics. Monte Carlo (and Quasi-Monte Carlo) methods often represent the only tool able to solve complex problems of these types. However, when the events of interest are rare, simulation requires a special attention, for two reasins: the need in accelerating the occurrence of those events and in getting unbiased estimators of them with a sufficiently small relative variance. This is the main problem in the area. Dionysos' work focuses then in dealing with the rare event situation. Rare event simulation has been reviewed in [22].

Multidimensional integrals. In [20], we present a versatile Monte Carlo method for estimating multidimensional integrals, with applications to rare-event probability estimation. The method combines two distinct and popular Monte Carlo simulation techniques, Markov chain Monte Carlo and Importance Sampling, into a single algorithm. We show that for some applied numerical examples the proposed Markov Chain Importance Sampling algorithm performs better than methods based solely on Importance Sampling or MCMC.

Static models. Static reliability analysis has been the topic of an extensive activity in the group for years. Exact evaluation of static network reliability parameters belongs to the NP-hard family and Monte Carlo simulation is therefore a relevant tool to provide estimations for them.

In [67], we first review a Recursive Variance Reduction (RVR) estimator which approaches the unreliability metric by recursively reducing the graph from the random choice of the first working link on selected cuts. We show that the method does not verify the bounded relative error (BRE) property as reliability of individual links goes to one, i.e., that the estimator is not robust in general to high reliability of links. We then propose to use the decomposition ideas of the RVR estimator in conjunction with the Importance Sampling technique.

Two new estimators are presented: the first one, called Balanced Recursive Decomposition estimator, chooses the first working link on cuts uniformly, while the second, called Zero-Variance Approximation Recursive Decomposition estimator, tries to mimic the estimator with variance zero for this technique. We show that in both cases the BRE property is verified and, moreover, that a Vanishing Relative Error property can be obtained for the Zero-Variance Approximation RVR under specific sufficient conditions. A numerical illustration of the power of the methods is provided on several benchmark networks.

The same problem is also analyzed in [19] by a novel method that exploits a generalized splitting (GS) algorithm. We show that the proposed GS algorithm can accurately estimate extremely small unreliabilities and we exhibit large examples where it performs much better than existing approaches. Remarkably, it is also flexible enough to dispense with the frequently made assumption of independent edge failures.

On the same type of model, we propose in [51] an adaptive parameterized method to approximate the zero-variance change of measure for the evaluation of static network reliability models, with links subject to failures. The method uses two rough approximations of the unreliability function, conditional on the states of any subset of links being fixed. One of these approximation, based on mincuts, under-estimates the true unknown unreliability, whereas the other one, based on minpaths, over-estimates it. Our proposed change of measure takes a convex linear combination of the two, estimates the optimal (graph-dependent) coefficient in this combination from pilot runs, and uses the resulting conditional unreliability approximation at each step of a dynamic Importance Sampling algorithm. This new scheme is more general and more flexible than a previously-proposed zero-variance approximation one, which is based on mincuts only and which was shown to be robust asymptotically when unreliabilities of individual links decrease toward zero. Our numerical examples show that the new scheme is often more efficient when the unreliabilities of the individual links are not so small but the overall unreliability is small because the system can fail in many ways. Part of these results are in the PhD [13].

In [43], we present a generalization of the above static models to cases for which the component failures are not independent. To model the dependence and also to develop effective simulation methods that estimate the system unreliability, we extend the static model into an auxiliary dynamic one where the components fail at random times, according to a Marshall-Olkin multivariate exponential distribution. We examine and compare different versions of this model and develop efficient unreliability estimation methods based on conditional Monte Carlo and on a generalized splitting methodology.

In [28], a different splitting algorithm is proposed for solving the same static problem, which is converted into a dynamic one by means of the Creation Process of Elperin, Gerbtsbakh and Lomonosov. The classic splitting technique is then applied, and the obtained results are explored through several numerical experiments. The relative error and the covering properties of the obtained estimator are particularly studied.

In [29], a generalization of the basic model is studied using Monte Carlo. The idea is that the system (the network) works when the terminal nodes are connected *by at least one path whose length is less than or equal to a given parameter* d. This is called Diameter Constrained Reliability. If the parameter d is greater than or equal to the longest path in the network (or between terminals), the problem is the classic one. The paper proposes a variance reduction technique for the estimation of the system's reliability in this setting. In [21], we analyze the particular case of $d = 2$ using exact techniques. These results are part of the thesis [14].

Finally, in [34] and [36] we made general presentations on the rare event problem in general, and on some of the team's results concerning the design of efficient techniques to analyze them.

6.8. Analytic models

Participants: Raymond Marie, Bruno Sericola, Gerardo Rubino, Laura Aspirot.

New books about Markovian models and applications. The book [65] is the french version of the book [66]. Markov chains are a fundamental class of stochastic processes. They are the main modeling tool used in our team. They are widely used to solve problems in a large number of domains such as operations research, computer science, communication networks and manufacturing systems. The success of Markov chains is mainly due to their simplicity of use, the large number of available theoretical results and the quality of algorithms developed for the numerical evaluation of many metrics of interest. The books present the theory of both discrete-time and continuous-time homogeneous Markov chains. They examine the explosion phenomenon, the Kolmogorov equations, the convergence to equilibrium and the passage time distributions to a state and to a subset of states. These results are applied to birth-and-death processes. A detailed study of the uniformization technique by means of Banach algebra results is also developed. This technique is used for the transient analysis of several queuing systems.

Another book entitled “Markov Chains and Dependability Theory” will be published soon by Cambridge University Press (see http://www.amazon.fr/Markov-Chains-Dependability-Theory-Gerardo/dp/1107007577/). Dependability metrics are omnipresent in every engineering field, from simple ones through to more complex measures combining performance and dependability aspects of systems. The book presents the mathematical basis of the analysis of these metrics in the most used framework, Markov models, describing both basic results and specialised techniques. It presents both discrete and continuous time Markov chains before focusing on dependability measures, which necessitate the study of Markov chains on subsets of states representing different user satisfaction levels for the modelled system. Topics covered include Markovian state lumping, analysis of sojourns on subset of states of Markov chains, analysis of most dependability metrics, fundamentals of performability analysis, and bounding and simulation techniques designed to evaluate dependability measures. The book is of interest to graduate students and researchers in all areas of engineering where the concepts of lifetime, repair duration, availability, reliability and risk are important.

Fluid models. In [53] and [44] we propose a new way of transporting video flows on a peer-to-peer architecture of the Bit-Torrent type. We analyze the performance obtained by our proposal by means of fluid views of the systems, that is, by representing them using differential equations. In [53] the basic idea is to select the downloading peers according to their progress in the downloading process: a given peer only sends chunks to other peers that are downloading at least roughly in the same “area” of the stream. The system is improved in [44] where the main resource (the available bandwidth) is distributed differently among the peers, giving some kind of priority to those nodes remaining more time connected.

In [39], we look at the problem of approximating Markovian views of the Machine Repaiman Model where life-times and repair times have Phase-type distributions, by differential equations. The machine population goes to infinity, and we analyze the properties of the limiting differential equation (once the Markovian sequence of models is properly scaled) and their relations with the initial models. In [63] we describe these results and other results concerning the same type of limiting processes, but concerning peer-to-peer networks. We discuss here the convergence aspects; the properties of the fluid models themselves are discussed in the two papers [53] and [44] mentioned before.

DISCO Project-Team

6. New Results

6.1. Equidimensional block-triangular representation of linear functional systems

Participant: Alban Quadrat.

In [30], it is shown that every linear functional system (e.g., PD systems, differential time-delay systems, difference systems) is equivalent to a linear functional system defined by an upper block-triangular matrix of functional operators: each diagonal block is respectively formed by a generating set of the elements of the system satisfying a purely i-codimensional system. Hence, the system can be integrated in cascade by successively solving (inhomogeneous) i-codimensional linear functional systems to get a Monge parametrization of its solution space [120]. The results are based on an explicit construction of the grade/purity filtration of the module associated with the linear functional system. This new approach does not use complex Grothendieck spectral sequence arguments as is done in the literature of modern algebra [86], [87]. To our knowledge, the algorithm obtained in [30] is the most efficient algorithm existing in the literature of non-commutative algebra. It was implemented in the PURITYFILTRATION package developed in Maple (see Section 5.6) and in the `homalg` package of GAP 4 (see Section 5.7). Classes of overdetermined/underdetermined linear systems of partial differential equations which cannot be directly integrated by Maple can be solved using the PURITYFILTRATION package.

6.2. Serre's reduction of linear functional systems and related problems

Participants: Alban Quadrat, Thomas Cluzeau [ENSIL, Univ. Limoges].

Given a linear multidimensional system (e.g., ordinary/partial differential systems, differential time-delay systems, difference systems), Serre's reduction aims at finding an equivalent linear multidimensional system which contains fewer equations and fewer unknowns. Finding Serre's reduction of a linear multidimensional system can generally simplify the study of structural properties and of different numerical analysis issues, and it can sometimes help solving the linear multidimensional system in closed form. The connection between Serre's reduction and the decomposition problem [94], which aims at finding an equivalent linear functional system which is defined by a block diagonal matrix of functional operators, is algorithmically studied in [41], [42]. Moreover, a characterization of isomorphic finitely presented modules in terms of certain inflations of their presentation matrices is obtained in [42]. This result yields a connection between a certain matrix completion problem and Serre's reduction [42].

6.3. Algorithmic study of linear PD systems and Stafford's theorems

Participants: Alban Quadrat, Daniel Robertz [Univ. Aachen].

In [121],[82], algorithmic versions of Stafford's results [124] (e.g., computation of unimodular elements, decomposition of modules, Serre's splitting-off theorem, Stafford's reduction, Bass' cancellation theorem, minimal number of generators) were obtained and implemented in the STAFFORD package [82]. In particular, we show how a determined/overdetermined linear system of partial differential equations with either polynomial, rational, formal power series or locally convergent power series coefficients is equivalent to a linear system of partial differential equations with at most two unknowns. This result is a large generalization of the cyclic vector theorem which plays an important role in the theory of linear ordinary differential equations.

6.4. Foundations of the behavioural approach

Participant: Alban Quadrat.

Within the algebraic analysis approach to behaviours [91], [113], in [34], we propose to consider a system not only as a behaviour $\mathrm{ext}^0_D(M, \mathcal{F})$ [107], where M is the finitely presented left D-module defined by the matrix defining the system and $\mathcal{F}$ the signal space, but as the set of all the $\mathrm{ext}^i_D(M, \mathcal{F})$'s, where $0 \leq i \leq n$, where n is the global dimension of D. In this new framework, using Yoneda product, the left D-homomorphims of M [94] and the internal symmetries of the behaviour $\mathrm{ext}^0_D(M, \mathcal{F})$ [94] are generalized to the full system $\{\mathrm{ext}^i_D(M, \mathcal{F})\}_{i=0,\ldots,n}$ In particular, a system-theoretic interpretation of the Yoneda product is given.

In [117], we study the construction of a double complex leading to a Grothendieck spectral sequence converging to the obstructions $\mathrm{tor}^i_D(N, \mathcal{F})$'s for the existence of a chain of successive parametrizations starting with the behaviour $\mathrm{ext}^0_D(M, \mathcal{F})$, where N is the Auslander transpose of M. These obstructions $\mathrm{tor}^i_D(N, \mathcal{F})$ can be studied by means of a long process starting with the $\mathcal{F}$-obstructions $\mathrm{ext}^j_D(\mathrm{ext}^k_D(N, D), \mathcal{F})$'s for the solvability of certain inhomogeneous linear systems defined by the algebraic obstructions $\mathrm{ext}^k_D(N, D)$'s measuring how far M is for being a projective left D-module. Hence, the algebraic properties of the left D-module M, defining the behaviour $\mathrm{ext}^0_D(M, \mathcal{F})$, and the functional properties of the signal space $\mathcal{F}$ can be simultaneously used to study the obstructions for the existence of a chain of successive parametrizations starting with the behaviour $\mathrm{ext}^0_D(M, \mathcal{F})$. These results can be used to find again the different situations studied in the literature (e.g., cases of an injective or a flat left D-module $\mathcal{F}$). Finally, setting $\mathcal{F} = D$, the above results can be used to find again the characterization of the grade/purity filtration of M by means of a Grothendieck spectral sequence. See Section 6.1 and [86], [87], [30].

Within the algebraic analysis approach to behaviours [91], [113], in [116], we explain how the concept of inverse image of a finitely presented left D-module M, defining the behaviour $\mathrm{ext}^0_D(M, \mathcal{F})$ [107], can be used to study the problem of characterizing the restriction of the behaviour $\mathrm{ext}^0_D(M, \mathcal{F})$ to a non characteristic submanifold of $\mathbb{R}^n$. In particular, we detail the explict construction of inverse images of left D-modules for standard maps.

6.5. Boundary value problems for linear ordinary integro-differential equations

Participants: Alban Quadrat, Georg Regensburger.

In [35], we study algorithmic aspects of linear ordinary integro-differential operators with polynomial coefficients. Even though this algebra is not noetherian and has zero divisors, Bavula recently proved in [85] that it is coherent, which allows one to develop an algebraic systems theory. For an algorithmic approach to linear systems theory of integro-differential equations with boundary conditions, computing the kernel of matrices is a fundamental task. As a first step, we have to find annihilators, which is, in turn, related to polynomial solutions. We present an algorithmic approach for computing polynomial solutions and the index for a class of linear operators including integro-differential operators. A generating set for right annihilators can be constructed in terms of such polynomial solutions. For initial value problems, an involution of the algebra of integro-differential operators also allows us to compute left annihilators, which can be interpreted as compatibility conditions of integro-differential equations with boundary conditions. These results are implemented in MAPLE based on the `IntDiffOp` and `IntDiffOperations` packages.

6.6. Noncommutative geometry approach to infinite-dimensional systems

Participant: Alban Quadrat.

In [112], [111], [110], it was shown how the fractional representation approach to analysis and synthesis problems developed by Vidyasagar, Desoer, Callier, Francis, Zames..., could be recast into a modern algebraic analysis approach based on module theory (e.g., fractional ideals, algebraic lattices) and the theory of Banach algebras. This new approach successfully solved open questions in the literature. Basing ourselves on this new approach, we explain in [114], [115] why the non-commutative geometry developed by Alain Connes is a natural framework for the study of stabilizing problems of infinite-dimensional systems. Using the 1-dimensional quantized calculus developed in non-commutative geometry and results obtained in [112], [111], [110], we show that every stabilizable system and their stabilizing controllers naturally admit geometric structures such as connections, curvatures, Chern classes, ... These results developed in [114], [115] are the first steps toward the use of the natural geometry of the stabilizable systems and their stabilizing controllers in the study of the important H_∞ and H_2-problems.

6.7. Stability analysis of fractional neutral systems with commensurate delays

Participants: Le Ha Vy Nguyen, Catherine Bonnet.

Fractional neutral systems with commensurate delays have chains of poles aymptotic to vertical lines. The case where the imaginary axis is an asymptotic one is interesting. Indeed, if the system has some chains of poles asymptotic to the imaginary axis, then the fact that all poles lie in the open left half-plane does not guarantee the H_∞-stability of the system.

This kind of systems was studied in [97], [104]. In [97], systems with single chains of poles asymptotic to the imaginary axis was considered and necessary and sufficient conditions for H_∞-stability were derived. Some particular systems with multiple chains have been examined in [104]. We have extended this year this study to more general systems with multiple chains of poles approaching the imaginary axis.

6.8. Stabilization of fractional neutral systems with commensurate delays

Participants: Le Ha Vy Nguyen, Catherine Bonnet.

We consider fractional neutral systems with commensurate delays which may have chains of poles asymptotic to vertical lines lying in the open left half-plane and have chains clustering the imaginary axis. Due to the latter, the system may possess infinitely many poles in the right half-plane. We prove that a class of rational fractional controllers cannot stabilize this kind of systems in the sense of H_∞ except in a simple case. For this case, thanks to the fractional PI controller given in [1], a parametrization of stabilizing controllers is derived [105].

6.9. Stabilization of MISO fractional systems with delays

Participants: Le Ha Vy Nguyen, Catherine Bonnet.

In order to yield the set of all the stabilizing controllers of a class of MISO fractional systems with delays by mean of Youla-Kucera parametrization regarding H_∞-stability, we are interested in determining coprime factorizations of the transfer function. Explicit expressions of left coprime factorizations and left Bézout factors have been derived in [103]. We have continued this year to search for explicit expressions of right coprime factorizations for some classes of systems [63].

6.10. Interval Observer

Participants: Frédéric Mazenc [correspondent], Thach Ngoc Dinh, Silviu Iulian Niculescu.

We made several progresses in the domain of the construction of state estimators called interval observers.

1) In [18], we have shown how interval observers can be constructed for nonlinear (and not Lipschitz) systems possessing a special triangular system.

2) The contributions [20] and [55] present a new major result for the design of interval observers for discete-time systems with input and output: it is explained how two classical Luenberger obsevers can be used, even in the absence of the positivity property as interval observer, provided two appropriate output, which compose the lower and the upper bound of the interval observer, are selected. In [19], coordinate transformations which change an arbitrary linear discrete-time system into a positive one and general nonlinear design of interval observers for nonlinear systems (satisfying a restricitive stability assumption) are proposed.

3) The paper [54] presents the first construction of continuous-discrete interval observer for linear continuous-time systems with discrete measurements. The importance in engineering applications of this result is clear: most of the time the measured variables are available at discrete instants only. The result relies on the design of changes of coordinates which transform a linear system into a nonnegative one, but the dynamic part of interval observers is not cooperative.

6.11. Reduction model approach: new advances

Participants: Frédéric Mazenc [correspondent], Michael Malisoff [Louisiana State University], Silviu Iulian Niculescu, Dorothé Normand-Cyrot [L2S, CNRS].

We solved several distinct problems entailing to the celebrated reduction model approach. Let us recall that this technique makes it possible to stabilize systems with arbitrarily large pointwise or distributed delay in the input.

1) We proposed in [25] a new construction of exponentially stabilizing sampled feedbacks for continuous-time linear time-invariant systems with an arbitrarily large constant pointwise delay in the inputs. Stability is guaranteed under an assumption on the size of the largest sampling interval. The proposed design is based on an adaptation of the reduction model approach. The stability of the closed loop systems is proved through a Lyapunov-Krasovskii functional of a new type, from which is derived a robustness result

2) The paper [59] presents several results pertaining to the stabilization with feedbacks given by an explicit formula of linear time varying systems in the case where there is a constant delay in the input. In addition, it establishes input-to-state stability with respect to additive uncertainties. As a particular case, we considered a large class of rapidly time varying systems and provided a lower bound on the admissible rapidness parameters. We illustrated our results using a pendulum model.

3) The paper [24], which is devoted to the original problem of stabilizing nonlinear systems with input with distributed delay, is actually not an extension of the reduction model approach, but it complements it and uses operators which have been inspired by those used in the classical context of the reduction model theory.

6.12. Neutral systems and integral equations

Participants: Frédéric Mazenc [correspondent], Hiroshi Ito [Kyushu Institute of Technology], Pierdomenico Pepe [Univ. of L'Aquila].

1) For nonlinear systems with delay of neutral type, we developped a new technique of stability and robustness analysis. It relies on the construction of functionals which make it possible to establish estimates of the solutions different from, but very similar to, estimates of ISS or iISS type. These functionals are themselves different from, but very similar to, ISS or iISS Lyapunov-Krasovskii functionals. The approach applies to systems which do not have a globally Lipschitz vector field and are not necessarily locally exponentially stable. We apply this technique to carry out a backstepping design of stabilizing control laws for a family of neutral nonlinear systems [21].

2) In a second paper [57], we extended the previous results to the problem of deriving stability and stabilizability conditions for nonlinear systems with delay interconnected with an integral equation via the construction of a Lyapunov-Krasovskii functional.

6.13. Nonlinear systems with delay

Participants: Frédéric Mazenc [correspondent], Michael Malisoff [Louisiana State University], Thach Ngoc Dinh.

We obtained new results on the robustness analysis of nonlinear systems belonging to a general family when they are globally stabilized by a state feedback corrupted by the presence of a delay and sampling [22], [58]. The result is based on the construction of a non-quadratic Lyapunov-Krasovskii functional.

In [23], a problem of state feedback stabilization of time-varying feedforward systems with a pointwise delay in the input is solved. The approach we adopted relies on a time-varying change of coordinates and Lyapunov-Krasovskii functionals. The result applies for any given constant delay, and provides uniformly globally asymptotically stabilizing controllers of arbitrarily small amplitude. The closed-loop systems enjoy input-to-state stability properties with respect to additive uncertainty on the controllers. The work was illustrated through a tracking problem for a model for high level formation flight of unmanned air vehicles.

6.14. Set theoretic fault detection and isolation

Participant: Sorin Olaru.

Fault-tolerant control theory is a well-studied topic but the use of the sets in detection, isolation and/or reconfiguration is rather tangential. Sorin Olaru together with his collaborators (and principally with F. Stoican) conducted a systematic analysis of the set-theoretic elements and devise approaches which exploit advanced elements within the field. The main idea is to translate fault detection and isolation conditions into those conditions involving sets. Furthermore, these are to be computed efficiently using positive invariance and reachability notions. Constraints imposed by exact fault control are used to define feasible references (which impose persistent excitation and, thus, non-convex feasible sets). Particular attention is given to the reciprocal influences between fault detection and isolation on the one hand, and control reconfiguration on the other. The recent results on this topic are gathered in the recent book [81].

A new result has been obtained by the use of controlled invariance for the separation of faulty/healthy invariant sets in the detection and isolation [32] based on the necessary and sufficent conditions of George Bitsoris.

In a series of recent papers [67], [68], [70], [69], the link between the interval observers and the invariant sets have been investigated by establishing a series of formal results on the limit behaviour with potential applications in the detection and isolation of actuators faults

6.15. Model Predictive Control: distributed formulations and collision avoidance problems

Participant: Sorin Olaru.

In [78], the mixt integer techniques have been analysed in the distributed model predictive control context, underlining the dependence of collision avoidance mechanism on the obstacle modeling and susequently on their treatement inside optimization-based control techniques as MPC (model predictive control). On the same topic of adversary consraints, a geometrical conditions has been establised in [71] for the local stabilization of a linear dynamics on a boundary of a forbidden region in the state space.

The theoretical developments from the last two years on the MPC design for multi-agent control problem led to the succesful application of receding horizon flight control for trajectory tracking of autonomous aerial vehicles [28]. In the same line or research, the predictive control for trajectory tracking and decentralized navigation of multi-agent formations has been presented in [29].

In [66] a Characterization of the Relative Positioning of Mobile Agents for Full Sensorial Coverage in an Augmented Space with Obstacles is presented in view of a MPC control design.

A predictive control-based algorithm for path following of autonomous aerial vehicles has been proposed in [65] to improve the previous trajectory tracking mechanism. The ultimate goal oof both schemes is to avoid the real-time infeasibility problems in MPC.

The distributed predictive control mechanisms have been used for the control of a four interconnected tanks benchmark [48], proving the versatility of a nonlinear Distributed MPC technique previously proposed by A. Grancharova.

In [62] the distributed Model Predictive Control of Leader-Follower Systems has been studied using an interior point method with efficient computations leading to simple tuning mechanisms for the cost functions and the terminal sets of the local MPC sub-problems.

6.16. Invariant sets in control

Participant: Sorin Olaru.

The longstanding research interest on the positive invaraince of a set with respect to the trajectories of the dynamical systems allowed recently the statement of explicit invariant approximation of the maximal robust positive invariant ser for LTI dynamics with zonotopic disturbances [51].

In the class of hybrid dynamical systems, explicit robustness and fragility margins for discrete-time linear systems with PWA control has been established in [64] by means of positive invariance arguments.

In [37] a series of new results on the linear constrained regulation problem have been presented by completing the classical results with the case of active constraints for the equilibrium point.

6.17. Optimization of mu-analysis parameterization

Participant: Guillaume Sandou [correspondent].

The robustness against parametric uncertainties can be studied using the structured singular value mu. In that case, a normalization of the uncertain parameters is performed, and the mu analysis provides the larger parallelepiped centered in the nominal and included in the stability domain. However, results depend on the initial normalization. In this study, the normalization is optimized so as to get the largest guaranteed stability domain. The corresponding problem being highly nonlinear, a metaheuristic method, Particle Swarm Optimization, is used for that purpose. An academic and a real life example, namely the pendulum in the cart problem, have been used to prove the viability of the approach.

6.18. Optimal weight tuning in Hinfinity loop-shaping with PSO considering time constraints

Participants: Guillaume Sandou [correspondent], Gilles Duc [Supélec, E3S], Philippe Feyel [Sagem].

Hinfinity loop-shaping controllers have proven their efficiency to solve problems based on complex industrial specifications. However, the tuning of the weighting filters is a time consuming task. This work deals with the use of metaheuristics optimization for this weighting filter tuning. Whereas this topic has already been investigated in lot of works, all of them assume a particular pole/zero/damping/pulse expression for the searched transfer function. But choosing the best weight structure is not trivial and may lead to suboptimal solutions for the design process. That is why, we propose to enhance the weight selection problem by relaxing the structure constraints of transfer functions. The developed methodology is tested, using a real industrial example and leads to satisfactory results.

6.19. mu-synthesis with dynamic D-Scalings using Quantum Particle Swarm Optimization

Participants: Guillaume Sandou [correspondent], Gilles Duc [Supélec, E3S], Philippe Feyel [Sagem].

This study proposes to revisit the mu-synthesis problem with a recent and efficient meta-heuristic called Quantum Particle Swarm Optimization (QPSO). This algorithm allows us to optimizing dynamics (or static) D-scalings without fitting them which leads to robust performance controllers. This method has been applied to an industrial problem and has been proven to be better than the classical D-K iteration method.

6.20. Stabilization of time-delay systems

Participants: Alban Quadrat, Arnaud Quadrat [SAGEM, MASSY].

In [118], [122], we study the stabilization problem of a linear system formed by a simple integrator and a time-delay system. We show that the stabilizing controllers of such a system can be rewritten as the closed-loop system defined by the stabilizing controllers of the simple integrator and a distributed delay system. This result is used to study tracking problems appearing in the study of inertially stabilized platforms for optical imaging systems. Moreover, an elementary proof for the parametrization [111] of all stabilizing controllers of a stabilizable plant – which does not necessarily admits doubly coprime factorizations – is given in [122].

6.21. A Stabilization problem in chemostats

Participants: Frédéric Mazenc [correspondent], Jérôme Harmand [LBE INRA, EPI MODEMIC].

We have considered the classical model of the chemostat (which is a bio-reactor) with one substrate, one species and a Haldane type growth rate function is considered. The input substrate concentration is supposed to be constant and the dilution rate is considered as the control. The problem of globally asymptotically stabilizing a positive equilibrium point of this system in the case where the measured concentrations are delayed and piecewise constant with a piecewise constant control is addressed. The result relies on the introduction of a dynamic extension of a new type. [56].

6.22. Control design for UAVs

Participants: Frédéric Mazenc [correspondent], Michael Malisoff [Louisiana State University].

In [14], we studied a kinematic model that is suitable for control design for high level formation flight of UAVs (Unmanned Aerial Vehicles). We designed controllers that give robust global tracking for a wide class of reference trajectories in the sense of the robustness notion called input-to-state stability. The control laws satisfy amplitude and rate constraints.

6.23. Modeling and control of Acute Myeloid Leukemia

Participants: José Luis Avila Alonso [correspondent], Annabelle Ballesta [BANG project-team], Frédéric Bonnans [COMMANDS project-team], Catherine Bonnet, Jean Clairambault [BANG project-team], Xavier Dupuis [COMMANDS project-team], Pierre Hirsch [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Jean-Pierre Marie [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Faten Merhi [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Silviu Iulian Niculescu, Hitay Özbay [Bilkent University, Ankara, Turkey], Ruoping Tang [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris].

In [75] we propose a new mathematical model of the cell dynamics in Acute Myeloid Leukemia (AML) which takes into account the four different phases of the proliferating compartment. The dynamics of the cell populations are governed by transport partial differential equations structured in age and by using the method of characteristics, we obtain that the dynamical system of equation can be reduced to two coupled nonlinear equations with four internal sub-systems involving distributed delays. Local stability conditions for a particular equilibrium point, corresponding to a positive cells, are derived in terms of a set of inequalities involving the parameters of the mathematical model. A parameter estimation of our model is being performed using biological data (Annabelle Ballesta).

We have also studied a coupled model for healthy and cancer cell dynamics in Acute Myeloid Leukemia consisting of two stages of maturation for cancer cells and three stages of maturation for healthy cells. The cell dynamics are modelled by nonlinear partial differential equations. The interconnection phenomenon between the healthy and cancer cells takes place on the re-introduction functions leaving the resting compartments to the proliferating compartments of both populations of cells at the first stage. For a particular healthy equilibrium point, locally stability conditions involving the parameters of the mathematical model are obtained [83], [84].

DOLPHIN Project-Team

6. New Results

6.1. Bi-level multi-objective optimization for pricing problems in long-haul transportation

Participants: M. Diaby, L. Brotcorne and E-G. Talbi

This work is concerned with the problem of pricing for a long-haul full load goods transportation. More precisely, we are interested in the situation where each vehicle delivers single request at a time. In this environment, we study the problem of pricing and valorization of unutilized capacity between two carriers. The first carrier B, cannot serve all the transportation requests and he thus needs to use outsourcing : second carrier A or his competitors. Carrier A, has to define the prices for carrier B transportation requests. Once carrier A has given its prices for the operations, it is B's decision to turn to A or to another carrier. This sequential and non-cooperative decision-making process can be adequately represented as a bilevel program : carrier B (the follower) wants to minimize transportation cost while A (the leader) seeks to maximize the revenue. Carrier A explicitly incorporates the reaction of carrier B in his optimization process.

Two types of models have been proposed : the bilevel mono-objective model and the bilevel biobjective model. More precisely, two objectives are simultaneously considered for the leader problem : the maximization of revenue and balancing the free load length (limiting the free load distances). We propose exact methods to solve moderate size instance of the problem and the heuristics to solve large-scale instances in reasonable time.

6.2. Approximating multi-objective scheduling problems

Participant: El-ghazali Talbi

External participants: Said Dabia, Tom Van Woensel, Tom De Kok (Eindhoeven Technical University)

In this contribution, we propose a generic approach to deal with multi-objective scheduling problems (MOSPs). The aim is to determine the set of Pareto solutions that represent the interactions between the different objectives. Due to the complexity of MOSPs, an efficient approximation based on dynamic programming is developed. The approximation has a provable worst case performance guarantee. Eventhough the approximate Pareto set consists of fewer solutions, it represents a good coverage of the true set of Pareto solutions. We consider generic cost parameters that depend on the state of the system. Numerical results are presented for the time-dependent multi-objective knapsack problem, showing the value of the approximation in the special case when the state of the system is expressed in terms of time [23].

6.3. Force-Based Cooperative Search Directions in Evolutionary Multi-objective Optimization

Participants: Bilel Derbel, Dimo Brockhoff, Arnaud Liefooghe

In order to approximate the set of Pareto optimal solutions, several evolutionary multi-objective optimization (EMO) algorithms transfer the multi-objective problem into several independent single-objective ones by means of scalarizing functions. The choice of the scalarizing functions' underlying search directions, however, is typically problem-dependent and therefore difficult if no information about the problem characteristics are known before the search process. In [46], we present new ideas of how these search directions can be computed *adaptively* during the search process in a *cooperative* manner. Based on the idea of Newton's law of universal gravitation, solutions attract and repel each other *in the objective space*. Several force-based EMO algorithms are proposed and compared experimentally on general bi-objective ρMNK landscapes with different objective correlations. It turns out that the new approach is easy to implement, fast, and competitive with respect to a $(\mu + \lambda)$-SMS-EMOA variant, in particular if the objectives show strong positive or negative correlations.

6.4. DYNAMO (DYNAmic programming using Metaheuristic for Optimization Problems)

Participants: Sophie Jacquin, Laetitia Jourdan, El-Ghazali Talbi

DYNAMOP (DYNAmic programming using Metaheuristic for Optimization Problems) is a new dynamic programming based on genetic algorithm to solve a hydro-scheduling problem. The representation which is based on a path in the graph of states of dynamic programming is adapted to dynamic structure of the problem and it allows to hybridize easily evolutionary algorithms with dynamic programming. DYNAMOP has been tested on two case studies of hydro-scheduling problem with different price scenarios. Experiments indicate that the proposed approach performs considerably better than classical genetic algorithms and dynamic programming.

6.5. MOCA-I: Multi-Objective Classification Algorithm for Imbalanced Data

Participants: Julie Jacques, Clarisse Dhaenens, Laetitia Jourdan

Dealing with Imbalanced data is a real challenge as predicting the minority class may be very difficult but has a great interest for medical applications for example. Therefore, we propose MOCA-I, a new multi-objective local search algorithm that is conceived to deal with class imbalancy, double meaning of missing data, volumetry and need of highly interpretable results all together [50]. MOCAI is a Pittsburgh multi-objective partial classification rule mining algorithm, using dominance-based multi-objective local search (DMLS). In comparison to state-of-the-art classification algorithms, MOCA-I obtains the best results on the 10 data sets of literature and is statistically better on the real data sets [50].

6.6. Neutrality Analysis is Graph coloring problem

Participants: Aymeric Blot, Clarisse Dhaenens, Laetitia Jourdan, Marie-Eleonore Marmion

Solving neutral problems is challenging as many optimization methods have difficulty to obtain good solutions. Hence, studying the neutrality in order to provide insights on the structure of the problem to be soved may be an answer. This has been done for the graph coloring problem (GCP) for which the neutrality of some hard instances has been quantified. This neutrality property has to be detected as it impacts the search process. Indeed, local optima may belong to plateaus that represent a barrier for local search methods. Then, we also aim at pointing out the interest of exploiting neutrality during the search. Therefore, a generic local search dedicated to neutral problems, NILS, is performed on several hard instances [78].

6.7. Neutrality in Multi-objective Local Search

Participants: Aymeric Blot, Clarisse Dhaenens, Laetitia Jourdan
External Participants: Hernan Aguirre, Kiyoshi Tanaka - Shinshu University, Japan

In multi-objective combinatorial optimization, the dominance-based local search algorithms are faced to sets of non-comparable solutions. In the absence of preferences, these solutions are equally good from the Pareto dominance perspective and can be considered neutral in term of quality, similar to the solutions who shares the same fitness value in mono-objective optimization. We propose two ideas to use the neutrality to improve the current local search algorithms. First, we analyze the distribution of neighbors for both small fully enumerable instances and hard large instances, to understand the distribution of neutral neighbors according to the rank of the solutions. Then, we compare the results of the proposed algorithms with the standard ones according to different indicators.

6.8. Biclustering for GWA data

Participants: Khedidja Seridi, Laetitia Jourdan, El-Ghazali Talbi

We have examined the possibilities of applying biclustering approaches to detect association between SNP markers and phenotype traits. There- fore, we have proposed a multiobjective model for biclustering problems in GWA context. Furthermore, we have proposed an adapted heuristic and meta- heuristic to solve it. The good performances of our algorithms are assessed using synthetic data sets.

6.9. Fitness Landscape Analysis for Multiobjective Optimization

Participant: Arnaud Liefooghe
External participants: Hernan Aguirre, Kiyoshi Tanaka (Shinshu Univ., Japan), Sébastien Verel (Univ. Littoral Côte d'Opale, France)

In [57], we investigate the correlation between the characteristics extracted from the problem instance and the performance of a simple evolutionary multiobjective optimization algorithm. First, a number of features are identified and measured on a large set of enumerable multiobjective NK-landscapes with objective correlation. A correlation analysis is conducted between those attributes, including low-level features extracted from the problem input data as well as high-level features extracted from the Pareto set, the Pareto graph and the fitness landscape. Second, we experimentally analyze the (estimated) running time of the global SEMO algorithm to identify a $(1+\epsilon)$-approximation of the Pareto set. By putting this performance measure in relation with problem instance features, we are able to explain the difficulties encountered by the algorithm with respect to the main instance characteristics.

In [38], we study the effects of population size on selection and performance scalability of two dominance-based algorithms applied to many-objective optimization. Our aim is to understand the relationship between the size of the Pareto optimal set, a characteristic of the many-objective problem at hand, the population size and the ability of the algorithm to retain Pareto optimal solutions in its population and find new ones. This work clarifies important issues of the dynamics of evolutionary algorithms on many-objective landscapes, particularly related to survival selection. It shows that optimal solutions are dropped from the population in favor of suboptimal solutions that appear non-dominated when survival selection is applied. It also shows that this selection lapse, the dropping of optimal solution, affects the discovery of new optimal solutions and is correlated to population size and the distribution of solutions that survival selection renders. Selection makes less mistakes with larger populations and when the distribution of solutions is better controlled. The results of this study will be helpful to properly set population size and have a clearer idea about the performance expectation of the algorithm.

6.10. On Set-based Local Search for Multiobjective Combinatorial Optimization

Participant from DOLPHIN: Arnaud Liefooghe
External participants: Matthieu Basseur, Adrien Goëffon (Univ. Angers, France), Sébastien Verel (Univ. Littoral Côte d'Opale, France)

In [42], we formalize a multiobjective local search paradigm by combining set-based multiobjective optimization and neighborhood-based search principles. Approximating the Pareto set of a multiobjective optimization problem has been recently defined as a set problem, in which the search space is made of all feasible solution-sets. We here introduce a general set-based local search algorithm, explicitly based on a set-domain search space, evaluation function, and neighborhood relation. Different classes of set-domain neighborhood structures are proposed, each one leading to a different set-based local search variant. The corresponding methodology generalizes and unifies a large number of existing approaches for multiobjective optimization. Preliminary experiments on multiobjective NK-landscapes with objective correlation validates the ability of the set-based local search principles. Moreover, our investigations shed the light to further research on the efficient exploration of large-size set-domain neighborhood structures.

6.11. Feature selection in high dimensional regression problems for genomics

Participants: Julie Hamon, Clarisse Dhaenens (External collaborator : Julien Jacques)

In the context of genomic selection in animal breeding, an important objective consists in looking for explicative markers for a phe-notype under study. In order to deal with a high number of markers, we propose to use combinatorial optimization to perform variable selection. Results show that our approach outperforms some classical and widely used methods on simulated and "closed to real" datasets [76]. Familial relationships have also been used in this specific context and allow to improve results.

6.12. Indicator-Based Multiobjective Optimization

Participant: Dimo Brockhoff
External Participants: Johannes Bader (formerly at ETH Zurich, Switzerland), Youssef Hamadi (Microsoft Research, Cambridge, UK), Souhila Kaci (Université Montpellier 2, France), Lothar Thiele (ETH Zurich, Switzerland), Heike Trautmann (University of Munster, Germany) Tobias Wagner (TU Dortmund, Germany), and Eckart Zitzler (PH Bern, Switzerland)

Indicator-based (evolutionary) multiobjective optimization algorithms have been first introduced in 2004 and typically use a quality indicator, assigning a solution set a real value, as a direct, internal performance criterion. Given that the indicator and the number μ of desired points is fixed, the optimization goal, also denoted by the term *optimal μ-distribution*, is then defined as the solution set(s) of size μ which maximizes the indicator value.

In 2013, we continued to investigate, theoretically and numerically, the optimal μ-distributions for the R2 indicator, an often recommended indicator based on scalarization functions [73]. We also proposed a new multiobjective optimizer with an R2-indicator-based selection [70]. With respect to the even more common *hypervolume indicator*, we combined the idea of the weighted hypervolume indicator with the idea of interactive algorithms and proposed a new algorithm that adapts the weighted hypervolume's weight function according to the user's preferences during the search. Last, we summarized our knowledge on the weighted hypervolume indicator and proposed a general framework of how to employ it within the hypervolume-based W-HypE algorithm [18].

6.13. A Hybrid Metaheuristic for Multiobjective Unconstrained Binary Quadratic Programming

Participant : Arnaud Liefooghe
External participants : Jin-Kao Hao (Univ. Angers, France), Sébastien Verel (Univ. Littoral Côte d'Opale, France)

The conventional Unconstrained Binary Quadratic Programming (UBQP) problem is known to be a unified modeling and solution framework for many combinatorial optimization problems. In [29], we extend the single-objective UBQP to the multiobjective case (mUBQP) where multiple objectives are to be optimized simultaneously. We propose a hybrid metaheuristic which combines an elitist evolutionary multiobjective optimization algorithm and a state-of-the-art single-objective tabu search procedure by using an achievement scalarizing function. Finally, we define a formal model to generate mUBQP instances and validate the performance of the proposed approach in obtaining competitive results on large-size mUBQP instances with two and three objectives.

6.14. Multi-core GPU-based parallel optimization

We have mainly investigated the design and implementation on multi-core GPU-based platforms of meta-heuristics and tree-based exact optimization methods focusing on Branch and bound (B&B) algorithms (Ph.D thesis of I. Chakroun).

- **GPU-based Metaheuristics**
  Participants: N. Melab, T-V. Luong, K. Boufaras and N. Melab

We came out with a pioneering work on single-solution methods. The hierarchy of parallel models has been rethought on GPU dealing with CPU-GPU data transfer optimization, thread control and automatic mapping of candidate solutions to threads. The implementation of the proposed approaches is provided through ParadisEO-GPU in [62] (nominated for Best Paper Award). High speedups ups have been achieved for some problems.

- **Multi-core GPU-based B&B algorithms**
  For exact optimization, we have revisited the design and implementation of highly irregular B&B algorithms on GPU dealing with hierarchical device memory optimization, on GPU combined with multi-core [45] dealing with CPU-GPU data transfer optimization and work partitioning, and on GPU-enhanced computational grids. Accelerations up to $\times 217$ are achieved on Tesla Nvidia C2050 on large Flow-Shop problems.

6.15. Energy-aware scheduling for clouds

Participants: Y. Kessaci, N. Melab, E-G. Talbi

High-performance computing (HPC) is moving from in-house to cloud-based HPC. One of the major issues of this later is the scheduling of HPC applications taking into account the energy criterion in addition to performance. In [54], we have addressed that issue (Ph.D thesis of Y. Kessaci). We have proposed several metaheuristics for cloud managers and experimented on OpenNebula using different (VMs) arrival scenarii and different hardware infrastructures. The results show that our approaches outperform the scheduler provided in OpenNebula by a significant margin in terms of energy consumption and number of scheduled VMs.

6.16. Heterogenous Multi-CPU Multi-GPU Parallel Branch-and-Bound Tree Search

Participants: Trong-Tuan Vu, Bilel Derbel, Nouredine Melab

In this work [71], we push forward the design of parallel and distributed optimization algorithms running on heterogenous systems consisting of multiple CPUs coupled with multiple GPUs. We consider parallel Branch-and-Bound (B&B), viewed as a generic algorithm searching in a dynamic tree representing a set of candidate solutions built dynamically at runtime. Given that several distributed CPUs and GPUs coming from possibly different clusters connected through a network can be used to parallelize the tree search, we give new insights into how to fully benefit from such a heterogeneous environment. More precisely, we describe a two-level generic and fully distributed parallel approach taking into account PU characteristics. In the first level, we use data streaming in order to allow parallelism between hosts and devices. The evaluation of tree nodes is done inside a GPU while the CPU-host is performing the pruning, selection and decomposition operations in parallel. In the second level, our approach incorporates an adaptive dynamic load balancing scheme based on distributed work stealing, in order to flow workloads efficiently from overloaded PUs to idle ones at runtime. We deployed our approach over a distributed system of up to 20 GPUs and 128 CPUs coming from three clusters. Different scales and configurations of PUs were experimented with the B&B algorithm and the well-known FlowShop combinatorial optimization problem as a case study. Firstly, on one single GPU, we improve on the running time of previous B&B GPUs implementation by at least a factor of two. More importantly, independently of CPUs or GPUs scale or power, our approach provides a substantial speed-up which is *nearly optimal* compared to the ideal performance one could expect in theory.

DRACULA Project-Team

6. New Results

6.1. Mathematical modeling in chronobiology

Circadian clocks are autonomous oscillators entrained by external Zeitgebers such as light-dark and temperature cycles. On the cellular level, rhythms are generated by negative transcriptional feedback loops. In mammals, the suprachiasmatic necleus (SCN) in the anterior part of the hypothalamus plays the role of the central circadian pacemaker. Coupling between individual neurons in the SCN leads to specify self-sustained oscillations even in the absence of external signals. These neuronal rhythms orchestrate the phasing of circadian oscillations in peripheral organs. Altogether, the mammalian circadian system can be regarded as a network of coupled oscillators. In order to understand the dynamic complexity of these rhythms, mathematical models successfully complement experimental investigations. In [19], we discuss basic ideas of modeling on three different levels: (i) rhythm generation in single cells by delayed negative feedbacks, (ii) synchronization of cells via external stimuli or cell-cell coupling, and (iii) optimization of chronotherapy.

6.2. Hybrid Models of Cell Population

The paper [20] is devoted to hybrid discrete-continuous models of cell populations dynamics. Cells are considered as individual objects which can divide, die by apoptosis, differentiate and move under external forces. Intra-cellular regulatory networks are described by ordinary differential equations while extracellular species by partial differential equations. We illustrate the application of this approach to some model examples and to the problem of tumor growth. Hybrid models of cell populations present an interesting nonlinear dynamics which is not observed for the conventional continuous models.

6.3. Multiscale Model in Biology

Biological processes span several scales in space, from the single molecules to organisms and ecosystems. Multiscale modelling approaches in biology are useful to take into account the complex interactions between different organisation levels in those systems. We present in [6] several single- and multiscale models, from the most simple to the complex ones, and discuss their properties from a multiscale point of view. Approaches based on master equations for stochastic processes, individual-based models, hybrid continuous-discrete models and structured PDE models are presented.

6.4. Model of hematopoiesis

In [2], a model of blood cell production in the bone marrow (hematopoiesis), has been investigated. It describes both the evolution of primitive hematopoietic stem cells and the maturation of these cells as they differentiate to form the three kinds of progenitors and mature blood cells (red blood cells, white cells and platelets). The three types of progenitors and mature cells are coupled to each other via their common origin in primitive hematopoietic stem cells compartment. The resulting system is composed by eleven age-structured partial differential equations. To analyze this model, we don't take into account cell age-dependence of coefficients, that prevents a usual reduction of the structured system to an unstructured delay differential system. We study the existence of stationary solutions: trivial, axial and positive steady states. Then we give conditions for the local asymptotic stability of the trivial steady state and by using a Lyapunov function, we obtain a sufficient condition for its global asymptotic stability. In some particular cases, we analyze the local asymptotic stability of the positive steady state by using the characteristic equation. Finally, by numerical simulations, we illustrate our results and we show that a change in the duration of cell cycle can cause oscillations.

6.5. Analysis of radiocarbon to facilitate identification of unknown decedents

The characterization of unidentified bodies or suspected human remains is a frequent and important task for forensic investigators. However, any identification method requires clues to the person's identity to allow for comparisons with missing persons. If such clues are lacking, information about the year of birth, sex and geographic origin of the victim, is particularly helpful to aid in the identification casework and limit the search for possible matches. We present in [4] results of stable isotope analysis of $(13)C$ and $(18)O$, and bomb-pulse $(14)C$ analyses that can help in the casework. The $(14)C$ analysis of enamel provided information of the year of birth with an average absolute error of 1.8 ± 1.3 years. We also found that analysis of enamel and root from the same tooth can be used to determine if the $(14)C$ values match the rising or falling part of the bomb-curve. Enamel laydown times can be used to estimate the date of birth of individuals, but here we show that this detour is unnecessary when using a large set of crude $(14)C$ data of tooth enamel as a reference. The levels of $(13)C$ in tooth enamel were higher in North America than in teeth from Europe and Asia, and Mexican teeth showed even higher levels than those from USA. DNA analysis was performed on 28 teeth, and provided individual-specific profiles in most cases and sex determination in all cases. In conclusion, these analyses can dramatically limit the number of possible matches and hence facilitate person identification work.

DREAM Project-Team

6. New Results

6.1. Diagnosis of large scale discrete event systems

Participants: Marie-Odile Cordier, Christine Largouët, Sophie Robin, Laurence Rozé, Yulong Zhao.

The problem we deal with is monitoring complex and large discrete-event systems (DES) such as an orchestration of web services or a fleet of mobile phones. Two approaches have been studied in our research group. The first one consists in representing the system model as a discrete-event system by an automaton. In this case, the diagnostic task consists in determining the trajectories (a sequence of states and events) compatible with the sequence of observations. From these trajectories, it is then easy to determine (identify and localize) the possible faults. In the second approach, the model consists in a set of predefined characteristic patterns. We use temporal patterns, called chronicles, represented by a set of temporally constrained events. The diagnostic task consists in recognizing these patterns by analyzing the flow of observed events.

6.1.1. Distributed monitoring with chronicles - Interleaving diagnosis and repair - Making web services more adaptive

Our work addresses the problem of maintaining the quality of service (QoS) of an orchestration of Web services (WS), which can be affected by exogenous events (i.e., faults). The main challenge in dealing with this problem is that typically the service where a failure is detected is not the one where a fault has occurred: faults have cascade effects on the whole orchestration of services. We have proposed a novel methodology to treat the problem that is not based on Web service (re)composition, but on an adaptive re-execution of the original orchestration. The re-execution process is driven by an orchestrator Manager that takes advantage of an abstract representation of the whole orchestration and may call a diagnostic module to localize the source of the detected failure. It is in charge of deciding the service activities whose results can be reused and may be skipped, and those that must be re-executed.

This year, we have improved the prototype, adding the visualization of the roadmap and the activities that do not have to be reexecuted. This work has been published in ICWS2013 [15] and we are working on a journal paper that will be submitted in 2014.

6.1.2. Scenario patterns for exploring qualitative ecosystems

This work aims at giving means of exploring complex systems, in our case ecosystems, and more recently agrosystems, specifically herd management systems. We proposed to transform environmental questions about future evolution of ecosystems into formalized queries that can be submitted to a simulation model. The system behavior is represented as a discrete event system described by a set of interacting timed automata, the global model corresponding to their composition on shared events. To query the model, we have defined high-level generic query patterns associated to the most usual types of request scenarios. These patterns are then translated into temporal logic formulas. The answer is computed thanks to model-checking techniques that are efficient for analyzing large-scale systems. Five generic patterns have been defined using TCTL (Timed Computation Tree Logic) “WhichStates”, “WhichDate”, “Stability”, “Always”, “Safety”. Three of them have been implemented using the model-checker UPPAAL.

The approach has first been experimented on a marine ecosystem under fishing pressure. The model describes the trophodynamic interactions between fish trophic groups as well as interactions with the fishery activities and with an environmental context. A paper has been previously published in the Environmental Modelling Software Journal [65]. More recently, a similar approach has been experimented on agrosystems, specifically herd management systems, for which a hybrid model has been built using hierarchical timed automata. This later work has been achieved in the context of Yulong Zhao’s PhD thesis [6] and done in collaboration with our colleagues of INRA.

6.1.3. Controler synthesis for dealing with “How to” queries

We extended the approach to deal with “How to” queries. As before, we rely on a qualitative model in the form of timed automata and on model-checking tools to answer queries. We proposed and compared two approaches to answer questions such as “How to avoid a given situation ?”(safety query). The first one exploits controller synthesis and the second one is a “generate and test” approach. We evaluated these two approaches in the context of an application that motivates this work, i.e. the management of a marine ecosystem and the evaluation of fishery management policies. The results have beenpreviously published in [88].

More recently, we used similar methodological tools to analyze in the context of herd management on a catchment. An hybrid model has been built using hierarchical timed automata and scenarios can be simulated and evaluated using the approach presented in the previous paragraph. In this context, the goal is to identify and analyse the best/optimal farming practices in order to reduce nitrate pollution due to livestock effluents. We proposed to use controler synthesis tools and to couple them with machine learning techniques in order to get the best strategies and to put them on easy-to-use form. This work has been made in the context of Yulong Zhao’s PhD thesis [6] and in collaboration with our colleagues of INRA (UMR PEGASE).

6.2. Machine learning for model acquisition

Participants: Sid Ahmed Benabderrahmane, Marie-Odile Cordier, Thomas Guyet, Simon Malinowski, René Quiniou.

Model acquisition is an important issue for model-based diagnosis, especially while modeling dynamic systems. We investigate machine learning methods for temporal data recorded by sensors or spatial data resulting from simulation processes. Our main objective is to extract knowledge, especially sequential and temporal patterns or prediction rules, from static or dynamic data (data streams). We are particularly interested in mining temporal patterns with numerical information and in incremental mining from sequences recorded by sensors.

6.2.1. Representing and mining time series

Time series are sequences of numerical values, e.g. recorded by sensors. Since these series can be huge and subject to noise, they are often transformed into sequences of symbols. The best known symbolic transformation method is SAX (Symbolic Aggregate approXimation) [68]. SAX is based on a piecewise constant approximation method that does not take into account the slope of the time series values in successive windows. We have extended the SAX method by adding a symbolic slope information to the SAX symbols. We have experimented our new representation, 1d-SAX, on three mining tasks. In most of these experiments 1d-SAX leads to a better accuracy than SAX [19].

We have also investigated a probabilistic representation of temporal patterns based on the latent Dirichlet allocation model (LDA). Such patterns can approximate the dynamics of a set of similar multivariate time series. We have experimented the method on hydrological flood time series to extract temporal patterns [7]. The extracted patterns were considered relevant and easy to understand by experts of the domain.

6.2.2. Incremental sequential mining

Sequential pattern mining algorithms operating on data streams generally compile a summary of the data seen so far from which they compute the set of actual sequential patterns. We propose another solution where the set of actual sequential patterns are incrementally updated as soon as new data arrive on the input stream. Our work stands in the framework of mining an infinite unique sequence. Our method [60] provides an algorithm that maintains a tree representation (inspired by the PSP algorithm [71]) of frequent sequential patterns and their minimal occurrences [69] in a window that slides along the input data stream. It makes use of two operations: deletion of the itemset at the beginning of the window (obsolete data) and addition of an itemset at the end of the window (new data). The experiments were conducted on simulated data and on real data of instantaneous power consumption. The results show that our incremental algorithm significantly improves the computation time compared to a non-incremental approach [61].

Recently, we have worked on the adaptation of our algorithm to closed sequential patterns. A closed pattern is a local maximal pattern such there exists no extension of this pattern having the same support. Closed patterns are known to provide a condensed represention of the solution patterns and lead to more efficient algorithms without losing information or completeness on extracted patterns. The tree of closed-patterns is less deep than the pattern-tree but the transformations of the tree by addition or deletion of items are more complex. The algorithm is under evaluation. We plan to submit a paper in 2014.

6.2.3. Multiscale segmentation of satellite image time series

Satellite images allow the acquisition of large-scale ground vegetation. Images are available along several years with a high acquisition frequency (1 image every two weeks). Such data are called satellite image time series (SITS). In [58] , we presented a method to segment an image through the characterization of the evolution of a vegetation index (NDVI) on two scales: annual and multi-year. The main issue of this approach was the required computation resources (time and memory). We first propose to adapt image segmentation algorithm to SITS. Segmented images reduces the number of time series to analyze and the computation time. We secondly applied 1D-SAX to reduce data dimensionality [20]. We evaluated this approach on the supervised classification of large SITS of Senegal and we showed that 1D-SAX approaches the classification results of time series while significantly reducing the required memory storage of the images.

6.2.4. Analysis of landscape based on spatial patterns

Researchers in agro-environment need a great variety of landscapes to test the agro-ecological models of their scientific hypotheses. Real landscapes are difficult to acquire and do not enable the agronomist to test all their hypothesis. Working with simulated landscapes is then an alternative to get a sufficient variety of experimental data. Our objective is to develop an original scheme to generate realistic landscapes. This approach is based on a spatial representation of landscapes by a graph expressing the spatial relationships between the agricultural parcels (as well as the roads, the rivers, the buildings, etc.), of a specific geographic area. We extract spatial patterns from a real geographic area and we use these patterns to generate new realistic landscapes. Using patterns preserves the interface properties between parcels.

We have begun the exploration of graph mining techniques, such as gSPAN [87], to discover the relevant spatial patterns present in a spatial-graph. But the graph-mining techniques are very time-consuming in comparison to sequence mining.

This year, we would like to test if using a path instead of a graph would be a faithful representation of the spatial organization of the landscape. In [17], we compare the potential expressivity of graphs and Hilbert-Peano curves [66] to characterize an agricultural landscape. The results show that mining frequent patterns in Hilbert-Peano curves would be as discriminant as mining frequent patterns in graphs.

The perception of the environment is an important dimension of the landscape we live in. One of our objectives is to study the relationships between the landscape patterns and their perception. We cope with this dimension by analysing the textual content of ‘atlas du paysage" (landscape atlas), that are produce by each french administrative regions. This year we worked on the construction of an ontology of landscape perception [21].

6.2.5. Subdimensional clustering for fast similarity search over time series data. Application to Information retrieval tasks

Information retrieval and similarity search tasks in time series databases remains a challenge that require to discover relevant pattern-sequences that are recurrent over the overall time series sequences, and to find temporal associations among these frequently occurring patterns. Previous work on information retrieval and similarity search in time series has been performed in different contexts such as diagnosis or failure detection of industrial materials. In whole query matching, a time series given as query is entirely compared to every time series of a database. The series should have same length, and a similarity measure is used to retrieve either a most similar time series or the top-k ranked time series. However, theses methods suffer from a lack of flexibility of the used similarity measures, a lack of scalability of the representation model, and a penalizing runtime to retrieve the information. Moreover, in some real world applications, one can be interested in retrieving specific interesting subsequences that are frequently present at different instants.

Motivated by these observations, we have designed a framework tackling the query by content problem on time series data, ensuring (i) fast response time, (ii) multi-level information representation, and (iii) representing temporal associations between extracted patterns. During the preparation step, all the multi-valued time series present in the database are transformed into a multi-resolution symbolic representation thus ensuring a lower dimensionality. Then, to accelerate and enhance the similarity search and the retrieval over the database, our model creates an index over recurrent patterns in the time series collection. These patterns can be generated by different techniques. Finally, the extracted patterns are grouped by clustering and the resulting clusters are indexed in a table within their centroids. A paper presenting the preliminary results is under submission to an international journal.

6.2.6. Knowledge Extraction from Heterogeneous Data

Recently, mining microarrays data has became a big challenge due to the growing sources of available data. We are using machine learning methods such as clustering, dimensionality reduction, association rules discovery on transcriptomic data, by combining a domain ontology as source of knowledge, in order to supervise the KDD process. Our objectives concern the identification of genes that could participate in the development of tumors. A two-way classification method was proposed, combining genes expression levels, represented as numerical data, and Gene Ontology (GO) annotations as symbolic data. The hopeful results obtained with genes clustering, through GO annotations, are an encouraging track to predict transcriptional regulatory networks, and for refining the existing sets of genes [11], [12].
We also introduced a new method for extracting enriched biological functions from transcriptomic databases using an integrative bi-classication approach. The initial gene datasets are firstly represented as a formal context (objects attributes), where objects are genes, and attributes are their expression profiles and complementary information of different knowledge bases. Formal Concept Analysis (FCA) is applied for extracting formal concepts regrouping genes having similar transcriptomic profiles and functional behaviors. An enrichment analysis is then performed in order to identify the relevant formal concepts from the generated Galois lattice, and to extract biological functions that could participate in the proliferation of cancers. Preliminary results seem very promising, and could help experts during the identification of degenerated biological functions [13].

6.3. Decision aiding with models and simulation data

Participants: Louis Bonneau de Beaufort, Tassadit Bouadi, Marie-Odile Cordier, Véronique Masson, René Quiniou.

Models can be very useful for decision aiding as they can be used to play different plausible scenarios for generating the data representing future states of the modeled process. However, the volume of simulation data may be huge. Thus, efficient tools must be investigated in order to store the simulation data, to focus on relevant parts of the data and to extract interesting knowledge from these data.

6.3.1. A datawarehouse for simulation data

The ACASSYA project 8.2.1 aims at providing experts or stakeholders or farmers with a tool to evaluate the impact of agricultural practices on water quality. As the simulations of the deep model TNT2 are time-consuming and generate huge data, we have proposed to store these simulation results in a datawarehouse and to extract relevant information, such as prediction rules, from the stored data. We have devised a general architecture for agro-environmental data on top of the framework Pentaho.

This year we have been working on the efficient computation of OLAP queries related to realistic scenarios proposed by experts in the domain. Precisely, we have devised indexing schemes to access the data in the OLAP cube. We have also worked on the visualization by a GIS (Geographical Information System) of the query results on maps of the geographical area under interest. A paper have been submitted to the COMPAG Journal. This work is detailed in Tassadit Bouadi's thesis [5].

6.3.2. Efficient computation of skyline queries in an interactive context

Skyline queries retrieve from a database the objects that maximize some criteria, related to user preferences for example, or objects that are the best compromises satisfying these criteria. When data are in huge volumes, such objects may shed light on interesting parts of the dataset. However, computing the skylines (i.e. retrieving the skyline points) may be time consuming because of many dominance tests. This is, especially the case in an interactive setting such as querying a data cube in the context of a datawarehouse.

We have worked at improving the formal setting of the partial materialization of skyline queries when dynamic preferences are refined online by the user. We have explicited which parts of the skyline evolve (which point are added or removed) when a new dimension is introduced in the computation. This led to an efficient incremental method for the online computation of the skyline corresponding to new user preferences [46]. An extended version of this paper is published in Journal "Transactions on Large Scale Data and Knowledge Centered Systems" (TLDKS) [8] and in Tassadit Bouadi's thesis [5].

6.3.3. Hierarchical skylines

Conventional skyline queries retrieve the skyline points in a context of dimensions with a single hierarchical level. However, in some applications with multidimensional and hierarchical data structure (e.g. data warehouses), skyline points may be associated with dimensions having multiple hierarchical levels. Thus, we have proposed an efficient approach reproducing the effect of the OLAP operators "drill-down" and "roll-up" on the computation of skyline queries. It allows the user to navigate along the dimensions hierarchies (i.e. specialize / generalize) while ensuring an online calculation of the associated skyline. The method is described in Tassadit Bouadi's thesis [5]. A paper describing this contribution is currently under submission to the "Very Large Data Bases (VLDB 2014)" conference.

6.3.4. Modeling influence propagation by Bayesian causal maps

The goal of this project is modeling shellfish fishing to assess the impact of management pollution scenarios on the *Rade de Brest*. Cognitive maps were built from interviews with fishermen. To represent and reason about these cognitive maps, we propose to use Bayesian Causal Maps making use of fishermen knowledge, particularly to perform influence propagation [82].

However, this model does not take into account the variety of influences asserted by the fishermen, but only the "mean" causal map. A report describing the project is available [28]. An approach that could combine individual knowledge with belief functions in the way of Philippe Smets's Transferable Belief Model [83] has been proposed. A report describing the projectis available [28].

This work is done in the framework of the RADE2BREST project, involving Agrocampus Ouest and CNRS (GEOMER/LETG), funded by "Ministère de l'Ecologie" (This project is not mentioned in section 8.2 because DREAM is not an official partner of this project.).

6.3.5. Recommending actions from classification rules

In the framework of the SACADEAU project, a paper dedicated to building actions from classification rules has been published in the KAIS Journal [9]. Our goal is to burden of analysing a large set of classification rules when the user is confronted to an "unsatisfactory situation" and needs help to decide about the appropriate actions to remedy to this situation. The method consists in comparing the situation to a set of classification rules. For this purpose, we propose DAKAR, a new framework for learning action recommendations dealing with complex notion of feasibility and quality of actions.

Sacadeau-Software, which is the decision support tool implemented with F. Ployette (former Inria engineer in the EPI Dream, now retired) in the SACADEAU project, has been published in the RIA Journal [10]. Sacadeau-Software allows to run simulations throughout a watershed and obtain the transfer rate of pollution through the catchment. Classification rules, characterizing the sub-parts of the watershed with pollution and the sub-parts without pollution, are automatically learned from the simulations. A visualization tool enables to relate the learned rules to the examples characterized by these rules. Finally, a user can select a situation of pollution and the action recommendation tool analyses the learned rules and proposes actions that improve this situation of pollution.

6.4. Diagnostic, causal reasoning and argumentation

Participants: Philippe Besnard, Marie-Odile Cordier, Yves Moinard.

Stemming on [38], [39], [40], [41], [42], we have designed an inference system based on causal statements. This is related to diagnosis (observed symptoms explained by faults). The aim is to produce possible explanations for some observed facts. Previously existing proposals were ad-hoc or, as in [45], [57], they were too close to standard logic to make a satisfactory diagnosis. A key issue for this kind of work is to distinguish logical implication from causal links and from ontological links. This is done by introducing a simple causal operator, and an *is-a* hierarchy. These two operators are added to a restricted first order logic of the Datalog kind (no function symbols). Then, our system produces elementary *explanations* for some set of observed facts. Each explanation links some facts to the considered observation, together with a set of atoms called the *justifications*: The observation is explained from these facts, provided the justifications are possible (not contradicted by the available data). This formalism has been translated into answer set programming [72], [73]). It is able to deal with complex problems such as finding explanations for the hurricane Xynthia (2010, February 28). In such situations, there are many data and many possible elementary explanations can be examined. This involves an extension of our formalism, in order to deal with more complex chains of causations and *is-A* links. Our formalism makes precise what all these possible explanations are. Then, in order to deal with so many possible complex explanations, we integrate this causal formalism into an argumentation framework. Logic-based formalizations of argumentation [43] take pros and cons for some conclusion into account. These formalizations assume a set of formulae and then exhaustively lay out arguments and counterarguments. This involves providing an initiating argument for the inference and then providing undercuts to this argument, and then undercuts to undercuts. So here our causal formalism provides a (rather large) set of explanations, and the argumentation part allows to select the best ones, under various criteria [22], [14].

Then, since answer set programming can easily deal with logical formalisms, the argumentation part will be incorporated into our already existing answers set programming translation of the causal formalism. Regarding this field of knowledge representation and reasoning, and more generally, artificial intelligence, we have participated to several chapters in the to be published "Panorama de l'intelligence artificielle. Ses bases méthodologiques, ses développements" [27], [26], [23], [24].

DREAMPAL Team

6. New Results

6.1. Language-Independent Symbolic Execution, Program Equivalence, and Program Verification

A significant part of our research project consists in applying formal techniques for symbolically executing and formally verifying HiHope programs, as well as for formally proving the equivalence of HiHope programs with the corresponding HoMade assembly and machine-code programs obtained by compilation of HiHope.

- Symbolic execution will detect bugs (e.g., stack undeflow) in HiHope programs. Additionaly, symbolic execution is the natural execution manner of HiHope programs as soon as they contain (typically, underspecified) hardware IPs;
- program verification will guarantee the absence of bugs (with respect to specified properties, e.g., no stack underflow, no invocation of unavailable IPs, ...);
- program equivalence will guarantee that such above-mentioned bugs are also absent from the HoMade assembly and machine-code programs obtained by compilation of HiHope source code.

Since these languages (especially HiHope) are not completely defined yet, we decided to work (together with our colleagues from Univ. Iasi, Romania) on language-independent symbolic execution, program-equivalence, and program-verification techniques. In this way, when all the languages in our project become stable, we will be readily able to instantiate the above generic techniques on (the K formal definitions of) the languages in question. We note that all the techniques described below are also independent of K: they are applicable to other language-definition frameworks that use similar rewriting-based formal operational semantics.

6.1.1. Symbolic Execution

In [9] we propose a language-independent symbolic execution framework for languages endowed with a formal operational semantics based on term rewriting. Starting from a given definition of a language, a new language definition is automatically generated, which has the same syntax as the original one but whose semantics extends data domains with symbolic values and adapts semantical rules to deal with these values. Then, the symbolic execution of concrete programs is, by definition, the execution of programs with the new symbolic semantics, on symbolic input data. We prove that the symbolic execution thus defined has the properties naturally expected from it. A prototype implementation of our approach was developed in the K framework. We demonstrate the genericity of our tool by instantiating it on several languages, and show how it can be used for the symbolic execution and model checking of several programs.

6.1.2. Program Equivalence

In [12] we propose a logic and a deductive system for stating and automatically proving the equivalence of programs in deterministic languages having a rewriting-based operational semantics. The deductive system is circular in nature and is proved sound and weakly complete; together, these results say that, when it terminates, our system correctly solves the program-equivalence problem as we state it. We show that our approach is suitable for proving the equivalence of both terminating and non-terminating programs, and also the equivalence of both concrete and symbolic programs. The latter are programs in which some statements or expressions are symbolic variables. By proving the equivalence between symbolic programs, one proves in one shot the equivalence of (possibly, infinitely) many concrete programs obtained by replacing the variables by concrete statements or expressions. We also report on a prototype implementation of the proposed deductive system in the K framework.

6.1.3. Program Verification

In [14] we present an automatic and language-independent program verification approach based on symbolic execution. The specification formalism we consider is Reachability Logic, a language-independent logic that constitutes an alternative to Hoare logics. Reachability Logic has a sound and relatively complete deduction system, which offers a lot of freedom (but no guidelines) for constructing proofs. Hence, we propose symbolic execution as a strategy for proof construction. We show that, under reasonable conditions on the semantics of programming languages, our symbolic-execution based Reachability-Logic formula verification is sound. We present a prototype implementation of the resulting language-independent verifier as an extension of a generic symbolic execution engine that we are developing in the K framework. The verifier is illustrated on programs written in languages also formally defined in K.

6.2. Master-Slave Control Structure for MP-SoC Architectures

Our Synchronous Communication Asynchronous Computation (SCAC) model is a data-parallel execution model dedicated to the Massively Parallel System-on-Chip. This model proposes a novel control structure, referred to as master-slave control [11]. Its concept departs from the centralized configuration. However, instead of a uni-processor master controlling a set of parallel processing elements (PE), the master cooperates with a grid of parallel slave controllers which supervises the activities of cluster of PEs.

The control structure in SCAC model is presented by two hierarchical control levels:

- The Master Control Unit (MCU), which controls the order execution in the whole system. It is a simple processor, which fetches and decodes program instruction and broadcasts execution orders to Slave Control Unit. It controls the end execution to establish synchronous communication.
- The Slave Control Unit (SCU), which controls: local node and PEs activities, parallel instructions execution and synchronous communication. It is a crucial component in the master-slave control structure. The SCUs grid allows independent parallel execution.

The hardware architecture is composed of a single MCU and multiple Slave controllers (SCUs) combined with local processing element (PE) (or a cluster of 16 PEs), known collectively as Nodes. The MCU and SCU array are connected through single level hierarchical bus and the SCUs are connected together through X-net interconnection network [2]. This network is clocked synchronously with the SCUs and respectively with the PEs. SCU controllers in the grid care for the instruction execution activities that involve a large degree of parallelism and the communication activities that need to coordinate all the PEs in the grid. The structure of master-slave control should be distinguished from other hierarchical or clustered approaches proposed for parallel computing. Such proposals are usually motivated by memory latency considerations and the desire to build a scalable system. The use of two control levels is therefore visible to the user in its effect on the communication between various processors. With master-slave control structure, the PEs in massively parallel system can execute independently and then can communicate synchronously. Such a construction has the advantage of allowing the designer to optimize distinct processors for their intended tasks and to implement simple interconnection network without additionally buffers and complex routing algorithms.

The aim of these last works is to design a master-slaves control structure for SCAC architecture to allow autonomous processing with simple and regular communication. This control structure based on IP blocks which offers good flexibility and scalability was implemented in synthesizable VHDL code. It is simulated and synthetized for Xilinx Virtex q6 (XC6VLX240T) board. The difficulty of designing a master-slave structure is a compromise between an optimal execution time and high flexibility, while reducing power consumption and silicon area.

6.3. Toward a Massively Parallel and Reflective Execution Model

FPGAs are undoubtedly suited to the definition of what could be called a DSHA (Domain Specific Hardware Architecture). Similarity with the DSSA (Domain Specific Software Architecture) an assembly of functional components performs basic transformations on data, while a software / hardware infrastructure ensures the ordering of these transformations. The HoMade processor is designed with this in mind: it can be seen as an IP integrator offering a mechanism for interprocess communication IPs via a battery and a scheduler of IPs via dedicated instructions for flow control. In this control we find two particular instructions for flow control designed for a massively parallel execution model for SPMD, and a new instruction can make HoMade reflexive . With this instruction, you can at runtime change the behavior of a virtual component by dynamically associating it to a particular HoMade instruction sentence and in particular IP triggering instructions. Same components can successively after applying this instruction, trigger a hardware IP, a software function which itself can trigger a flow of execution of hardware IPs. This intercession [2] feature , parts of HoMade core, is valid for one processor or for all HoMade slave components in a massively parallel architecture. We demonstrated on a FPGA board which computes the Fibonacci sequence with three different methods, but always through a single call to a unique Virtual Component.

6.4. Power Estimation at System-Level for MPSoC Based Platforms

Shifting the design entry point up to the system level is the most important countermeasure adopted to manage the increasing complexity of Multiprocessor System on Chip (MPSoC). The reason is that decisions taken at this level, early in the design cycle, have the greatest impact on the final design in terms of power and energy efficiency. However, taking decisions at this level is very difficult, since the design space is extremely wide and it has so far been mostly a manual activity. Efficient system-level power estimation tools are therefore necessary to enable proper Design Space Exploration (DSE) based on power/energy and timing. We propose a tool based on efficient hybrid system level power estimation methodology for MPSoC. In this methodology, a combination of Functional Level Power Analysis (FLPA) and system level simulation technique are used to compute the power of the whole system. Basically, the FLPA concept is proposed for processor architecture in order to obtain parameterized arithmetic power models depending on the consumption of the main functional blocks. In this work, FLPA is extended to set up generic power models for the different parts of the platform. In addition, a simulation framework is developed at the transactional level to evaluate accurately the activities used in the related power models. The combination of the above two parts leads to a hybrid power estimation, that gives a better trade-off between accuracy and speed. The proposed methodology has several benefits: It considers the power consumption of the embedded system in its entirety; and Leads to accurate estimates without a costly and complex material. The proposed methodology is also scalable for exploring complex embedded architectures. Based on the proposed methodology, our Power Estimation Tool at System-Level (PETS) is developed. The usefulness and effectiveness of our PETS tool is validated through a typical mono-processor and multiprocessor embedded system designed around the TI OMAP (3530 and 5912) and the Xilinx Virtex II Pro FPGA boards. This methodology is demonstrated and evaluated by using a variety of basic programs to complete media benchmarks. Estimated power values are compared to real board measurements for both simple and multiprocessor architectures. Our obtained power estimation results provide less than 3% of error for mono-processor, 3.8% for homogeneous multiprocessor system and 4.3% for heterogeneous multiprocessor system and 70x faster compared to the state-of-the-art power estimation tools. These results have been presented in the PhD of Santhosh Kumar Rethinagiri [2] and published in [4].

6.5. Dynamically reconfigurable CPU/FPGA architecture for the testing and simulation of avionic systems

Real-time computing systems are increasingly used in aerospace and avionic industries. In the face of power wall and real-time requirements, hardware designers are directed towards reconfigurable computing with the usage of heterogeneous CPU/FPGA systems. However, there is a lack of real-time environments able to deal with the execution of applications on such heterogeneous systems dedicated to avionic Testing and Simulation (T&S). This year, we addressed the problem of soft real-time environments for CPU/FPGA systems and we proposed first a high-performance hardware architecture used to implement intimately coupled hardware and software avionic models. Second, we developed an efficient real-time software environment for the model's

[2] Wikipedia definition: intercession is the ability of a program to modify its own execution state or alter its own interpretation or meaning.

execution, the multi-core CPU monitoring and the runtime task re-allocation to avoid the timing constraint violation. Experimental results underpin the industrial relevance of the presented approach for avionic T&S systems with real-time support. These results are presented in the PhD of George Afonso [1] and in different publications [7] [10] [8].

6.6. A custom reconfiguration controller for partial and dynamic reconfiguration in HoMade based systems

In all Xilinx devices supporting dynamic reconfiguration, such a functionality is realized using a hardware reconfiguration port called ICAP, that moves bitstreams from the reconfiguration memory to the programmable logic. ICAP is initialized by a Xilinx HW controller driven exclusively by a Microblaze processor and thus connected to a PLB or AXI bus.

This makes the partial and dynamic reconfiguration a very tedious task, as it implies using several Xilinx tools (XPS, ISE, PlanAhead,..etc). PDR becomes also resources and time consuming due to the fact that it uses very large interfaces and a static Xilinx architecture (in addition to the system that we want to design) including specific processors, buses, controllers,..etc.

Our contribution is the design of a custom ICAP controller, driven only by a HoMade processor, without any additional processors, buses or controllers. This ensures that our HoMade reconfigurable systems consumes fewer resources on the FPGA and does not require other tools than the standard ISE and PlanAhead tools in order to be designed.

6.7. Hardware control for partially and dynamically reconfigurable systems: from modelling to implementation

This work proposes a control design methodology for FPGA-based reconfigurable systems aiming at increasing control design productivity and guaranteeing implementation efficiency. This methodology is based on a semi-distributed control model [5] composed of a set of modular distributed controllers executing each observation, decision-making and reconfiguration tasks for a reconfigurable region of the system, and a coordinator between the distributed controllers decisions in order to respect global systems constraints and objectives. This semi-distributed decision-making is based on the mode-automata formalism. The proposed combination between modularity, control splitting and formalism-based design allows to enhance the flexibility, reusability and scalability of the control design. Another point that can be added to this combination, to enhance design productivity, is design automation. For this, the proposed methodology is based on Model-Driven Engineering approach [5] allowing to automate code generation from high-level models. This approach makes use of the UML MARTE (Modeling and Analysis of Real-Time and Embedded Systems) standard profile, allowing to make low-level technical details transparent to designers and to automate the VHDL code generation for hardware implementation of the modeled control systems in order to guarantee their performance. The generated control systems were validated using simulation. Synthesis results showed an acceptable time and resource overhead for systems having different numbers of controllers. A control system composed of four controllers and a coordinator was also validated through physical implementation in an FPGA system for an image processing application.

6.8. A model-based approach for dynamically reconfigurable systems design: from MARTE to RecoMARTE

This work is done in the context of the ANR FAMOUS project. It proposes a co-design methodology of dynamically reconfigurable systems based on FPGA. Our methodology is based on the Engineering Model Driven approach (MDE) and the models specification is done in the UML MARTE profile. It aims at ensuring flexibility, reusability and automation to facilitate the work of the designer and improve his productivity. The first contribution related is identifying parts of dynamically reconfigurable FPGA that can be modeled at the high abstraction levels. So, we defined a design flow based on the MDE to ensure the automation of code generation. According to this flow, several models are created mainly through MARTE profile concepts.

However, the modeling concepts of dynamic reconfiguration on FPGAs required extensions in MARTE. Thus, we identified the missing concepts to be integrated in a new profile that extends MARTE called RECOMARTE. The second contribution allows the chain automation and experimental validation. To integrate our design flow and to automate code generation, a processing chain was used. The final model resulting from MARTE proposed design flow is given as input to this chain.

We thereby move from MARTE to RECOMARTE models via an intermediate description according to the IP-XACT standard to finally generate files describing the complete system in the Xilinx XPS environment. This automation will accelerate the design phase and avoid errors due to the direct manipulation of these details. Finally, an example of application of image processing has been developed to demonstrate and validate our methodology.

DYLISS Project-Team

6. New Results

6.1. Data integration

Participants: Jacques Nicolas, Andres Aravena, Charles Bettembourg, Jérémie Bourdon, Jeanne Cambefort, Guillaume Collet, Olivier Dameron, Damien Eveillard, Julie Laniau, Sylvain Prigent, Anne Siegel, Sven Thiele, Valentin Wucher.

Metabolic network reconstruction: combinatorial gap-filling method We introduced an exhaustive gap-filling procedure on the first metabolic network for a macroalgea (Ectocarpus Siliculosus). As this species is a non benchmark model, this issue is related to hard combinatorial optimization problems. To that matter, we took advantages of the latest improvement of Answer Set Programming solvers (combination of clasp and unclasp) and introduced a new model of the network expansion problem. [*G. Collet, D. Eveillard, S. Prigent, A. Siegel, S. Thiele*] [27]

Identification of functional gene units in non benchmark models We introduced the concept of "shortest genome segments" (SGS) to detect functional units on exotic species, such as extremophiles, that are by nature unrefined. They correspond to genome portion which contain a large density of genes coding for enzymes which regulate successive reactions of metabolic pathways. There identification is a hard optimization combinatorial problem. We relied on the declarative modeling power of answer set programming (ASP) to encode the identification of shortest genome segments and prove that SGS are stable in (i) computational time and (ii) ability to predict functional units when one deteriorates the biological knowledge [*D. Eveillard, A. Siegel, S. Thiele*] [26]

Refinement of regulatory network from genomic, expression data and functional unit data We integrated heterogeneous information from two types of network predictions to determine a causal explanation for the observed gene co-expression. We modeled this integration as a combinatorial optimization problem. We demonstrated that this problem belongs to the NP-hard complexity class. We proposed an heuristic approach to have an approximate solution in a practical execution time. Our evaluation showed that the E.coli regulatory network resulting from the application of this method has higher accuracy than the putative one built with traditional tools. Applications to the mining bacterium *Acidithiobacillus ferrooxidans* allowed analyzing the relevance of central regulators. [*A. Aravena, D. Eveillard, A. Siegel*] [23], [13] [Thesis]

Reconstruction of a protein interaction network for archaebacteria To gain insights into genomic maintenance processes in hyperthermophilic archaea, a protein-interaction network centered on informational processes of *Pyrococcus abyssi* was generated by affinity purification coupled with mass spectrometry. We have proposed a graph theoretic analysis of this network including statistical (e.g. clusterisation coefficients) and topological aspects (bicluster analysis, search of a maximal interaction skeleton), which helps network interpretation in terms of formation of complexes or interaction dynamics. [*J. Nicolas*] [20] [Online publication]

Knowledge evolution in ontologies We studied the impact of an ontology evolution on its structural complexity. As a case study we used sixty monthly releases of the Gene Ontology and its three independent branches i.e. biological processes (BP), cellular components (CC) and molecular functions (MF). For each release, we measured complexity by computing metrics related to the size, the nodes connectivity and the hierarchical structure. We showed that the variation of the number of classes and relations in an ontology does not provide enough information about the evolution of its complexity. However, connectivity and hierarchy-related metrics revealed different patterns of values as well as of evolution for the three branches of the Gene Ontology [*O. Dameron, C. Bettembourg*] [17], [14] [Online publication][Thesis]

Treatment process representation for breast cancer patients. The general cancer registry of Poitou-Charentes developed a multiple source information system covering diseases, anatomical structures and cytopathology. We proposed an algorithm for representing and analyzing the patient's treatment process. An expert compared the original data with our representation and computed a score of dissimilarity. The results showed that an integrated information system can successfully analyze the data to determine whether they comply with the guidelines [*O. Dameron*] [31].

AphidAtlas project We began a collaboration with the AphidAtlas project for defining the structure of an ontology of aphids anatomy and development [*O. Dameron*] [30].

6.2. Asymptotic dynamics

Participants: Anne Siegel, Oumarou Abdou-Arbi, Geoffroy Andrieux, Jérémie Bourdon, Jeanne Cambefort, Damien Eveillard, Michel Le Borgne, Vincent Picard, Sven Thiele, Santiago Videla.

Learning families of boolean signaling networks We propose the use of ASP to explore the space of feasible logic models of a signaling network. To that matter, we exhaustively enumerate the set of sub-optimal boolean logical models which are compatible with both the topology of a knowledge-based influence graph and the observed response of the system to several perturbations (phosphorylation datasets). We illustrate the importance of characterizing such a family of models in a global and exhaustive manner by revisiting a model of pro-growth and inflammatory pathways in human liver cells and studying the variability with the set of compatible models. [*A. Siegel, S. Thiele, S. Videla*] [18] [Online publication]

Control the steady-state response of qualitative signaling networks: intervention sets The minimal intervention set problem roughly consists in identifying the perturbation that can be undergone over a signaling network to predict a fixed expected behavior. We have provided a precise characterization of the minimal intervention set problem relying on three-valued logic and fixpoint semantics. We address this problem within ASP and using real-world biological benchmarks we show that it greatly outperforms previous work using dedicated algorithms. [*A. Siegel, S. Videla*] [19] [Online publication]

Reachability in dynamical signaling networks: cut sets In the scope of discrete finite-state models of interacting components, we present a novel algorithm for identifying sets of local states of components whose activity is necessary for the reachability of a given local state. Those sets are referred to as cut sets; they provide potential therapeutic targets that are proven to prevent molecules of interest to become active, up to the correctness of the model. Our method is based on the so-called Graph of Local Causality and form an under-approximation of the complete minimal cut sets of the dynamics. It makes tractable the formal analysis of very large scale networks. [*G. Andrieux, M. Le Borgne*] [28], [12] [Online publication][Thesis]

Exploring metabolism flexibility through quantitative study of precursor sets for system outputs We extended a Flux-Balanced-Analysis approach to quantify the precursor composition of each system output and to discuss the biological relevance of a set of flux in a given metabolic network. The composition is called contribution of inputs over outputs [AIO]. In order to further investigate metabolic network flexibility, we have proposed an efficient local search algorithm computing the extremal values of AIO coefficients. This approach enables to discriminate diets without making any assumption on the internal behaviour of the system. [*J. Bourdon, O. Abdou-Arbi, A. Siegel*] [15], [11] [Thesis]

6.3. Sequence annotation

Participants: François Coste, Aymeric Antoine-Lorquin, Catherine Belleannée, Guillaume Collet, Gaëlle Garet, Clovis Galiez, Laurent Miclet, Olivier Quenez, Jacques Nicolas, Valentin Wucher.

Refinement of mi-RNA regulation network thanks to concept analysis MicroRNAs (miRNAs) are small RNA molecules that bind messenger RNAs (mRNAs) to silence their expression. To improve the discrimination between true and false interactions during their prediction, we defined a repair process based on the hypothesis that the true graph is formed by interaction modules represented by formal concepts, i.e. set of miRNAs having the same regulation profile. To validate our hypothesis and method, we have extracted parameters from a biological miRNA/mRNA network and used them to build random networks. Each repaired network can be evaluated with a score balancing the number of edge changes and the conceptual adequacy in the spirit of the minimum description length principle. [*J. Nicolas, V. Wucher*] [32]

Analogical proportions and the factorization of information in distributive lattices. We have conducted theoretical studies to elucidate whether formal concept lattices can have properties that could be used in further studies. In this direction, analogical proportions are statements involving four entities, of the form 'A is to B as C is to D'. They play an important role in analogical reasoning. They have been formalized in both a propositional logic setting and an algebraic setting. We define and study analogical proportions in the general setting of lattices, and more particularly of distributive lattices. We discussed the decomposition of analogical proportions in canonical proportions as well as the resolution of analogical proportion equations, and illustrate especially on the case of Boolean lattices, which reflects the logical modeling. [*L. Miclet*] [24], [29]

Bioinformatics and Artificial Intelligence In this book chapter, we introduce the main objects studied in Bioinformatics at different levels (the macromolecules, their interactions as well as the knowledge formalization or extraction) and present meanwhile a survey of the contribution and influence of Artificial Intelligence to this research field on related key tasks (gene prediction, functional annotation, structure prediction, transcriptomics analysis, network acquisition and analysis, knowledge integration and formalization, information retrieval and extraction from documents, ...). [*F. Coste*] [33]

Genome studies: fast assembly and SNP identification This work is a follow-up of collaborations with the GenScale team and the GenOuest platform. We reported the first identification of a set of SNPs isolated from the genome of *I. ricinus* - an important vector of pathogens in Europe, by applying a reduction of genomic complexity, pyrosequencing and new bioinformatics tools[21] [Online publication]. We also contributed to show that the genome assembly program MINIA is successfully able to assemble a 100 Mbp genome on a very low-end, low-power system with 512 MB RAM and a 32 GB flash drive such as a Raspberry Pi. [*G. Collet, O. Quenez*] [34][Online publication]

DYOGENE Project-Team

6. New Results

6.1. Ancillary service to the grid from deferrable loads: the case for intelligent pool pumps in Florida

Renewable energy sources such as wind and solar power have a high degree of unpredictability and time-variation, which makes balancing demand and supply challenging. One possible way to address this challenge is to harness the inherent flexibility in demand of many types of loads. In [28], we focus on pool pumps, and how they can be used to provide ancillary service to the grid for maintaining demand-supply balance. A Markovian Decision Process (MDP) model is introduced for an individual pool pump. A randomized control architecture is proposed, motivated by the need for decentralized decision making, and the need to avoid synchronization that can lead to large and detrimental spikes in demand. An aggregate model for a large number of pools is then developed by examining the mean field limit. A key innovation is an LTI-system approximation of the aggregate nonlinear model, with a scalar signal as the input and a measure of the aggregate demand as the output. This makes the approximation particularly convenient for control design at the grid level. Simulations are provided to illustrate the accuracy of the approximations and effectiveness of the proposed control approach.

6.2. Impact of Storage on the Efficiency and Prices in Real-Time Electricity Markets

In [19] we study the effect of energy-storage systems in dynamic real-time electricity markets. We consider that demand and renewable generation are stochastic, that real-time production is affected by ramping constraints, and that market players seek to selfishly maximize their profit. We distinguish three scenarios, depending on the owner of the storage system: (A) the supplier, (B) the consumer, or (C) a stand-alone player. In all cases, we show the existence of a competitive equilibrium when players are price-takers (they do not affect market prices). We further establish that under the equilibrium price process, players' selfish responses coincide with the social welfare-maximizing policy computed by a (hypothetical) social planner. We show that with storage the resulting price process is smoother than without. We determine empirically the storage parameters that maximize the players' revenue in the market. In the case of consumer-owned storage, or a stand-alone storage operator (scenarios B and C), we find that they do not match socially optimal parameters. We conclude that consumers and the stand-alone storage operator (but not suppliers) have an incentive to under-dimension their storage system. In addition, we determine the scaling laws of optimal storage parameters as a function of the volatility of demand and renewables. We show, in particular, that the optimal storage energy capacity scales as the volatility to the fourth power.

6.3. Risk-Aware SLA Negotiation

In order to assure Quality of Service (QoS) connectivity, Network Service Providers (NSPs) negotiate Service Level Agreements (SLAs). However, a committed SLA might fail to respect its QoS promises. In such a case, the customer is refunded. To maximize their revenues, the NSPs must deal with risks of SLA violations, which are correlated to their network capacities. Due to the complexity of the problem, we first study in [21], a system with one NSP provider and give a method to compute its risk-aware optimal strategy using (max; +)-algebras. Using the same method, we study the case where two NSPs collaborate and the case where they compete, and we derive the Price of Anarchy. This method provides optimal negotiation strategies but, when modeling customers' reaction to SLA failure, analytical results do not hold. Hence, we propose a learning framework that chooses the NSP risk-aware optimal strategy under failures capturing the impact of reputation. Finally, by simulation, we observe how the NSP can benefit from such a framework.

6.4. Impact of Rare Alarms on Event Correlation

Nowadays, telecommunication systems are growing more and more complex, generating a large amount of alarms that cannot be effectively managed by human operators. The problem is to detect significant combinations of alarms describing an issue in real-time. In [18], we present a powerful heuristic algorithm that constructs dependency graphs of alarm patterns. More precisely, it highlights patterns extracted from an alarm flow obtained from a learning process with a small footprint on network management system performance. This algorithm helps to detect issues in real-time by effectively delivering concise alarm patterns. Furthermore, it allows the proactive analysis of the functioning of a network by computing the general trends of this network. We evaluate our algorithm on an optical network alarm data set of an existing operator. We find similar results as the expert analysis performed for this operator by Alcatel-Lucent Customer Services.

6.5. Some Synchronization Issues in OSPF Routing

A routing protocol such as OSPF has a cyclic behavior to regularly update its view of the network topology. Its behavior is divided into periods. Each period produces a flood of network information messages. We observe a regular activity in terms of messages exchanges and filling of receive buffers in routers. [17] examines the consequences of possible overlap of activity between periods, leading to a buffer overflow. OSPF allows "out of sync" flows by considering an initial delay (phase). We study the optimum calculation of these offsets to reduce the load, while maintaining a short period to ensure a protocol reactive to topology changes. Such studies are conducted using a simulated Petri net model. A heuristic for determining initial delays is proposed. A core network in Germany serves as illustration.

6.6. Exact Worst-case Delay in FIFO-multiplexing Feed-forward Networks

In this paper we compute the actual worst-case end-to-end delay for a flow in a feed-forward network of FIFO-multiplexing service curve nodes, where flows are shaped by piecewise-affine concave arrival curves, and service curves are piecewise affine and convex. We show that the worst-case delay problem can be formulated as a mixed integer-linear programming problem, whose size grows exponentially with the number of nodes involved. Furthermore, we present approximate solution schemes to find upper and lower delay bounds on the worst-case delay. Both only require to solve just *one* linear programming problem, and yield bounds which are generally more accurate than those found in the previous work, which are computed under more restrictive assumptions.

6.7. Fast weak–KAM integrators for separable Hamiltonian systems

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.

6.8. Probabilistic cellular automata, invariant measures, and perfect sampling

A probabilistic cellular automaton (PCA) can be viewed as a Markov chain. The cells are updated synchronously and independently, according to a distribution depending on a finite neighborhood. In [9], we investigate the ergodicity of this Markov chain. A classical cellular automaton is a particular case of PCA. For a one-dimensional cellular automaton, we prove that ergodicity is equivalent to nilpotency, and is therefore undecidable. We then propose an efficient perfect sampling algorithm for the invariant measure of an ergodic PCA. Our algorithm does not assume any monotonicity property of the local rule. It is based on a bounding process which is shown to also be a PCA. Last, we focus on the PCA majority, whose asymptotic behavior is unknown, and perform numerical experiments using the perfect sampling procedure.

6.9. Density Classification on Infinite Lattices and Trees

Consider an infinite graph with nodes initially labeled by independent Bernoulli random variables of parameter p. In [7], 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 to the uniform configuration with only 0's if p<1/2, and only 1's if p>1/2. We present solutions to the problem on the regular grids of dimension d, for any d>1, and on the regular infinite trees. For the bi-infinite line, we propose some candidates that we back up with numerical simulations.

6.10. Semi-infinite paths of the radial spanning tree

In the paper [4], in collaboration with David Coupier and Viet Chi Tran of Lille 1, we study the semi-infinite paths of the radial spanning tree (RST) of a Poisson point process in the plane using Stochastic Geometry. We first show that the expectation of the number of intersection points between semi-infinite paths and the sphere with radius r grows sublinearly with r. Then, we prove that in each (deterministic) direction, there exists with probability one a unique semi-infinite path, framed by an infinite number of other semi-infinite paths of close asymptotic directions. The set of (random) directions in which there are more than one semi-infinite paths is dense in $[0, 2\pi)$. It corresponds to possible asymptotic directions of competition interfaces. We show that the RST can be decomposed in at most five infinite subtrees directly connected to the root. The interfaces separating these subtrees are studied and simulations are provided.

6.11. Generating functionals of random packing point processes

In the paper [45], we study the generating functionals of a class of random packing point processes of the Matérn type. Consider a symmetrical conflict relationship between the points of a point process. The Matérn type constructions provide a generic way of selecting a subset of this point process which is conflict-free. The simplest one consists in keeping only conflict-free points. There is however a wide class of Matérn type processes based on more elaborate selection rules and providing larger sets of selected points. The general idea being that if a point is discarded because of a given conflict, there is no need to discard other points with which it is also in conflict. The ultimate selection rule within this class is the so called Random Sequential Adsorption, where the cardinality of the sequence of conflicts allowing one to decide whether a given point is selected is *not* bounded. The present paper provides a sufficient condition on the span of the conflict relationship under which all the above point processes are well defined when the initial point process is Poisson. It then establishes, still in the Poisson case, a set of differential equations satisfied by the probability generating functionals of these Matérn type point processes. Integral equations are also given for the Palm distributions.

6.12. Clustering and percolation of point processes

We are interested in phase transitions in certain percolation models on point processes and their dependence on clustering properties of the point processes. In [5], we show that point processes with smaller void probabilities and factorial moment measures than the stationary Poisson point process exhibit non-trivial phase transition in the percolation of some coverage models based on level-sets of additive functionals of the point process. Examples of such point processes are determinantal point processes, some perturbed lattices, and more generally, negatively associated point processes. Examples of such coverage models are k-coverage in the Boolean model (coverage by at least k grains) and SINR-coverage (coverage if the signal-to-interference-and-noise ratio is large). In particular, we answer in affirmative the hypothesis of existence of phase transition in the percolation of k-faces in the Čech simplicial complex (also called clique percolation) on point processes which cluster less than the Poisson process. We also construct a Cox point process, which is “more clustered" than the Poisson point process and whose Boolean model percolates for arbitrarily small radius. This shows that clustering (at least, as detected by our specific tools) does not always “worsen” percolation, as well as that upper-bounding this clustering by a Poisson process is a necessary assumption for the phase transition to hold.

6.13. Using Poisson processes to model lattice cellular networks

An almost ubiquitous assumption made in the stochastic-analytic approach to study of the quality of user-service in cellular networks is Poisson distribution of base stations, often completed by some specific assumption regarding the distribution of the fading (e.g. Rayleigh). The former (Poisson) assumption is usually (vaguely) justified in the context of cellular networks, by various irregularities in the real placement of base stations, which ideally should form a lattice (e.g. hexagonal) pattern. In the first part of [14] we provide a different and rigorous argument justifying the Poisson assumption under sufficiently strong log-normal shadowing observed in the network, in the evaluation of a natural class of the typical-user service-characteristics (including path-loss, interference, signal-to-interference ratio, spectral efficiency). Namely, we present a Poisson-convergence result for a broad range of stationary (including lattice) networks subject to log-normal shadowing of increasing variance. We show also for the Poisson model that the distribution of all these typical-user service characteristics does not depend on the particular form of the additional fading distribution. Our approach involves a mapping of 2D network model to 1D image of it "perceived" by the typical user. For this image we prove our Poisson convergence result and the invariance of the Poisson limit with respect to the distribution of the additional shadowing or fading. Moreover, in the second part of the paper we present some new results for Poisson model allowing one to calculate the distribution function of the SINR in its whole domain. We use them to study and optimize the mean energy efficiency in cellular networks.

6.14. Compactification of the Action of a Point-Shift on the Palm Probability of a Point Process

In collaboration with Mir-Omid Haji-Mirsadeghi (Sharif University, Iran) [50], we analyzed the compactification of Palm probabilities by the action of a point–shift. A point-shift maps, in a translation invariant way, each point of a stationary point process Φ to some point of Φ. The initial motivation of this paper is the construction of probability measures, defined on the space of counting measures with an atom at the origin, which are left invariant by a given point-shift f. The point-shift probabilities of Φ are defined from the action of the semigroup of point-shift translations on the space of Palm probabilities, and more precisely from the compactification of the orbits of this semigroup action. If the point-shift probability is uniquely defined, and if f is continuous with respect to the vague topology, then the point-shift probability of Φ provides a solution to the initial question. Point-shift probabilities are shown to be a strict generalization of Palm probabilities: when the considered point-shift f is bijective, the point-shift probability of Φ boils down to the Palm probability of Φ. When it is not bijective, there exist cases where the point-shift probability of Φ is the law of Φ under the Palm probability of some stationary thinning Ψ of Φ. But there also exist cases where the point-shift probability of Φ is singular w.r.t. the Palm probability of Φ and where, in addition, it cannot be the law of Φ under the Palm probability of any stationary point process Ψ jointly stationary with Φ. The paper also gives a criterium of existence of the point-shift probabilities of a stationary point process and discusses uniqueness. The results are illustrated through several examples.

6.15. A Stochastic Geometry Framework for Analyzing Pairwise-Cooperative Cellular Networks

Cooperation in cellular networks has been recently suggested as a promising scheme to improve system performance, especially for cell-edge users. In [34], we use stochastic geometry to analyze cooperation models where the positions of Base Stations (BSs) follow a Poisson point process distribution and where Voronoi cells define the planar areas associated with them. For the service of each user, either one or two BSs are involved. If two, these cooperate by exchange of user data and channel related information with conferencing over some backhaul link. Our framework generally allows variable levels of channel information at the transmitters. In this paper we investigate the case of limited channel state information for cooperation (channel phase, second neighbour interference), but not the fully adaptive case which would require considerable feedback. The total per-user transmission power is further split between the two transmitters and a common message is encoded. The decision for a user to choose service with or without cooperation is directed by a family of geometric policies depending on its relative position to its two closest base stations. An exact expression of the network coverage probability is derived. Numerical evaluation allows one to analyze significant coverage benefits compared to the non-cooperative case. As a conclusion, cooperation schemes can improve system performance without exploitation of extra network resources.

6.16. SINR-based k-coverage probability in cellular networks with arbitrary shadowing

In [20], we give numerically tractable, explicit integral expressions for the distribution of the signal-to-interference-and-noise-ratio (SINR) experienced by a typical user in the down-link channel from the k-th strongest base stations of a cellular network modelled by Poisson point process on the plane. Our signal propagation-loss model comprises of a power-law path-loss function with arbitrarily distributed shadowing, independent across all base stations, with and without Rayleigh fading. Our results are valid in the whole domain of SINR, in particular for $SINR < 1$, where one observes multiple coverage. In this latter aspect our paper complements previous studies reported in [55].

6.17. Equivalence and comparison of heterogeneous cellular networks

In [15], we consider a general heterogeneous network in which, besides general propagation effects (shadowing and/or fading), individual base stations can have different emitting powers and be subject to different parameters of Hata-like path-loss models (path-loss exponent and constant) due to, for example, varying antenna heights. We assume also that the stations may have varying parameters of, for example, the link layer performance (SINR threshold, etc). By studying the *propagation processes* of signals received by the typical user from all antennas marked by the corresponding antenna parameters, we show that seemingly different heterogeneous networks based on Poisson point processes can be equivalent from the point of view a typical user. These neworks can be replaced with a model where all the previously varying propagation parameters (including path-loss exponents) are set to constants while the only trade-off being the introduction of an isotropic base station density. This allows one to perform analytic comparisons of different network models via their isotropic representations. In the case of a constant path-loss exponent, the isotropic representation simplifies to a homogeneous modification of the constant intensity of the original network, thus generalizing a previous result showing that the propagation processes only depend on one moment of the emitted power and propagation effects. We give examples and applications to motivate these results and highlight an interesting observation regarding random path-loss exponents.

6.18. How user throughput depends on the traffic demand in large cellular networks: a typical cell analysis and real network measurements

In [40], we assume a space-time Poisson process of call arrivals on the infinite plane, independently marked by data volumes and served by a cellular network modeled by an infinite ergodic point process of base stations. Each point of this point process represents the location of a base station that applies a processor sharing policy to serve users arriving in its vicinity, modeled by the Voronoi cell, possibly perturbed by some random signal propagation effects. User service rates depend on their signal-to-interference-and-noise ratios with respect to the serving station. Little's that allows to express the mean user throughput in any region of this network model as the ratio of the mean traffic demand to the steady-state mean number of users in this region. Using ergodic arguments and the Palm theoretic formalism, we define a global mean user throughput in the cellular network and prove that it is equal to the ratio of mean traffic demand to the mean number of users in the steady state of the “typical cell” of the network. Here, both means account for double averaging: over time and network geometry, and can be related to the per-surface traffic demand, base-station density and the spatial distribution of the signal-to-interference-and-noise ratio. This latter accounts for network irregularities, shadowing and cell dependence via some cell-load equations. Inspired by the analysis of the typical cell, we propose also a simpler, approximate, but fully analytic approach, called the mean cell approach. The key quantity explicitly calculated in this approach is the cell load. In analogy to the load factor of the (classical) M/G/1 processor sharing queue, it characterizes the stability condition, mean number of users and the mean user throughput. We validate our approach comparing analytical and simulation results for Poisson network model to real-network measurements.

6.19. Analysis of a Proportionally Fair and Locally Adaptive spatial Aloha in Poisson Networks

The proportionally fair sharing of the capacity of a Poisson network using Spatial-Aloha leads to closed-form performance expressions in two extreme cases: (1) the case without topology information, where the analysis boils down to a parametric optimization problem leveraging stochastic geometry; (2) the case with full network topology information, which was recently solved using shot-noise techniques. In [37], we show that there exists a continuum of adaptive controls between these two extremes, based on local stopping sets, which can also be analyzed in closed form. We also show that these control schemes are implementable, in contrast to the full information case which is not. As local information increases, the performance levels of these schemes are shown to get arbitrarily close to those of the full information scheme. The analytical results are combined with discrete event simulation to provide a detailed evaluation of the performance of this class of medium access controls.

6.20. Optimal Rate sampling in 802.11 Systems

In 802.11 systems, Rate Adaptation (RA) is a fundamental mechanism allowing transmitters to adapt the coding and modulation scheme as well as the MIMO transmission mode to the radio channel conditions, and in turn, to learn and track the (mode, rate) pair providing the highest throughput. So far, the design of RA mechanisms has been mainly driven by heuristics. In contrast, in [42], we rigorously formulate such design as an online stochastic optimisation problem. We solve this problem and present ORS (Optimal Rate Sampling), a family of (mode, rate) pair adaptation algorithms that provably learn as fast as it is possible the best pair for transmission. We study the performance of ORS algorithms in both stationary radio environments where the successful packet transmission probabilities at the various (mode, rate) pairs do not vary over time, and in non-stationary environments where these probabilities evolve. We show that under ORS algorithms, the throughput loss due to the need to explore sub-optimal (mode, rate) pairs does not depend on the number of available pairs, which is a crucial advantage as evolving 802.11 standards offer an increasingly large number of (mode, rate) pairs. We illustrate the efficiency of ORS algorithms (compared to the state-of-the-art algorithms) using simulations and traces extracted from 802.11 test-beds.

6.21. Flooding in Weighted Sparse Random Graphs

In [3], we study the impact of edge weights on distances in sparse random graphs. We interpret these weights as delays and take them as independent and identically distributed exponential random variables. We analyze the weighted flooding time defined as the minimum time needed to reach all nodes from one uniformly chosen node and the weighted diameter corresponding to the largest distance between any pair of vertices. Under some standard regularity conditions on the degree sequence of the random graph, we show that these quantities grow as the logarithm of n when the size of the graph n tends to infinity. We also derive the exact value for the prefactor. These results allow 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.

6.22. Viral Marketing On Configuration Model

In [38], we consider propagation of influence on a Configuration Model, where each vertex can be influenced by any of its neighbours but in its turn, it can only influence a random subset of its neighbours. Our (enhanced) model is described by the total degree of the typical vertex, representing the total number of its neighbours and the transmitter degree, representing the number of neighbours it is able to influence. We give a condition involving the joint distribution of these two degrees, which if satisfied would allow with high probability the influence to reach a non-negligible fraction of the vertices, called a *big (influenced) component*, provided that the source vertex is chosen from a set of *good pioneers*. We show that asymptotically the big component is essentially the same, regardless of the good pioneer we choose, and we explicitly evaluate the asymptotic relative size of this component. Finally, under some additional technical assumption we calculate the relative size of the set of good pioneers. The main technical tool employed is the "fluid limit" analysis of the joint exploration of the configuration model and the propagation of the influence up to the time when a big influenced component is completed. This method was introduced in [59] to study the giant component of the configuration model. Using this approach we study also a reverse dynamic, which traces all the possible sources of influence of a given vertex, and which by a new "duality" relation allows to characterise the set of good pioneers.

6.23. Pioneers of Influence Propagation in Social Networks

With the growing importance of corporate viral marketing campaigns on online social networks, the interest in studies of influence propagation through networks is higher than ever. In a viral marketing campaign, a firm initially targets a small set of pioneers and hopes that they would influence a sizeable fraction of the population by diffusion of influence through the network. In general, any marketing campaign might fail to go viral in the first try. As such, it would be useful to have some guide to evaluate the effectiveness of the campaign and judge whether it is worthy of further resources, and in case the campaign has potential, how to hit upon a good pioneer who can make the campaign go viral.

In [43], we present a diffusion model developed by enriching the generalized random graph (a.k.a. configuration model) to provide insight into these questions. We offer the intuition behind the results on this model, rigorously proved in [38], and illustrate them here by taking examples of random networks having prototypical degree distributions — Poisson degree distribution, which is commonly used as a kind of benchmark, and Power Law degree distribution, which is normally used to approximate the real-world networks. On these networks, the members are assumed to have varying attitudes towards propagating the information. We analyze three cases, in particular — (1) Bernoulli transmissions, when a member influences each of its friend with probability p; (2) Node percolation, when a member influences all its friends with probability p and none with probability $1-p$; (3) Coupon-collector transmissions, when a member randomly selects one of his friends K times with replacement.

We assume that the configuration model is the closest approximation of a large online social network, when the information available about the network is very limited. The key insight offered by this study from a firm's perspective is regarding how to evaluate the effectiveness of a marketing campaign and do cost-benefit analysis by collecting relevant statistical data from the pioneers it selects. The campaign evaluation criterion is informed by the observation that if the parameters of the underlying network and the campaign effectiveness are such that the campaign can indeed reach a significant fraction of the population, then the set of good pioneers also forms a significant fraction of the population. Therefore, in such a case, the firms can even adopt the naïve strategy of repeatedly picking and targeting some number of pioneers at random from the population. With this strategy, the probability of them picking a good pioneer will increase geometrically fast with the number of tries.

6.24. Peer-to-Peer Networks

In [12], in collaboration with I. Norros (VTT, Finland) and F. Mathieu (Bell Labs), we propose a new model for peer-to-peer networking which takes the network bottlenecks into account beyond the access. This model can cope with key features of P2P networking like degree or locality constraints together with the fact that distant peers often have a smaller rate than nearby peers. Using a network model based on rate functions, we give a closed form expression of peers download performance in the system's fluid limit, as well as approximations for the other cases. Our results show the existence of realistic settings for which the average download time is a decreasing function of the load, a phenomenon that we call super-scalability.

6.25. Stability of the bipartite matching model

In [8], we consider the bipartite matching model of customers and servers introduced by Caldentey, Kaplan and Weiss (2009). Customers and servers play symmetrical roles. There are finite sets C and S of customer and server classes, respectively. Time is discrete and at each time step one customer and one server arrive in the system according to a joint probability measure μ on C× S, independently of the past. Also, at each time step, pairs of matched customers and servers, if they exist, depart from the system. Authorized em matchings are given by a fixed bipartite graph (C, S, E⊂ C × S). A matching policy is chosen, which decides how to match when there are several possibilities. Customers/servers that cannot be matched are stored in a buffer. The evolution of the model can be described by a discrete-time Markov chain. We study its stability under various admissible matching policies, including ML (match the longest), MS (match the shortest), FIFO (match the oldest), RANDOM (match uniformly), and PRIORITY. There exist natural necessary conditions for stability (independent of the matching policy) defining the maximal possible stability region. For some bipartite graphs, we prove that the stability region is indeed maximal for any admissible matching policy. For the ML policy, we prove that the stability region is maximal for any bipartite graph. For the MS and PRIORITY policies, we exhibit a bipartite graph with a non-maximal stability region.

6.26. Matchings on infinite graphs

Elek and Lippner (Proc. Am. Math. Soc. 138(8), 2939–2947, 2010) showed that the convergence of a sequence of bounded-degree graphs implies the existence of a limit for the proportion of vertices covered by a maximum matching. In [6], we provide a characterization of the limiting parameter via a local recursion defined directly on the limit of the graph sequence. Interestingly, the recursion may admit multiple solutions, implying non-trivial long-range dependencies between the covered vertices. We overcome this lack of correlation decay by introducing a perturbative parameter (temperature), which we let progressively go to zero. This allows us to uniquely identify the correct solution. In the important case where the graph limit is a unimodular Galton–Watson tree, the recursion simplifies into a distributional equation that can be solved explicitly, leading to a new asymptotic formula that considerably extends the well-known one by Karp and Sipser for Erdős-Rényi random graphs.

6.27. Double-hashing thresholds via local weak convergence.

A lot of interest has recently arisen in the analysis of multiple-choice “cuckoo hashing” schemes. In this context, a main performance criterion is the load threshold under which the hashing scheme is able to build a valid hashtable with high probability in the limit of large systems; various techniques have successfully been used to answer this question (differential equations, combinatorics, cavity method) for increasing levels of generality of the model. However, the hashing scheme analysed so far is quite utopic in that it requires to generate a lot of independent, fully random choices. Schemes with reduced randomness exists, such as “double hashing”, which is expected to provide similar asymptotic results as the ideal scheme, yet they have been more resistant to analysis so far. In [22], we point out that the approach via the cavity method extends quite naturally to the analysis of double hashing and allows to compute the corresponding threshold. The path followed is to show that the graph induced by the double hashing scheme has the same local weak limit as the one obtained with full randomness.

6.28. Convergence of multivariate belief propagation, with applications to cuckoo hashing and load balancing

[23] is motivated by two applications, namely generalizations of cuckoo hashing, a computationally simple approach to assigning keys to objects, and load balancing in content distribution networks, where one is interested in determining the impact of content replication on performance. These two problems admit a common abstraction: in both scenarios, performance is characterized by the maximum weight of a generalization of a matching in a bipartite graph, featuring node and edge capacities. Our main result is a law of large numbers characterizing the asymptotic maximum weight matching in the limit of large bipartite random graphs, when

the graphs admit a local weak limit that is a tree. This result specializes to the two application scenarios, yielding new results in both contexts. In contrast with previous results, the key novelty is the ability to handle edge capacities with arbitrary integer values. An analysis of belief propagation algorithms (BP) with multivariate belief vectors underlies the proof. In particular, we show convergence of the corresponding BP by exploiting monotonicity of the belief vectors with respect to the so-called upshifted likelihood ratio stochastic order. This auxiliary result can be of independent interest, providing a new set of structural conditions which ensure convergence of BP.

6.29. Bypassing correlation decay for matchings with an application to XORSAT

Many combinatorial optimization problems on sparse graphs do not exhibit the correlation decay property. In such cases, the cavity method remains a sophisticated heuristic with no rigorous proof. In [24], we consider the maximum matching problem which is one of the simplest such example. We show that monotonicity properties of the problem allows us to define solutions for the cavity equations. More importantly, we are able to identify the 'right' solution of these equations and then to compute the asymptotics for the size of a maximum matching. The results for finite graphs are self-contained. We give references to recent extensions making use of the notion of local weak convergence for graphs and the theory of unimodular networks.

As an application, we consider the random XORSAT problem which according to the physics literature has a 'one-step replica symmetry breaking' (1RSB) glass phase. We derive new bounds on the satisfiability threshold valid for general graphs (and conjectured to be tight).

6.30. Sublinear-Time Algorithms for Monomer-Dimer Systems on Bounded Degree Graphs

For a graph G, let $Z(G,\lambda)$ be the partition function of the monomer-dimer system defined by $\sum_k m_k(G)\lambda^k$, where $m_k(G)$ is the number of matchings of size k in G. In [27], we consider graphs of bounded degree and develop a sublinear-time algorithm for estimating $\log Z(G,\lambda)$ at an arbitrary value $\lambda > 0$ within additive error ϵn with high probability. The query complexity of our algorithm does not depend on the size of G and is polynomial in $1/\epsilon$, and we also provide a lower bound quadratic in $1/\epsilon$ for this problem. This is the first analysis of a sublinear-time approximation algorithm for a $\#P$-complete problem. Our approach is based on the correlation decay of the Gibbs distribution associated with $Z(G,\lambda)$. We show that our algorithm approximates the probability for a vertex to be covered by a matching, sampled according to this Gibbs distribution, in a near-optimal sublinear time. We extend our results to approximate the average size and the entropy of such a matching within an additive error with high probability, where again the query complexity is polynomial in $1/\epsilon$ and the lower bound is quadratic in $1/\epsilon$. Our algorithms are simple to implement and of practical use when dealing with massive datasets. Our results extend to other systems where the correlation decay is known to hold as for the independent set problem up to the critical activity.

6.31. Reconstruction in the Labeled Stochastic Block Model

The labeled stochastic block model is a random graph model representing networks with community structure and interactions of multiple types. In its simplest form, it consists of two communities of approximately equal size, and the edges are drawn and labeled at random with probability depending on whether their two endpoints belong to the same community or not.

It has been conjectured that this model exhibits a phase transition: reconstruction (i.e. identification of a partition positively correlated with the true partition into the underlying communities) would be feasible if and only if a model parameter exceeds a threshold.

In [25], we prove one half of this conjecture, i.e., reconstruction is impossible when below the threshold. In the converse direction, we introduce a suitably weighted graph. We show that when above the threshold by a specific constant, reconstruction is achieved by (1) minimum bisection, and (2) a spectral method combined with removal of nodes of high degree.

6.32. Spectrum Bandit Optimisation

In [26], we consider the problem of allocating radio channels to links in a wireless network. Links interact through interference, modelled as a conflict graph (i.e., two interfering links cannot be simultaneously active on the same channel). We aim at identifying the channel allocation maximizing the total network throughput over a finite time horizon. Should we know the average radio conditions on each channel and on each link, an optimal allocation would be obtained by solving an Integer Linear Program (ILP). When radio conditions are unknown a priori, we look for a sequential channel allocation policy that converges to the optimal allocation while minimizing on the way the throughput loss or *regret* due to the need for exploring sub-optimal allocations. We formulate this problem as a generic linear bandit problem, and analyze it first in a stochastic setting where radio conditions are driven by a stationary stochastic process, and then in an adversarial setting where radio conditions can evolve arbitrarily. We provide, in both settings, algorithms whose regret upper bounds outperform those of existing algorithms for linear bandit problems.

6.33. Randomized Consensus with Attractive and Repulsive Links

In [29], we study convergence properties of a randomized consensus algorithm over a graph with both attractive and repulsive links. At each time instant, a node is randomly selected to interact with a random neighbor. Depending on if the link between the two nodes belongs to a given subgraph of attractive or repulsive links, the node update follows a standard attractive weighted average or a repulsive weighted average, respectively. The repulsive update has the opposite sign of the standard consensus update. In this way, it counteracts the consensus formation and can be seen as a model of link faults or malicious attacks in a communication network, or the impact of trust and antagonism in a social network. Various probabilistic convergence and divergence conditions are established. A threshold condition for the strength of the repulsive action is given for convergence in expectation: when the repulsive weight crosses this threshold value, the algorithm transits from convergence to divergence. An explicit value of the threshold is derived for classes of attractive and repulsive graphs. The results show that a single repulsive link can sometimes drastically change the behavior of the consensus algorithm. They also explicitly show how the robustness of the consensus algorithm depends on the size and other properties of the graphs.

6.34. Continuous-time Distributed Optimization of Homogenous Dynamics

This paper explores the fundamental properties of distributed minimization of a sum of functions with each function only known to one node, and a pre-specified level of node knowledge and computational capacity. We define the optimization information each node receives from its objective function, the neighboring information each node receives from its neighbors, and the computational capacity each node can take advantage of in controlling its state. It is proven that there exist a neighboring information way and a control law that guarantee global optimal consensus if and only if the solution sets of the local objective functions admit a nonempty intersection set for fixed strongly connected graphs. Then we show that for any tolerated error, we can find a control law that guarantees global optimal consensus within this error for fixed, bidirectional, and connected graphs under mild conditions. For time-varying graphs, we show that optimal consensus can always be achieved as long as the graph is uniformly jointly strongly connected and the nonempty intersection condition holds. The results illustrate that nonempty intersection for the local optimal solution sets is a critical condition for successful distributed optimization for a large class of algorithms.

6.35. Two-target Algorithms for Infinite-Armed Bandits with Bernoulli Rewards

In [16], we consider an infinite-armed bandit problem with Bernoulli rewards. The mean rewards are independent, uniformly distributed over $[0, 1]$. Rewards 0 and 1 are referred to as a success and a failure, respectively. We propose a novel algorithm where the decision to exploit any arm is based on two successive targets, namely, the total number of successes until the first failure and until the first m failures, respectively, where m is a fixed parameter. This two-target algorithm achieves a long-term average regret in $\sqrt{2n}$ for a large parameter m and a known time horizon n. This regret is optimal and strictly less than the regret achieved by the best known algorithms, which is in $2\sqrt{n}$. The results are extended to any mean-reward distribution whose support contains 1 and to unknown time horizons. Numerical experiments show the performance of the algorithm for finite time horizons.

E-MOTION Project-Team

5. New Results

5.1. Perception and Situation Awareness in Dynamic Environments

5.1.1. Sensor Fusion for state parameters identification

Participants: Agostino Martinelli, Chiara Troiani.

5.1.1.1. General theoretical results

We continued to investigate the visual-inertial structure from motion problem by further addressing the two important issues of observability and resolvability in closed form. Regarding the first issue, we extended our previous results published last year on the journal of Transaction on Robotics [44] by investigating the case when the visual sensor is not extrinsically calibrated. In order to deal with this case, we must augment the state to be estimated by including all the parameters that characterize the extrinsic camera calibration, i.e., the six parameters that describe the relative transformation between the frame attached to the camera and the frame attached to the Inertial Measurement Unit (IMU). On the other hand, because of the larger size of the resulting state, it became prohibitive a direct application of the method that we introduced two years ago (see [43]) in order to discover the observability properties for this new state. For this reason, our first novel contribution during this year was the introduction of new methodologies able to significantly reduce the computational burden demanded by the implementation of the method in [43]. These methodologies have been published in [22] and a deeper description of their use is currently under revision on the journal Foundations and Trends in Robotics. The new results obtained by using these methodologies basically state that also the new six parameters that describe the camera extrinsic calibration are observable. Finally, we started a new research that in the literature is known as the Unknown Input Observability (UIO) and it is investigated by the automatic control community. We started this new research since we investigated the observability properties of the visual inertial structure from motion as the number of inertial sensors is reduced. Specifically, instead of considering the standard formulation, which assumes a monocular camera, three orthogonal accelerometers and three orthogonal gyroscopes, the considered sensor suit only consists of a monocular camera and one or two accelerometers. This analysis has never been provided before. A preliminary investigation seems to prove that the observability properties of visual inertial structure from motion do not change by removing all the three gyroscopes and one accelerometer. By removing a further accelerometer, if the camera is not extrinsically calibrated, the system loses part of its observability properties. On the other hand, as the camera is extrinsically calibrated, the system maintains the same observability properties as in the standard case. This contribution clearly shows that the information provided by a monocular camera, three accelerometers and three gyroscopes is redundant. Additionally, it provides a new perspective in the framework of neuroscience to the process of vestibular and visual integration for depth perception and self motion perception. Indeed, the vestibular system, which provides balance in most mammals, consists of two organs (the *utricle* and the *saccule*) able to sense the acceleration only along two independent axes (and not three). In order to analyze these systems with a reduced number of inertial sensors, we had to consider control systems where some of the inputs are unknown. Indeed, the visual-inertial structure from motion problem can be characterized by a control system where the inputs are known thanks to the inertial sensors. Hence, to deal with the visual-inertial structure from motion as the number of inertial sensors is reduced, we had to introduce a new method able to address the more general UIO problem. We believe that our solution to the UIO is general and this is the reason because we started this new research domain in control theory. Preliminary results are currently under revision on the journal Foundations and Trends in Robotics and we also plan to present them at the next ICRA conference. Regarding the second issue, i.e., the problem resolvability in closed form, a new simple closed form solution to visual-inertial structure from motion has been derived. This solution expresses the structure of the scene and the motion only in terms of the visual and inertial measurements collected during a short time interval. This allowed us to introduce deterministic algorithms able to simultaneously determine the structure of the scene together with the motion without the need for any initialization or prior knowledge. Additionally, the closed-form solution allowed us to identify the conditions under which the visual-inertial structure from motion has a finite number of solutions. Specifically, it is shown that the problem can have a unique solution, two distinct solutions or infinite solutions depending on the trajectory, on the number of point-features and on their arrangement in the 3D space and on the number of camera images. All the results have been published on the international journal of Computer Vision [15].

5.1.1.2. Applications with a Micro Aerial Vehicle

We introduced a new method to localize a micro aerial vehicle (MAV) in GPS denied environments and without the usage of any known pattern [26]. The method exploits the planar ground assumption and only uses the data provided by a monocular camera and an inertial measurement unit. It is based on a closed solution which provides the vehicle pose from a single camera image, once the roll and the pitch angles are obtained by the inertial measurements. Specifically, the vehicle position and attitude can uniquely be determined by having two point features. However, the precision is significantly improved by using three point features. The closed form solution makes the method very simple in terms of computational cost and therefore very suitable for real time implementation. Additionally, because of this closed solution, the method does not need any initialization. Results of experimentation show the effectiveness of the proposed approach.

We proposed a novel method to estimate the relative motion between two consecutive camera views, which only requires the observation of a single feature in the scene and the knowledge of the angular rates from an inertial measurement unit, under the assumption that the local camera motion lies in a plane perpendicular to the gravity vector [27]. Using this 1-point motion parametrization, we provide two very efficient algorithms to remove the outliers of the feature-matching process. Thanks to their inherent efficiency, the proposed algorithms are very suitable for computationally-limited robots. We test the proposed approaches on both synthetic and real data, using video footage from a small flying quadrotor. We show that our methods outperform standard RANSAC-based implementations by up to two orders of magnitude in speed, while being able to identify the majority of the inliers.

5.1.2. A new formulation of the Bayesian Occupancy Filter : an hybrid sampling based framework

Participants: Lukas Rummelhard, Amaury Nègre.

The Bayesian Occupancy Filter (BOF) is a discretized grid structure based bayesian algorithm, in which the environment is subdivised in cells to which random variables are linked. These random variables represent the state of occupancy and the motion field of the scene, without any notion of object detection and tracking, making the updating part of the filter an evaluation of the distribution of these variables, according to the new data acquisition. In the classic representation of the BOF, the motion field of each cell is represented as a neighborhood grid, the probability of the cell moving from the current one to another of the neighborhood being stocked in an histogram. If this representation is convenient for the update, since the potential antecedents of any cell is exactly determined by the structure, and so the propagation model is easily parallelizable, it also raises determinant issues :

- the structure requires the process rate to be constant, and a priori known.
- in the case of a moving grid, such as an application of car perception, many aliasing problems can appear, not only in the occupation grid, but in the motion fields of cells. A linear interpolation in 4-dimension field to fill each value of the histograms can quickly become unreasonable.
- to be able to match the slowest moves in the scene and the tiniest objects, the resolution of the grid and the motion histogram must be the high. On the other hand, since the system must be able to evaluate the speed of highly dynamic objects (typically, a moving car), the maximum encoded speed is to be high as well. This results in a necessary huge resolution grid, which prevent the system from being used with satisfying results on an embedded device. This huge grid is also mostly empty (most of the motion field histogram for a occupied cell will be empty). On top of that, the perception system being used to represent the direct environment of a moving car, the encoded velocity is a relative velocity, which implies, if we consider the maximal speed of a car to be V_max, to maintain

a motion field able to represent speeds from $-2 * V_{max}$ to $2 * V_{max}$. The necessity of such a sized structure is a huge limitation of practical use of the method.

Considering those limitations, a new way to represent the motion field has been developped. To do so, a new formulation of the BOF has been elaborated. This new version allow to introduce in the filter itself a distinction between static and dynamic parts, and so adapt the computation power. The main idea of this new representation is to mix two forms of sampling : a uniform one, represented as a grid, for the static objects and the empty areas, and a non uniform one, based on particles drawn from dynamic regions. The motion field in a cell will be represented as a set of samples from the distribution for values which are not null, and a weight given to the static hypothesis. The use of a set of samples to represent the motion field leads to a important decrease of the needed memory space, as well as the classification between dynamic objects and static objects or free areas. In the updating process, the antecedent of a cell can be either from the static configuration or from the dynamic configuration, which are both way easier to project in the new reference frame of the moving grid. The first results are stimulating, in term of occupancy evaluation and mostly in term of velocity prediction, being way more accurate and responsive than the older version. Those improvements will soon be presented in detail in upcoming papers, one being currently in the process of redaction.

5.1.3. DATMO

Participants: Dung Vu, Mathias Perrollaz, Amaury Nègre.

In the current work, we have been developing a general framework for tracking multiple targets from lidar data.

In the past decades, multiple target tracking has been an active research topic. When object observations are known, object tracking becomes a data association (DA) problem. Among popular DA methods, multiple hypothesis tracking (MHT) is widely used. MHT is a multi-frame tracking method that is capable of handling ambiguities in data association by propagating hypotheses until they can be solved when enough observations are collected. The main disadvantage of MHT is its computational complexity since the number of hypotheses grows exponentially over time. The joint probabilistic data association (JPDA) filter is more efficient but prone to make erroneous decision since only single frame is considered and the association made in the past is not reversible. Other sequential approaches using particle filters share the same weakness that they cannot reverse time back when ambiguities exist. All DA approaches mentioned above requires a strong assumption of one-to-one mapping between targets and observations which is usually violated in real environments. For instance, a single object can be seen by several observations due to occlusion, or multiple moving objects can be merged into a single observation when moving closely.

In the research, we propose a new data association approach that deals with split/merge nature of object observations. In addition, our approach also tackles ambiguities by taking into account a sequence of observations in a sliding window of frames. To avoid the high computational complexity, a very efficient Markov Chain Monte Carlo (MCMC) technique is proposed to sample and search for the optimum solution in the spatio-temporal solution space. Moreover, various aspects including prior information, object model, motion model and measurement model are explicitly integrated in a theoretically sound framework.

5.1.4. Visual recognition for intelligent vehicles

Participants: Alexandros Makris, Mathias Perrollaz, Christian Laugier.

We have developped an object class recognition method. The method uses local image features and follows the part-based detection approach. It fuses intensity and depth information in a probabilistic framework. The depth of each local feature is used to weigh the probability of finding the object at a given distance. To train the system for an object class, only a database of images annotated with bounding boxes is required, thus automatizing the extension of the system to different object classes. We apply our method to the problem of detecting vehicles from a moving platform. The experiments with a data set of stereo images in an urban environment show a significant improvement in performance when using both information modalities.

In 2013, the method has been published in IEEE Transactions on Intelligent Transportation Systems [14].

../../../../projets/e-motion/IMG/pbog0_cam.jpg

../../../../projets/e-motion/IMG/pbog1_cam.jpg

(a) Camera image, two-way street

(b) Camera image, highway

../../../../projets/e-motion/IMG/pbog0_occ_speed.jpg

../../../../projets/e-motion/IMG/pbog

(c) Occupancy and velocity result, two-way street

(d) Occupancy and velocity result, highway

Figure 1. Results of the new algorithm : the free cells are black, the occupied ones are white. Unknown areas are grey. When a cell is seen as dynamic, a vector representing the average of the associated motion field is drawn in red.

5.1.5. Experimental platform for road perception

5.1.5.1. Experimental platform material description

Our experimental platform for road perception is shown in Figure 2 . This platform is a commercial Lexus car LS600h equipped with a variety of sensor, including two IBEO Lux lidars placed toward the edges of the front bumper, a TYZX stereo camera plus a high resolution color camera situated behind the windshield, and an Xsens MTi-G inertial sensor with GPS. To do online data computation and data acquisition, a standard computer is located on the back of the car.

This platform allows us to conduct experimentation and data acquisition in various road environments (country roads, downtown and highway), at different time of the day, with various driving situations (light traffic, dense traffic, traffic jams).

../../../../projets/e-motion/IMG/plateforme_lexus.png

Figure 2. Lexus LS600h car equipped with two IBEO Lux lidars, a stereo plus a monocular camera, and an Xsens MTi-G inertial sensor with GPS.

5.1.5.2. Software architecture

The perception and situation awareness software architecture is integrated in the ROS framework. ROS (http://www.ros.org) is an open source robotics middleware designed to be distributed and modular. For the Lexus platform, we developed a set of ROS module for each sensor and for each perception component. Each perception module can be dynamically connected with the required drivers or other perception modules. The main architecture of the perception components is illustrated on Figure 3 .

../../../../projets/e-motion/IMG/architecture_lexus.png

Figure 3. Architecture of the main perception components in the Lexus platform.

5.1.6. Software and Hardware Integration for Embedded Bayesian Perception

Participants: Mathias Perrollaz, Christian Laugier, Qadeer Baig, Dizan Vasquez, Lukas Rummelhard, Amaury Nègre.

The objective of this recently started research work is to re-design in a highly parallel fashion our Bayesian Perception approach for dynamic environments (based on the BOF concept), in order to deeply integrate the software components into new multi-processor hardware boards. The goal is to miniaturize the software/hardware perception system (i.e., to reduce the size, the load, the energy consumption and the cost, while increasing the efficiency of the system).

For supported this research, we began to work in the “Perfect” project. This project, included in the IRT-Nano program, involves the CEA-LETI DACLE lab and ST-Microelectronics. Perfect is focusing onto the second integration objectives (6 years) and the development of integrated open platforms in the domain of transportation (vehicle and infrastructure) and in a second step in the domain of health sector (mobility of elderly and handicapped people, monitoring of elderly people at home. . .). The objective of e-Motion in this project is to transfer and port its main Bayesian perception modules from traditional computing system to embedded low-power multi-processors board. The targeted board is a STHorm from ST Microelectronics which is has a many-core architecture with a very low consumption. In 2013 we worked with the CEA to obtain a first implementation of the Bayesian occupancy grid filter on STHorm. Those preliminary results demonstrated the feasibility of the concepts but highlighted some key points to improve such as the memory footprint we need to reduce to obtain real-time accurate results.

../../../../projets/e-motion/IMG/BOF_Integration.png

Figure 4. First objective for software/hardware of Bayesian perception : developing and using multiple processor board from mobile technologies. The approach will be validated with real demonstrators..

5.2. Dynamic Change Prediction and Situation Awareness

5.2.1. Vision-based Lane Tracker

Participants: Mathias Perrollaz, Amaury Nègre.

In the field of vehicle risk assessment system, vehicle to road localization is an essential information to predict drivers behaviors as well as collision risk evaluation. To achieve this task, we have developed a vision based lane tracker to estimate the geometry of the lane using the line markers. Previous development was based on a Monte-Carlo particle filter to estimate simultaneously the road plane orientation, the lane curvature and the camera position. To perform the particle evaluation, the algorithm first process a ridge extraction of the camera image and then projected the left and right marker line represented by the particle on the ridge image.

The first improvement of the tracker consists on dynamically adapting the scale of the ridge filter to improve the efficiency and the precision of the particle evaluation. The second improvement is capability the perform multi-lane tracking for example in highway environment. To solve this problem, the position of the lane is added to the state and the particle evaluation consider the total number of line marker (a-priori known). Figure 5 shows the results of the lane tracker program on a highway environment.

5.2.2. Vision-based Lane Change Prediction

Participants: Suryansh Kumar, Dizan Vasquez, Mathias Perrollaz, Stephanie Lefevre, Amaury Nègre, Maiwen Gault.

For both Advanced Driving Assistance Systems and Autonomous Vehicles, it is very important to have the capability of predicting and understanding the driver's behavior. This work addresses this subject in a bottom-up fashion by first detecting low-level "atomic" maneuvers which can be used as a building block for more complex behavior. Concretely, we have developed a learning-based approach that uses lane tracking data to predict lane changes.

Most works in the literature address this as a classification problem, and often use some version of Support Vector Machines (SVM) to solve it. The problem with this approach is that it is sensitive to noise and can yield high-frequency oscillations in the obtained predictions, moreover, they do not provide any information concerning the Time To Change (TTC). Other approaches use a filtering approaches, using Hidden Markov Models (HMM), for example. Although they produce smoother predictions and, in some cases, even a TTC estimate, some studies [33] suggest that HMM-based approaches are less accurate than those based on SVMs.

Our work combines the advantages of both approaches in a hierarchical fashion. First, lane tracking data (i.e. lane-relative yaw, lane-relative lateral position and their fist-order derivatives) are used as an input of a multi-class SVM. Then, the Bradley-Terry model is used to translate the SVM output into a probability which is used as the observation model of a Bayesian filter (Fig. 6).

This work has been published in the Intelligent Vehicles conference [18]. Since then, we have worked on an improved HMM-only approach which addresses the shortcomings of similar approaches by using a continuous observation model. In our preliminary experiments, this approach leads to improved predictions over the hybrid one.

5.2.3. Feature-based human behavior modeling

Participants: Suryansh Kumar, Yufeng Yu, Dizan Vasquez.

When it comes to modeling and learning complex human behavior, the preferred approach in the literature is to try to learn the typical motion patterns that people or vehicles get engaged into in a given environment. This has, however, a major drawback: the learned patterns only apply to the environment where they have been learned. This means that, for a new environment, previous knowledge cannot be used and patterns should be learned from scratch.

This situation has been recently addressed [49] by a family of approaches which rely on two complementary hypotheses:

- *Agents behave like planners*. The idea is that when people or other intelligent agents move around, they plan in order to minimize a cost function. Thus, if this function is known, it fully determines the agent's behavior.

../../../../projets/e-motion/IMG/lane_tracker_camera.jpg

(a) Camera image

../../../../projets/e-motion/IMG/lane_tracker_particles.jpg

../../../../projets/e-motion/IMG/lane_tracker_r

(b) Particles cloud

(c) Lane tracker result

Figure 5. Visual Particle based lane tracking.

../../../../projets/e-motion/IMG/lane_change-hybrid.png

Figure 6. Overview of our hybrid lane change prediction approach.

- *The cost is a function of local features*. This extends the previous hypothesis by assuming that the cost function does not only depends on the agent's state but also in a number of external features (e.g. local people density, nearby traffic signs). Since the cost function depends on the features, it is possible to compute it even for previously unseen environments, as long as they contain the same kind of features.

Under these assumptions, the problem becomes that of learning the unknown cost function by observing how people move. This is often called *apprenticeship learning* and, when the underlying planning model is a Markov Decision Process, *inverse reinforcement learning*.

We have been working on a software library and evaluation testbed for different features and cost-function learning algorithms. We have conducted, in collaboration with the University of Freiburg, a first round of experiments concerning people moving in crowds. The results will be the subject of a submission to the IROS 2014 conference.

../../../../projets/e-motion/IMG/torcs.png

Figure 7. The TORCS racing simulator.

We have also been working on an application for intelligent vehicles and ADAS. As a first step, we have developed a ROS interface for the TORCS racing simulator, as well as a road simulation using the same platform (Fig. 7). This work has been the product of a collaboration with the Beijing University and IIIT Hyderabad.

5.2.4. Safety applications at road intersections for connected vehicles

Participants: Stéphanie Lefèvre, Christian Laugier.

From a safety perspective, road intersections are the most dangerous areas in the road network. They are also the most complex. Because of the extended situational awareness that they provide, wireless vehicular communications (or Vehicle-to-X communications, V2X) could greatly reduce the rate of intersection accidents. However, numerous research challenges remain before the full use of this technology can be achieved. A PhD was started on this topic in 2009 in collaboration with Renault, and was successfully defended in 2012 [42]. The purpose was to formulate and develop a probabilistic reasoning framework which would allow combining the information shared by the vehicles to estimate the situation and the associated risk as a vehicle negotiates an intersection. The first contribution of the PhD was to model the motion of vehicles using a Dynamic Bayesian Network where the maneuvers of different vehicles influence each other via an "expected maneuver". This "expected maneuver" represents what a driver is expected to do given the state of the other vehicles in the area and the traffic rules which apply at the intersection. Thanks to the use of a probabilistic framework, uncertainties related to sensor errors and interpretation ambiguities are handled. The second contribution was a novel approach to risk estimation based on the comparison between what drivers intend to do and what they are expected to do. The reasoning is carried out by performing inference on the Dynamic Bayesian Network introduced earlier, using a particle filter. The approach was validated with field trials using Renault passenger vehicles equipped with vehicle-to-vehicle wireless communication modems [41], and in simulation [40]. The results show that the algorithm is able to detect dangerous situations early and complies with real-time constraints. We also developed a theoretical extension of the model to generalize it to arbitrary traffic situations [29]. This work is still ongoing thanks to an Inria@SiliconValley fellowship granted to S. Lefevre at the end of her PhD. Since January 2013 she is working in the Teleimmersion group at the University of California Berkeley, as a postdoctoral researcher. The research conducted there lead to two new developments on the topic of "Safety applications at road intersections for connected vehicle".
The first development concerns probabilistic decision making for Collision Avoidance (CA) systems. In the processing chain of a CA system, the "Decision making" module follows the "Risk assessment" module. The research done during the PhD stopped at the "Risk assessment" module, and we now address the challenges present in the "Decision making" module. We identified two main challenges:

1. The first one is that the decision making module has to make decisions based on uncertain knowledge. Sensors provide noisy measurements, digital maps contain errors, and interpreting a vehicle's motion in terms of driver intention is uncertain. These uncertainties propagate to the risk assessment module and to the decision making module, but the latter is still required to make a decision from that uncertain knowledge.
2. The second one is that the timing of interventions is critical. If an intervention is triggered at a time when the uncertainty about the occurrence of a collision is too large, there is a chance that it will end up being a false alarm. High false alarm rates are detrimental to the driver acceptance of safety systems and can lead to the user losing trust in the system. On the other hand, if the system waits until the last moment (certainty about the occurrence of a collision) to trigger an intervention, it might be too late to avoid the accident.

In recent work [20] we proposed to introduce the possibility for a CA system to postpone making a decision. Our objective is to implement the fact that in some situations the new observations obtained by waiting will reduce the uncertainty about the occurrence of a collision, therefore the decision will be more reliable if it is made later using this additional information. The important question to solve is whether the potential gain brought by the additional information outweighs the cost of waiting. In order to answer this question, our decision making approach runs a *preposterior analysis* to determine the expected value and cost of the additional information. The value of the additional information can be quantified by means of the Expected Value of Sample Information (EVSI). It corresponds to the additional expected payoff possible through knowledge of the additional information and is computed by subtracting the expected costs of deciding with and without additional information. The cost of the additional information is quantified by means of the Expected Cost of Waiting (ECW). It is computed as the difference between the probability that the CA system will be able to avoid the potential collision if it intervenes now and if it intervenes at time t+1. Our decision making strategy is to postpone the decision making process to time t+1 if and only if the EVSI is positive and the ECW is null. The algorithm was tested in simulation at a two-way stop intersection for collision scenarios

and no-collision scenarios involving two vehicles. A comparative evaluation with a decision making strategy which does not allow postponing decisions showed that our approach generates fewer false alarms and avoids as many collisions. These results were published at the conference IEEE IROS'13 [20] and a patent application was filed with UC Berkeley [32].
The second development addresses the relationship between privacy strategies for V2X and safety applications which rely on V2X. User privacy is a requirement for wireless vehicular communications, and a number of privacy protection strategies have already been developed and standardized. In particular, methods relying on the use of temporary pseudonyms and silent periods have proved their ability to confuse attackers who would attempt to track vehicles. However, these privacy protection schemes are not without consequences for safety applications. Such applications make decisions (e.g. warning drivers of an upcoming danger) based on their current estimation of the state of the real world, and this representation is created from the information contained in beacons received from other vehicles. Therefore, interruptions in the transmission of information will impact the decision-making process. If a silent period is scheduled to start at a safety-critical moment, it could result in safety systems not intervening when they should have, namely a "missed intervention". From a user and safety perspective, this is not acceptable. In this work we address this issue and evaluate the impact of pseudonym change strategies on V2X-based Intersection Collision Avoidance (ICA) system. We use the ICA system developed during the PhD and simulated 3 different privacy protection strategies:

1. The "*Fixed ID*" strategy assigns a fixed pseudonym to a vehicle for the entire duration of a trip (i.e. a new pseudonym is assigned to the vehicle every time it starts). Testing this case gives us a reference for how well the collision avoidance system performs when there is no pseudonym change and no silent period during a trip, which is what was assumed in our PhD work.
2. The "*Baseline*" strategy follows the recommendations of the SAE J2735 standard for V2X communications. Pseudonyms are changed every 120 seconds and are followed by a silent period of random duration.
3. The "*Adaptive*" strategy is a modified version of the *Baseline* strategy where the risk of the situation is taken into account to decide whether or not a vehicle should be allowed to change pseudonym at time t. It relies on the estimation of the current ability of the collision avoidance system to keep the vehicle on a collision-free trajectory. The idea here is to authorize a pseudonym change and silent period only if it will not affect the performance of the safety application.

Simulations were conducted using the same simulator and the same scenarios as the tests run during the PhD. The performance of the three privacy strategies was evaluated both in terms of privacy and in terms of successful interventions of the ICA system. The results show that the ICA application requires silent periods to be shorter than two seconds in order to operate correctly in conjunction with the SAE J2735 standard. They also indicate that the addition of simple rules which authorize or not a pseudonym change depending on the context leads to major safety improvements compared to the SAE J2735 standard alone (see Figure 8). These results, which were published at the conference IEEE VNC'13 [21], highlight the necessity of a joint design. That is, the requirements of safety applications should be taken into account when designing privacy strategies, and pseudonym change schemes should be accounted for when designing safety applications which rely on V2X communications. This collaboration is necessary in order to ensure that vehicular communications and safety applications do not neutralize each other, but instead, work together toward safer roads.

5.2.5. Guidance for Uncertain shooting domain

Participant: Emmanuel Mazer.

This study is made in collabortion with MBDA (Monsieur Le Menec) and Probayes (Monsieur Laurent Saroul) under the ITP framework financed by the british MOD and the french DGA.

From N°51 of Vector Magazine (copyright MBDA) *It's an enduring question facing those in military conflict: 'when do you pull the trigger?' However, for pilots in air combat there is an added question: 'Once you've pulled the trigger, when do you break the link with your missile?' A new answer to that problem was a highlight of the MCM-ITP conference at Lille in May (see Vector 50). Entitled 'Guidance in Uncertain Shooting Domains', this joint project between MBDA and French company Probayes has produced a set of algorithms*

../../../../projets/e-motion/IMG/Image_1.png

Figure 8. Percentage of missed interventions, avoided collisions, and failed interventions as a function of the duration of the silent period for the Baseline strategy (left) and the Adaptive strategy (right).

to help a pilot decide when to break a telemetry link. It's a decision fraught with danger. As pilots approach a target, the longer they keep the link, the greater the chance of their missile finding its target. However, closing in on the target usually means entering the enemy's own kill zone. So, a calculation of the best trade-off between mission success and pilot safety needs to be made.

"At the moment, when a pilot is engaging an enemy aircraft, he's obliged to do sums in his head," explained Graham Wallis, MBDA UK's Chief Technologist. "What we're trying to do is to take that away, and hide it behind a probability display." The problem arises because a seeker's scanning range is often far shorter than the maximum travelling distance, or kinematic range, of a missile. Though less of an issue at short range, it's a problem for medium-range weapons – where the target is likely to manoeuvre beyond the initial seeker scan area; hence the need for guidance from the pilot, who may have the target in radar sight.

SUCCESS PROBABILITY Stepping back, there are two sides to the firing equation. As customers require an air-to-air missile that will find its target almost 100 of the time, MBDA provides distance data (known as Launch Success Zone tables) to achieve that. They are understandably conservative. Clearly, though, the final decision to break the datalink can only be taken by the pilot, who is also making mental calculations to minimise his or her exposure to deadly risk. GUSD effectively offers a 'probability meter' to help reconcile those two imperatives.

Physically, GUSD could be a display with four bars and the circle of a pie-chart. Each bar represents the probability in percentage terms of an enemy pilot adopting one of four typical behaviours during air combat: flying head-on to attack; turning tail and heading for home, and either turning left or right – along with the probabilities that the enemy has just launched its own missile. The pie chart gives a single percentage – of the MBDA missile's chances of hitting its target. "The figures displayed are effectively our computer trying to read the mind of the enemy pilot," said Graham Wallis, whose team also drew on the experience of MBDA's former air force pilots.

QUICK CALCULATIONS Not surprisingly, the computations behind GUSD are hugely complex. The main input is the realtime radar tracking data of the enemy aircraft, although other elements such as seeker acquisition data, missile dynamics and the missile's inertial navigation errors are also included. A set of algorithms (and their associated mathematical methods, see box) then process this input – with the three key algorithms covering target behaviour and identification (Hidden Markov Models); the generation of bundles of trajectories (Markov Chains and Monte Carlo techniques) and trajectory collision checking (R-Tree).

Currently at Technology Readiness Level 3, GUSD's future depends on being incorporated into a programme. The generic data used to date would be replaced by real missile and seeker data, requiring GUSD to move out of the open forum of MCM-ITP and into the area covered by defence secrecy. "Implementation of research and technology is particularly difficult when it comes to jet fighters," Graham Wallis declared. "But I'm optimistic it will get into a future upgrade of current aircraft, and could even replace the Launch Success Zone tables as a firing cue for pulling that trigger."

MATHEMATICAL MODELS Named after 19th century Russian mathematician Andrey Markov, the Markov Chain uses a set of rules to predict what will happen next in a situation, when all the variables are known. In a game of bridge, if all the hands were known to all players and they had no discretion over which card to play next, every card laid down could be predicted as a Markov Chain. With a Hidden Markov Model, a player doesn't know what an opponent holds but can infer that, for example, they no longer have any trump cards because of how they have played their last hand. The Model takes multiple states e.g. 'Opponent Has Trumps' or 'Opponent Doesn't Have Trumps' and establishes a percentage weighting of probability to each – which can then be used for a decision over the best card to play. Though it is only an analogy, GUSD's algorithms use the mass of input data to infer the probability of an enemy pilot's tactical moves. R-trees are a way of handling multi-dimensional information - in GUSD's case by using geometric models that drastically cut the computation load.

OTHER APPLICATIONS Even before its own future is decided, GUSD could already lead to a successor MCM-ITP project with Probayes in the area of mission planning for long-range missiles. Called Rapid Mission

Planning and Rehearsal, it could automate much of what is currently a labour-intensive process and to provide the customer with different options for mission routes – each with its own probability rating for success.

../../../../projets/e-motion/IMG/Guidance.jpg

Figure 9. artistic view of GUSD (copyright MBDA)

5.3. Human Centered Navigation in the physical world

5.3.1. Social Mapping

Participants: Panagiotis Papadakis, Anne Spalanzani, Christian Laugier.

With robots technology shifting towards entering human populated environments, the need for augmented perceptual robotic skills emerges that complement to human presence. In this integration, perception and adaptation to the implicit human social conventions plays a fundamental role. Toward this goal, we introduce a novel methodology to detect and analyse complex spatial interactions of multiple people and encode them in the form of a social map, whose structure is obtained by computing a latent space representation of human proxemic behaviour. We accomplish this by appointing to humans distinct, skew-normal density functions that quantify social sensitivity and by using them in the sequel to induce a training set for regressing a collective density function of social sensitivity (see fig. 10). Finally, we extract level-sets of constant social sensitivity levels within the social map by which we can effectively and efficiently analyse individual as well as shared interaction zones of varying shape and size. Extensive experiments on human interaction scenarios demonstrate the feasibility and utility of the proposed approach in diverse conditions and promote its application to social mapping of human-populated environments. This work was published at IROS [23] and submitted to RAS journal.

../../../../projets/e-motion/IMG/SocialMap.png

Figure 10. A representative example of mapping human social interactions using the proposed methodology. The regressed global sociality density along with isocontours of constant social comfort are superimposed on the corresponding scene.

5.3.2. Goal oriented risk based navigation in dynamic uncertain environment

Participants: Anne Spalanzani, Jorge Rios-Martinez, Arturo Escobedo-Cabello, Procopio Silveira-Stein, Gregoire Vignon, Alejandro Dizan Vasquez Govea, Christian Laugier.

Navigation in large dynamic spaces has been adressed often using deterministic representations, fast updating and reactive avoidance strategies. However, probabilistic representations are much more informative and their use in mapping and prediction methods improves the quality of obtained results. Since 2008 we have proposed a new concept to integrate a probabilistic collision risk function linking planning and navigation methods with the perception and the prediction of the dynamic environments [36]. Moving obstacles are supposed to move along typical motion patterns represented by Gaussian Processes or Growing HMM. The likelihood of the obstacles' future trajectory and the probability of occupation are used to compute the risk of collision. The proposed planning algorithm, call RiskRRT (see Figure 11 for an illustration), is a sampling-based partial planner guided by the risk of collision. Results concerning this work were published in [37] [38] [39]. [47] and [48]. In 2013, Jorge Rios defended his phD on this topic. We obtained an Inria ADT to optimize and share the RiskRRT algorithm.

../../../../projets/e-motion/IMG/grow2.png

Figure 11. Social navigation example. RiskRRT selected a plan (red line) to the goal (red arrow).

5.3.3. Navigation Taking Advantage of Moving Agents

Participants: Procopio Silveira-Stein, Anne Spalanzani, Christian Laugier.

In this work, we proposes a different form of robotic navigation in dynamic environments, where the robot takes advantage of the motion of pedestrians, in order to improve its own navigation capabilities. The main idea is that, instead of treating persons as dynamic obstacles that should be avoided, they should be treated as special agents with an expert knowledge of navigating in dynamic scenarios. To benefit from the motion of pedestrians, this work proposes that the robot select and follow them, so it can move along optimal paths, deviate from undetected obstacles, improve navigation in densely populated areas and increase its acceptance by other humans. To accomplish this proposition, novel approaches are developed in the area of leader selection, where two methods are explored. The first uses motion prediction approaches while the second uses a machine learning method, to evaluate the leader quality of subjects, which is trained with real examples. Finally, the leader selection methods are integrated with motion planning algorithms and experiments are conducted in order to validate the proposed techniques. One of the most relevant application is navigation among crowds. Figure 12 illustrates the concept.

../../../../projets/e-motion/IMG/LeaderFollow.png

Figure 12. Navigation among crowds. The robot (rectangle) needs to reach the right side of the corridor. 2 groups of people (one in yellow, the other in blue) are crossing in this corridor. A classical motion planning would not find a path. The robot choses a leader (represented by a green circle around a person) and follows him.

The work is published in [25] [24] and Procopio Stein defended his phD the 11th of december 2013 at the Aveiro University (phD co-directed by Anne Spalanzani and Vitor Santos).

5.3.4. Autonomous Wheelchair for Elders Assistance

Participants: Arturo Escobedo-Cabello, Gregoire Vignon, Anne Spalanzani, Christian Laugier.

The aging of world's population is bringing the need to provide robotic platforms capable to assist elder people to move [46]. It is necessary that such transportation is reliable, safe and comfortable. People with motor disabilities and elders are expected to benefit from new developments in the field of autonomous navigation robotics.

Autonomously driven wheelchairs are a real need for those patients who lack the strength or skills to drive a normal electric wheelchair. The services provided by this kind of robots can also be used to provide a service of comfort, assisting the user to perform difficult tasks as traversing a door, driving in a narrow corridor etc.

Simple improvements of the classical powered wheelchair can often diminish several difficulties while driving. This idea of comfort has emerged as a design goal in autonomous navigation systems, designers are becoming more aware of the importance of the user when scheming solution algorithms. This is particularly important when designing services or devices intended to assist people with some disability. In order for the robot to have a correct understanding of the intention of the user (when moving around) it is necessary to create a model of the user that takes into account his habits, type of disability and environmental information. The ongoing research project is centered in the understanding of the intentions of the user while driving an autonomous wheelchair, so that we can use this information to make this task easier.

During the last two years the work was centered in the improvement of the usability of the system. A review of the state of the art in user's intention estimation algorithms was made and a new model to infer the intentions of the user in a known environment was presented [34], [35].

The algorithm models the intention of the user as 2D topological goals in the environment. Those places are selected according to how frequently they are visited by the user (user habits). The system was designed so that the user can give orders to the wheelchair by using any type of interface, as long as he can show the direction of the intended movement (joystick, head tracking, brain control, etc). As shown in figure13 , the chosen approach uses a Dynamic Bayesian Network to model and infer the intentions. The main contribution of this work is to model the intention of the user as topological goals instead of normal trajectory-based methods, therefore the model is simpler to deal with.

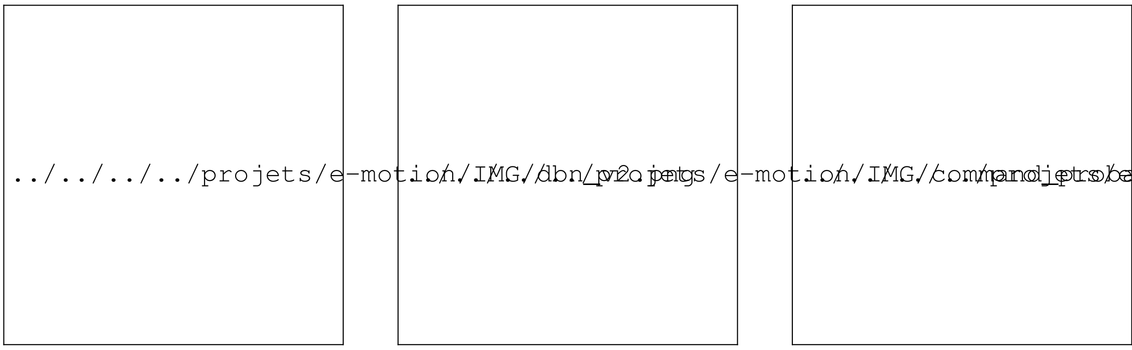


Figure 13. User's Intention Estimation Algorithm Left:User's intention model. The Bayesian network used to estimate the current user's intended; Center: The probability value for a given command C_t (big arrow) is proportional to the angle a_i formed respect to each goal g_i in the environment. Right: The user is looking to the left (in the direction of his desired goal). Once that the user's intention estimation system computes the goal with the highest probability, the autonomous navigation module plans the path and controls the movement of the wheelchair to take the user to the destination.

In 2013 the results of the user intention destination method were published in the IROS conference [16]. The presented work covered the following aspects:

User Intention Estimation: Two different methods to drive the wheelchair were compared, a semi-autonomous and a manual mode. In semi-autonomous mode the user's intention is estimated from the position of the face and the wheelchair takes care of all the planning to arrive there while avoiding obstacles. In manual mode the wheelchair is driven using the face without assistance from the robotic controller.

The Bayesian estimator shown in Fig.13 was used to do the inference of the desired destination of the user in semi-autonomous mode. The user's intention was modeled as a set of destinations commonly visited by the user and the task consisted in finding the destination targeted by the user.

Interfaces: People with motor disabilities and elders often have problems using joysticks and other standard control devices. Under this consideration our experimental platform was equipped with different types of user-interfaces to provide a multi-modal functionality as described in [35]. A face pose interface allows to control the wheelchair's motion by changing the face direction, while voice recognition interface is used to guarantee an adequate control of the wheelchair for those commands that otherwise would be difficult to give by only using the face (Stop, start, etc). This exploitation of more natural and easy-to-use human machine interfaces was one of the main contributions of the work presented in [16].

Experimental evaluation: Experiments were done in the hall of the Inria Rhône-Alpes laboratory. People in the scene were tracked to detect the most visited destinations in this setting (red circles). Those typical destinations were then placed in the map used by the robotic wheelchair. Each destination has a probability value related to the number of times that it is visited by people. The extracted typical destinations and related probabilities are used as prior knowledge when inferring the user's desired destination. In Fig.14 (Right) the spheres represent the typical destinations placed in the internal map o the wheelchair and the size of the sphere represents the probability of being the desired destination of the user in the wheelchair given it's position in the map and direction of the face (blue arrow).

To evaluate the performance of the method different persons were asked to drive the wheelchair in both “manual” and “semi-autonomous” mode. The trajectories of followed by the wheelchair were recorded and evaluated in [16], In Fig.15 one sample of those trajectories is presented where it can be appreciated how those trajectories executed with assistance of the robot are considerably softer than those obtained in manual mode as explained in [16].

Human aware navigation: Current work is being done in the construction of a social cost-map that is able to work with different open source path planning algorithms. This plug-in was developed using the method presented by Rios in [45].

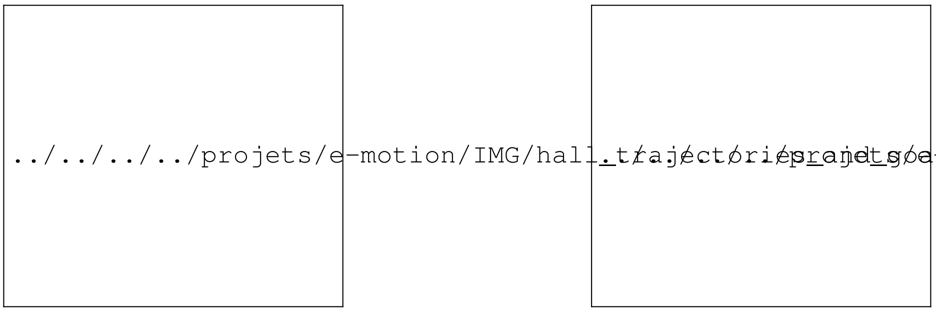


Figure 14. Left: The users of a normal environment move between typical destinations that can be learned. Right: The typical destinations marked in the map used by the wheelchair, The probability for each destination given the position of the wheelchair and direction of the command is proportional to the size of the sphere.

5.3.5. Bayesian modelling to implement and compare different theories of speech communication

Participants: Raphael Laurent, Pierre Bessière, Julien Diard, Jean-Luc Schwartz.

A central issue in speech science concerns the nature of representations and processes involved in communication. The search for phoneme or syllable specific invariants led to three major sets of approaches: motor, auditory and perceptuo-motor theories. They have been widely argued for and against, but the theoretical debate appears to be stagnating. It is our belief that computationnal models designed within a rigorous mathematical framework may allow to put forward new arguments to support either theory, and new ideas for experiments to be carried out on human subjects.

We have designed an integrative Bayesian model which allows to study auditory, motor and perceptuo-motor aspects of speech production and perception. This model was tested on perception tasks on evaluation corpora with more and more variability compared to the learning corpus. This showed a really high robustness of the purely motor model, which contained more information that it is the case in practise, due to unrealistic learning methods. The work was then focused on more realistic learning algorithms, where speech motor gestures are unsupervisedly learned through imitation, by generating motor gestures trying to reach auditory targets, and memorising the acoustics corresponding to these motor commands. This work was published in 2013 [19]. Raphael Laurent defended his phD in 2013.

5.3.6. Bayesian computing

Participants: Emmanuel Mazer, Pierre Bessière.

A book and the compagnon software on bayesian programming have been released this year :

The book “Bayesian Programming” is available at http://www.crcpress.com/product/isbn/9781439880326 and the associated software at http://www.probayes.com/fr/Bayesian-Programming-Book/

../../../../projets/e-motion/IMG/trajectories.png

Figure 15. Experimental evaluation of the user's intention method. Some samples of the resulting trajectories are presented. (a) and (c) show the results when using the assistance of the user's intention estimation system. (b) and (d) were achieved by driving the wheelchair using the face without any assistance. Here we can it can be appreciated how those trajectories executed with assistance of the robot are considerably softer than those obtained in manual mode.

Features

Presents a new modeling methodology and inference algorithms for Bayesian programming Explains how to build efficient Bayesian models Addresses controversies, historical notes, epistemological debates, and tricky technical questions in a dedicated chapter separate from the main text Encourages further research on new programming languages and specialized hardware for computing large-scale Bayesian inference problems Offers an online Python package for running and modifying the Python program examples in the book

Summary

Probability as an Alternative to Boolean Logic While logic is the mathematical foundation of rational reasoning and the fundamental principle of computing, it is restricted to problems where information is both complete and certain. However, many real-world problems, from financial investments to email filtering, are incomplete or uncertain in nature. Probability theory and Bayesian computing together provide an alternative framework to deal with incomplete and uncertain data.

Decision-Making Tools and Methods for Incomplete and Uncertain Data

Emphasizing probability as an alternative to Boolean logic, Bayesian Programming covers new methods to build probabilistic programs for real-world applications. Written by the team who designed and implemented an efficient probabilistic inference engine to interpret Bayesian programs, the book offers many Python examples that are also available on a supplementary website together with an interpreter that allows readers to experiment with this new approach to programming.

Principles and Modeling

Only requiring a basic foundation in mathematics, the first two parts of the book present a new methodology for building subjective probabilistic models. The authors introduce the principles of Bayesian programming and discuss good practices for probabilistic modeling. Numerous simple examples highlight the application of Bayesian modeling in different fields.

Formalism and Algorithms

The third part synthesizes existing work on Bayesian inference algorithms since an efficient Bayesian inference engine is needed to automate the probabilistic calculus in Bayesian programs. Many bibliographic references are included for readers who would like more details on the formalism of Bayesian programming, the main probabilistic models, general purpose algorithms for Bayesian inference, and learning problems.

FAQs Along with a glossary, the fourth part contains answers to frequently asked questions. The authors compare Bayesian programming and possibility theories, discuss the computational complexity of Bayesian inference, cover the irreducibility of incompleteness, and address the subjectivist versus objectivist epistemology of probability.

The First Steps toward a Bayesian Computer A new modeling methodology, new inference algorithms, new programming languages, and new hardware are all needed to create a complete Bayesian computing framework. Focusing on the methodology and algorithms, this book describes the first steps toward reaching that goal. It encourages readers to explore emerging areas, such as bio-inspired computing, and develop new programming languages and hardware architectures.

../../../../projets/e-motion/IMG/book.jpg

Figure 16. Bayesian Programming book covert

ESPRESSO Project-Team

6. New Results

6.1. A pivot in between synchrony and asynchrony

Participants: Thierry Gautier, Paul Le Guernic, Jean-Pierre Talpin.

Our time modeling framework requires a pivot specification paradigm to materialise a spectrum of models of computation and communication ranging from synchrony to asynchrony, from software to hardware, and accommodate with (abstractions of) software behaviors (software, functional blocks, tasks) and requirements (temporal properties, contracts, regular expressions) through logical, periodic, multi-periodic or affine time. We aim at developing a framework comprising dataflow networks (communications) and synchronous automata (computations) controlled by synthesised wrapper enforcing abstractions of specified constraints from the software viewpoint (timing requirements).

Relations between Kahn networks and classes of synchronous dataflow graphs (SDF) as well as synchronous languages have been studied in the past (e.g. Lustre), yet never in the full generality of relating the domain-theoretic model of Kahn networks to it most general synchronous incarnation (one that at least allows to express several clock domains) [17]. We are currently elaborating such a model to characterise morphisms between untimed asynchronous networks and multi-clocked, synchronous, dataflow networks. In this prospect, we developed the first constructive operational semantics of Signal [21], which opens to further investigations on its full abstraction relation with a denotational characterisation using Kahn networks over a polychronous domain.

6.2. New functionalities of Polychrony

Participants: Loïc Besnard, Thierry Gautier, Paul Le Guernic.

We have developed and integrated in the Signal toolbox some clock computations useful for optimizations: *assignment clocks* and *utility clocks*. These information may be used to reduce the frequency of the computations and the communications for distributed code generation.

Assignment clock. A given signal is supposed to be computed at the instants of its clock, defined by the clock of the expression of its definition. For a signal x, the expression of its definition can always be rewritten as $x := (E_1\,\texttt{when}\,h_1)\,\texttt{default}\,...\,\texttt{default}\,(E_{n-1}\,\texttt{when}\,h_{n-1})\,\texttt{default}\,(x\,\$\,\texttt{when}\,k)$. If we assume that the signal keeps, between two consecutive instants, the last computed value, the assignment of $(x\,\$)$ to x is unnecessary. Then, the assignment clock is then defined by $h_1\hat{}+...\hat{}+h_{n-1}$, smaller than the clock of x defined by $(h_1\hat{}+...\hat{}+h_{n-1})\hat{}+k$.

Utility clock. The utility clock defines the instants at which a signal is necessary. For a signal x, the utility clock, $hu(x)$ is defined by:

- the clock of x if x is an input, an output, a memory, or if it is used to define an undersampling clock $(\texttt{when}\,f(x))$;
- otherwise, it is defined, for $x \to y_1\,\texttt{when}\,h_1$, ..., $x \to y_n\,\texttt{when}\,h_n$, by $\sum_{i=1,n}(hu(y_i)\hat{}*h_i)$.

If we rewrite the Signal program by sampling the signals (except for inputs/outputs) by their utility clock, the new Signal program is equivalent to the previous one, with respect to its behavior with the external world. Note that the utility clock can be used only when this transformation does not introduce cycles in the graph.

6.3. Formal Verification of Synchronous Dataflow Program Transformations Toward Certified Compilers

Participants: Van-Chan Ngo, Jean-Pierre Talpin, Thierry Gautier, Paul Le Guernic, Loïc Besnard.

Translation validation [49], [48] is a technique that attempts to verify that program transformations preserve the program semantics. A compiler generally involves several phases during its compilation process. For instance, the Signal compiler [2], [8], in its first two phases, *calculates the clock information*, makes *Boolean abstraction*, and makes *static scheduling*. The final phase is the executable code generation. One can try to prove globally that the input program and its final transformed program have the same semantics. However, we believe that a better approach consists in separating the concerns and proving for each phase the preservation of different kinds of semantic properties. In the case of the Signal compiler, the preservation of the semantics can be decomposed into the preservation of clock semantics, data dependence, and value-equivalence of variables.

Translation Validation for Clock and SDGs Transformations. This work focuses on proving the preservation of clock semantics in the first two phases of the Signal compiler. In order to do that we encode the clock semantics and data dependence as *clock models* and *synchronous dependence graphs* (SDGs). Then we show that a transformation is correct if and only if there exist *refinements* between clock models, and between SDGs, written as $\Phi(P_2) \sqsubseteq_{clk} \Phi(P_1)$ and $SDG(P_2) \sqsubseteq_{dep} SDG(P_1)$ [15]. We delegate the checking of the preservation to a SMT-solver [38], [54].

Translation Validation of Polychronous Dataflow Specifications: from Signal to C using Synchronous Dataflow Value-Graphs. In this work, we build a validator for the synchronous dataflow compiler of Signal. This validator tries to match the value-graph [53] of each output of the original program and its transformed counterpart. That ensures that every output of the original program and its counterpart in the transformed program have the same value whenever they are present. Our validator does not require any instrumentation and modification of the compiler, nor any rewriting of the source program.

The Signal program and its generated C program have been represented in the same shared synchronous dataflow value-graph (SDVG), in which the nodes for the same structures (variables, constants, operators) have been shared. For instance, the values of input signals and their corresponding variables in the generated C code are represented by the same nodes in the shared graph. Then, the shared graph is transformed following *predefined rules* to show that all output signal values in the Signal program and their counterparts in the generated C code are rooted at the same subgraph.

Consider the following process, where $IR(P)$ is the compiled code of the program P and TV(SDVG(P,IR(P))) is *true* when all output signal values in P and their counterparts in $IR(P)$ are the same:
if (Cp(P) is Error) then output Error; else
if ((Φ(IR(P)) $\sqsubseteq_{clk}$ Φ(P)) and (SDG(IR(P)) $\sqsubseteq_{dep}$ SDG(P)) and (TV(SDVG(P,IR(P))))) then output IR(P); else output Error.
This will provide formal guarantee as strong as that provided by a formally certified compiler w.r.t. the clock semantics and the data dependence in case the validator is certified formally.

Implementation and Experiments. At a high level, our tool *SigCert* [47] developed in OCaml checks the correctness of the compilation of the Polychrony Signal compiler w.r.t clock semantics, data dependence, and value-equivalence as shown in Figure 8 .

6.4. Exploring system architectures in AADL via Polychrony and SynDEx

Participants: Huafeng Yu, Loïc Besnard, Thierry Gautier, Jean-Pierre Talpin, Paul Le Guernic.

../../../../projets/espresso/IMG/verificationprocess.png

Figure 8. An overview of our integration within Polychrony toolset.

Architecture analysis & design language (AADL) has been increasingly adopted in the design of embedded systems, and corresponding scheduling and formal verification have been well studied. However, little work takes code distribution and architecture exploration into account, particularly considering clock constraints, for distributed multi-processor systems. Our approach [20], [16], [17] handles these concerns within the toolchain AADL-Polychrony-SynDEx. First, in order to avoid semantic ambiguities of AADL, the polychronous/multiclock semantics of AADL, based on a polychronous model of computation, is considered. Clock synthesis is then carried out in Polychrony, which bridges the gap between the polychronous semantics and the synchronous semantics of SynDEx [42]. The same timing semantics is always preserved in order to ensure the correctness of the transformations between different formalisms. Code distribution and corresponding scheduling is carried out on the obtained SynDEx model in the last step, which enables the exploration of architectures originally specified in AADL. Our contribution provides a fast yet efficient architecture exploration approach for the design of distributed real-time and embedded systems. The approach has been illustrated using an avionic case study.

6.5. A synchronous annex for the AADL

Participants: Loïc Besnard, Thierry Gautier, Paul Le Guernic, Jean-Pierre Talpin.

We propose a synchronous timing annex for the SAE standard AADL. Our approach consists of building a synchronous model of computation and communication that best fits the semantics and expressive capability of the AADL and its behavioral annex and yet requires little to know (syntactic) extension to it, i.e. to identify a synchronous core of the AADL (which prerequisites a formal definition of synchrony at hand) and define a formal design methodology to use the AADL in a way that supports formal analysis, verification and synthesis.

Our approach first identifies the core AADL concepts from which time events can be described. Then, is considers the behavior annex (BA) as the mean to model synchronous signals and traces through automata. Finally, we consider elements of the constraint annex to reason about abstractions of these signals and traces by clocks and relations among them. To support the formal presentation of these elements, we define a model of automata that comprises a transition system to express explicit transitions and constraints, in the form of a boolean formula on time, to implicitly constraint its behavior. The implementation of such an automaton amounts to composing its explicit transition system with that of the controller synthesised from its specified constraints.

6.6. Ongoing activities and results for integration of Polychrony with the P toolset

Participants: Christophe Junke, Loïc Besnard, Thierry Gautier, Paul Le Guernic, Jean-Pierre Talpin.

Current state of P. The P language is still under definition, notably for the software/hardware architectural description of systems. In late october 2013, technical partners (headed by Adacore) released the first beta version of the toolset. The main activites of the ESPRESSO team can be splitted in analysis and development activities:

- The analysis activites consisted in understanding what tasks shall be performed ultimately by the P toolset w.r.t. code generation and architecture, and how Polychrony could be used in the proposed workflow.
- The development activities consisted in introducing a modified block sequencing algorithm in P and starting the development of the P to Signal converter.

Co-modeling in P. First, P should import functional behavior from Simulink, Stateflow and UML class, activity and state machine diagrams. Those represent a strictly sequential semantics: "the code generated from functional behaviour language will be strictly sequential and void of tasking features" (P specification [1]).

[1] https://forge.open-do.org/plugins/moinmoin/p/

Second, imported architectural description languages are likely to be SysML, MARTE and AADL, which present concurrent semantics. Hence, "the code generated from architectural description languages may include concurrent semantics (thread, shared resources...)" (*ibid*). However, code generation from architectural description languages will consist of invocations to an underlying real-time API. The current target of code generation is the APEX ARINC-653 API, which provides real-time services like inter/intra-partition communication channels as well as task scheduling. Real-time properties of imported architectural elements, like task periods and scheduling policy, are used to configure those services.

Code distribution. Code generation should be able to distribute the functional blocks among architectural elements (processors/threads and buses/queues).

Polychrony offers a way to distribute Signal processes among different locations [31]. In general, such code distribution may lead to the synthesis of new input and output ports: when expressing synchronous communication with asynchronous protocols, some clock information might need to be added to resynchronize data-flows. Moreover, the computation model of Signal allows to order asynchronous read and write operations to avoid communication deadlock. The extended input/output interfaces of blocks could be reimported back to P in order to ensure the correctness of code distribution.

It appears however that the subset of Simulink that is imported in P, and the execution model of P functional models that is enforced by the P compiler, can be viewed as a composition of endochronous multi-rate nodes (all inputs of a node are computed before all of its outputs; this avoids deadlock problems when composing nodes). This model ends up being similar to a Lustre model of computation, where code distribution can be performed without adding communication flows and where read/write operations can be setup in a general way without introducing deadlocks [41].

Despite the above observations, it might be possible to extend the input/output interfaces of existing P models thanks to Polychrony. One approach is to ensure that block dependencies between Simulink blocks are effectively respected after code distribution. Indeed, functional blocks can be partially ordered thanks to user-defined priorities. If other partners see an interest with this approach, it could be possible to establish communication links between ordered blocks, so that the global execution order of blocks in a distributed setting is the same as the one modeled originally in the simulation environment.

Model clustering. Alternatively, it would be interesting from a Signal point of view to loosen the synchronization assumptions made by both Simulink and P so that only algebraic dependencies are taken into account (e.g. interpret all Simulink subsystems as virtual, ignore all non-strictly required dependencies...), while respecting clock constraints (e.g. sample time, controlled and enabled blocks...). In that case, the Signal compiler could perform code distribution for simulation purposes, or simply to provide another compilation scheme for P. Another step could be to apply an automatic code distribution mechanism into so-called *clusters*, and export those clusters back to P as architectural elements. The resulting P model would end-up being having possibly more tasks/threads and smaller functional blocks, which might be interesting. Those design decisions are still under consideration and must be discussed with other partners.

From P to Signal. The development activities in the P project currently consist in adding a P to Signal translator. It is being developed as a backend of the P toolset, which provides a number of facilities to access and perform computations on P models. The current prototype must be completed and refined according to what are the actual needs in the project, but can already be tested with input models.

In order to validate the approach, the existing test models of the P projects are all checked with this exporter (over two hundreds small models, a couple of big ones). The resulting SSME files are then converted to Signal files: this step required to generate a command-line version of the Eclispe Polychrony product, as well a a batch converter from SSME to Signal (this converter is integrated in the Polychrony environment). In addition to convert SSME files to Signal, this converter also validates the model against Ecore constraints. Finally, the resulting Signal files are compiled with the original `C++` Signal compiler to check typing and clock relationships (those tests are not performed at the SSME level). The resulting test toolchain gives useful feedbacks for the iterative development of the translator.

Partial orders in P. The exporter also needs to export block dependencies from functional models. Since Polychrony is also able to infer a total order while taking into account code distribution, it was not satisfactory to export the existing total order computed by the P toolset: it is more sensible to export the subset that is strictly necessary (or desired). In agreement with technical partners, we modified the existing sequencer so that it could be parameterized with block ordering criteria (for example, we might want to take into account dataflow dependencies as well as user-defined priority in Polychrony, but nothing more). The outcome is a single package responsible for computing partial and total orders inside the P toolset. This prevents other tools, like the P to Signal exporter, to compute partial order by themselves.

The implementation of the sequencer is based on a dependency matrix that helps computing the transitive closure of dependencies (to quickly check whether two blocks are dependent on each other) while keeping track of their transitive reduction (in order to export only the minimal set of relationships). Now that the first version of the P toolset is released, the sequencer will hopefully be integrated in the P toolset.

6.7. Real-Time Scheduling of Dataflow Graphs

Participants: Adnan Bouakaz, Jean-Pierre Talpin.

The ever-increasing functional and nonfunctional requirements in real-time safety-critical embedded systems call for new design flows that solve the specification, validation, and synthesis problems. Ensuring key properties, such as functional determinism and temporal predictability, has been the main objective of many embedded system design models. Dataflow models of computation (such as KPN [44], SDF [46], CSDF [34], etc.) are widely used to model stream-based embedded systems due to their inherent functional determinism. Since the introduction of the (C)SDF model, a considerable effort has been made to solve the static-periodic scheduling problem [28]. Ensuring boundedness and liveness is the essence of the proposed algorithms in addition to optimizing some nonfunctional performance metrics (e.g. buffer minimization, throughput maximization, etc.). However, nowadays real-time embedded systems are so complex that real-time operating systems are used to manage hardware resources and host real-time tasks. Most of real-time operating systems rely on priority-driven scheduling algorithms [51], [37] (e.g. RM, EDF, etc.) instead of static schedules which are inflexible and difficult to maintain. Our work [12], [18], [19] [35] addresses the real-time scheduling problem of dataflow graph specifications; i.e., transformation of the dataflow specification to a set of independent real-time tasks w.r.t. a given priority-driven scheduling policy such that the following properties are satisfied: (1) channels are bounded and overflow/underflow-free; (2) the task set is schedulable on a given uniprocessor (or multiprocessor) architecture. This problem requires the synthesis of scheduling parameters (e.g. periods, priorities, processor allocation, etc.) and channel capacities. Furthermore, our work considers two performance optimization problems: buffer minimization and throughput maximization.

6.8. Structure-Preserved Distribution of Synchronous Programs

Participants: Ke Sun, Loïc Besnard, Thierry Gautier, Paul Le Guernic, Jean-Pierre Talpin.

We propose an automatically structure-preserved distribution method, which is based on synchronous guarded actions [50] and component calls in an intermediate representation [36]. The guarded actions describe the local behavior. The component calls preserve the modular structure information of synchronous programs. Using this method, the designer can naturally blend the distribution design into the whole system design procedure: following the modular structure, a globally asynchronous locally synchronous (GALS) network over distributed nodes can be automatically constructed. Each node corresponds to a component and contains:

- a computing element, computing (as sender) or reacting to (as receiver) scheduling commands;
- a controlling element, called adaptor, adjusting the asynchronous communication between nodes.

The computing element focuses on the functional behaviors (i.e., value computation) in synchronous runs, which can be perfectly described by synchronous guarded actions. On the other hand, the controlling element mainly considers the temporal constraints (i.e., clock relation) under asynchronous communications. Guarded actions are not suitable for specifying clock relations, then we use polychronous specifications [8] to define the inter-node communications.

A perspective for future work would be the structure-preserved distribution of synchronous programs with multi-interaction. Multiple interactions in one logical instant are desynchronized and projected onto finer grained instants. Owing to this extension, it would provide more convenience for the designer to express the parallelism among components.

EXMO Project-Team

6. New Results

6.1. Ontology matching and alignments

We pursue our work on ontology matching and alignment support [5], [12] with contributions to evaluation and alignment semantics.

6.1.1. Evaluation

Participant: Jérôme Euzenat.

Since 2004, we run the Ontology Alignment Evaluation Initiative (OAEI) which organises evaluation campaigns for assessing the degree of achievement of actual ontology matching algorithms [2].

This year, we ran the OAEI 2013 evaluation campaign [7]. It offered 8 different test sets (7 of which under the SEALS platform). This issue brought the following results:

- Once again, more participants than ever (23);
- Most ontology matchers running on the SEALS platform (20);
- Increased performances in terms of precision and recall;
- Matchers are now faster and more scalable. There are also more matchers using networked resources.

We used again the our generator for generating new version of benchmarks [4]. The Alignment API was used for manipulating alignments and evaluating results.

A novelty of this year was the evaluation of interactive systems, included in the SEALS client. It brings interesting insight on the performances of such systems and should certainly be continued.

The participating systems and evaluation results were presented in the 8th Ontology Matching workshop, that was held in Sydney, Australia [13]. More information on OAEI can be found at http://oaei.ontologymatching.org/.

6.1.2. Algebras of relations in alignments

Participants: Armen Inants [Correspondent], Jérôme Euzenat.

We had previously shown that algebras of relations between concepts can be used for expressing relations in alignments. We have worked this year as extending them in two ways.

We increased the expressiveness of relations between concepts, not restricting the algebra to necessarily non empty concepts. This describes all taxonomical (as opposed to mereological) relation algebras, i.e., all those relations that have been used by matchers so far.

We also dealt with relations among different kinds of entities – individuals or concepts. For this, relation algebra structures are considered in an arbitrary one- or many-sorted logical theory. We established a sufficient condition for a set of dyadic formulas in a first-order theory to generate a relation algebra. This result is extended to many-sorted theories by means of Schröder categories.

This work is part of the PhD of Armen Inants.

6.2. Data interlinking

The web of data uses semantic web technologies to publish data on the web in such a way that they can be interpreted and connected together. It is thus critical to be able to establish links between these data, both for the web of data and for the semantic web that it contributes to feed. We consider this problem from different perspectives.

6.2.1. Interlinking cross-lingual RDF data sets

Participants: Tatiana Lesnikova [Correspondent], Jérôme David, Jérôme Euzenat.

Data interlinking is a difficult task in a cross-lingual environment like the Web. Even systems based on graph structure, ultimately rely on anchors based on language fragments. If languages are different, fragments have to be compared by more sophisticated techniques. In that context, we are developing an approach which represents RDF entities as (virtual) text documents and compare them using different strategies [9], [10]. We investigate two directions: (1) a translation-based approach where the virtual documents are automatically translated; (2) a language-independent approach where important terms found in documents are mapped to a terminological resource like Wordnet to compute document similarity.

This work is part of the PhD of Tatiana Lesnikova developed in the LINDICLE project (see §7.1.2).

6.2.2. Data interlinking from expressive alignments

Participants: Zhengjie Fan [Correspondent], Jérôme Euzenat.

In the context of the DATALIFT project, we are further developing the data interlinking module. We have developed an algorithm able to determine potential attribute correspondences of two classes depending on their features. For that purpose, we use k-means or k-medoids clustering. These correspondences are then used to construct a SILK script which generates an initial link set. Some of the links are presented to the user who assesses their validity. We then use an improvement of the disjunctive version space supervised learning method to learn a better script from the assessed links. Such a technique can be iterated until satisfactory links are found.

This work is part of the PhD of Zhengjie Fan, co-supervised with François Scharffe (LIRMM), and developed in the DATALIFT project (see §7.1.1).

6.2.3. Key and pseudo-key detection for web data set interlinking

Participants: Jérôme David [Correspondent], Manuel Atencia Arcas, Anthony Delaby, Jérôme Euzenat.

Keys are sets of properties which uniquely identify individuals (instances of a class). We have refined the notion of database keys in a way which is more adapted to the context of description logics and the openness of the semantic web. We have also refined the weaker notion of a linkkey introduced in [12]. Then we have shown how such keys, together with ontology alignments, and linkkeys may be used for deducing equality statements (links) between individuals across data sources in the web of data.

However, ontologies do not necessarily come with key descriptions, and never with linkkey assertions (which would hold across ontologies). But, these can be extracted from data by assuming that keys holding for specific data sets, may hold universally. We have extended these classical key extraction techniques for extracting linkkeys.

This work is developed partly in the LINDICLE and DATALIFT projects. A proof of concept implementation is available at http://rdfpkeys.inrialpes.fr/.

6.3. Ontology networks

Dealing with the semantic web, we are interested in ontology networks, i.e., sets of distributed ontologies that have to work together. One way for these systems to interact consists of exchanging queries and answers. For that reason, we pay particular attention to query systems.

6.3.1. Path queries and μ-calculus

Participant: Jérôme Euzenat.

Querying the semantic web is mainly done through the SPARQL language or its extensions through paths and entailment regimes [14]. Query containment is the problem of deciding if the answers to a query are included in those of another query for any queried data sources. This problem is very important for query optimisation purposes. In the SPARQL context, it can be equally useful for distributing federated queries or for implementing schema-based access control. In order to experimentally assess implementation strengths and limitations, we provided a first SPARQL containment test benchmark. We studied the query demographics on DBPEDIA logs to design benchmarks for relevant query containment solvers. We tested available solvers on their domain of applicability on three different benchmark suites [6] and found that (i) tested solutions are overall functionally correct, (ii) in spite of its complexity, SPARQL query containment is practicable for acyclic queries, (iii) state-of-the-art solvers are at an early stage both in terms of capabilities and implementation.

This work has been developed in collaboration with the TYREX team and within the PhD thesis of Melisachew Wudage Chekol now in the ORPAILLEUR team. The benchmarks, results and software are available at http://sparql-qc-bench.inrialpes.fr.

FLOWERS Project-Team

6. New Results

6.1. Robotic and Computational Models of Human Development

6.1.1. Computational models of information-seeking, curiosity and attention

Participants: Pierre-Yves Oudeyer, Manuel Lopes.

An associated team, called Neurocuriosity, was created between Flowers and the Cognitive Neuroscience lab of Jacqueline Gottlieb at Univ. Columbia, NY. The goal of this associated team is to investigate mechanisms of spontaneous exploration and learning in humans by setting up experiments allowing to confirm or falsify predictions made by computational models previously developped by the team. This constitutes a crucial collaboration between developmental robotics and cognitive neuroscience. This joint work already led to a major publication on curiosity and information seeking, in the prestigious Trends in Cognitive Science journal (impact factor: 16.5). [27]

Abstract: Intelligent animals devote much time and energy to exploring and obtaining information, but the underlying mechanisms are poorly understood. We review recent developments on this topic that have emerged from the traditionally separate fields of machine learning, eye movements in natural behavior, and studies of curiosity in psychology and neuroscience. These studies show that exploration may be guided by a family of mechanisms that range from automatic biases toward novelty or surprise to systematic searches for learning progress and information gain in curiosity-driven behavior. In addition, eye movements reflect visual information searching in multiple conditions and are amenable for cellular-level investigations. This suggests that the oculomotor system is an excellent model system for understanding information-sampling mechanisms.

6.1.1.1. Formalizing Imitation Learning

Participants: Thomas Cederborg, Pierre-Yves Oudeyer.

An original formalization of imitation learning was elaborated. Previous attempts to systematize imitation learning has been limited to categorizing different types of demonstrator goals (for example defining success in terms of the sequential joint positions of a dance, or in terms of environmental end states), and/or been limited to a smaller subset of imitation (such as learning from tele-operated demonstrations). The formalism proposed attempts to describe a large number of different types of learning algorithms using the same notation. Any type of algorithm that modifies a policy based on observations of a human, is treated as an interpretation hypothesis of this behavior. One example would be an update algorithm that updates a policy, partially based on the hypothesis that the demonstrator succeeds at demonstrations with probability 0.8, or an update algorithm that assumes that a scalar value is an accurate evaluation of an action compared to the latest seven actions. The formalism aims to give a principled way of updating these hypotheses, either rejecting some of a set of hypotheses regarding the same type of behavior, or set of parameters of an hypothesis. Any learning algorithm that modifies policy based on observations of a human that wants an agent to do something or act in some way, is describable as an interpretation hypothesis. If the learning algorithm is static, this simply corresponds to an hypothesis that is not updated based on observations. A journal article [26].

6.1.1.2. Self-Organization of Early Vocal Development in Infants and Machines: The Role of Intrinsic Motivation

Participants: Clément Moulin-Frier, Sao Mai Nguyen, Pierre-Yves Oudeyer.

We bridge the gap between two issues in infant development: vocal development and intrinsic motivation. We propose and experimentally test the hypothesis that general mechanisms of intrinsically motivated spontaneous exploration, also called curiosity-driven learning, can self-organize developmental stages during early vocal learning and explain several aspects observed in infants (Figure 20). We introduce a computational model of intrinsically motivated vocal exploration, which allows the learner to autonomously structure its own vocal experiments, and thus its own learning schedule, through a drive to maximize competence progress. This model relies on a physical model of the vocal tract, the auditory system and the agent's motor control, as well as vocalizations of social peers. We present computational experiments that show how such a mechanism can explain the adaptive transition from vocal self-exploration with little influence from the speech environment, to a later stage where vocal exploration becomes influenced by vocalizations of peers (Figure 21). Within the initial self-exploration phase, we show that a sequence of vocal production stages self-organizes, and shares properties with data from infant developmental psychology: the vocal learner first discovers how to control phonation, then focuses on vocal variations of unarticulated sounds, and finally automatically discovers and focuses on babbling with articulated proto-syllables (Figure 22). As the vocal learner becomes more proficient at producing complex sounds, imitating vocalizations of peers starts to provide high learning progress explaining an automatic shift from self-exploration to vocal imitation.

This work has been recently accepted in the journal *Frontiers in Psychology, Cognitive Science* [30].

6.1.1.3. Emergent Proximo-Distal Motor Development through Adaptive Exploration, applied to Reaching and Vocal Learning

Participants: Freek Stulp, Pierre-Yves Oudeyer, Jules Brochard, Clément Moulin-Frier.

Life-long robot learning in the high-dimensional real world requires guided and structured exploration mechanisms. In this developmental context, we have investigated the use of the PI^2-CMAES episodic reinforcement learning algorithm, which is able to learn high-dimensional motor tasks through adaptive control of exploration. By studying PI^2-CMAES in a reaching task on a simulated arm, we observe two developmental properties. First, we show how PI^2-CMAES autonomously and continuously tunes the global exploration/exploitation trade-off, allowing it to re-adapt to changing tasks. Second, we show how PI^2-CMAES spontaneously self-organizes a maturational structure whilst exploring the degrees-of-freedom (DOFs) of the motor space. In particular, it automatically demonstrates the so-called *proximo-distal maturation* observed in humans: after first freezing distal DOFs while exploring predominantly the most proximal DOF, it progressively frees exploration in DOFs along the proximo-distal body axis. These emergent properties suggest the use of PI^2-CMAES as a general tool for studying reinforcement learning of skills in life-long developmental learning contexts. This work was published in the Paladyn Journal of Behavioral Robotics [36].

This model of emergent developmental freezing and unfreezing of degrees of freedom was then applied to infant vocal development. For this aim, we used an articulatory synthesizer which is a computer model of the human vocal tract and the ear. While testing different possibilities, the algorithm eventually creates learning structures, which are more efficient that random motor babbling. Using the algorithm with a vocal synthesizer, we show that it can reproduce a babbling infant's characteristic: the predominance of the jaw over the other articulators, namely the canonical babbling.

This is the first study to our knowledge of emergent maturation in speech. Without presupposing any biological or social constraint, we give a new explanation of the jaw predominance in babbling, based on freezing and freeing the degrees of freedom in an adaptive maturation scheme to improve learning. This provides an original hypothesis regarding the emergence of canonical babbling in infant vocal development.

This last work was performed during the internship of Jules Brochard in 2013 and a journal article is currently being written.

6.1.2. COSMO ("Communicating about Objects using Sensory-Motor Operations"): a Bayesian modeling framework for studying speech communication and the emergence of phonological systems

Participants: Clément Moulin-Frier, Jean-Luc Schwartz, Julien Diard, Pierre Bessiã¨re.

../../../../projets/flowers/IMG/VocalChronol.png

Figure 20. Rapid view of the first year of infant vocal development.

../../../../projets/flowers/IMG/percentageStrategy.png

Figure 21. Our model displays an adaptive transition from vocal self-exploration with little influence from the speech environment, to a later stage where vocal exploration becomes influenced by vocalizations of peers.

../../../../projets/flowers/IMG/dev_seq.png

Figure 22. Within the self-exploration phase, our model first discovers how to control phonation, then focuses on vocal variations of unarticulated sounds, and finally automatically discovers and focuses on babbling with articulated proto-syllables.

This work began with the PhD thesis of Clement Moulin-Frier at GIPSA-Lab, Grenoble, France, supervised by Jean-Luc Schwartz (GIPSA-Lab, CNRS), Julien Diard (LPNC, CNRS) and Pierre Bessière (College de France, CNRS). A few papers were finalized during his post-doc at FLOWERS in 2012. Firstly, an international journal paper based on the PhD thesis work of Raphael Laurent (GIPSA-Lab), extending Moulin-Frier's model, was published [108], as well as a commentary in *Behavioral and Brain Sciences* [97]. Both these papers provide computational arguments based on a sensory-motor cognitive model to feed the age-old debate of motor vs. auditory theories of speech perception. Secondly, in another journal paper under the submission process, we attempt to derive some properties of phonological systems (the sound systems of human languages) from the mere properties of speech communication. We introduce a model of the cognitive architecture of a communicating agent, called COSMO (for "Communicating about Objects using Sensory-Motor Operations") that allows expressing in a probabilistic way the main theoretical trends found in the speech production and perception literature. This allows a computational comparison of these theoretical trends, helping to identify the conditions that favor the emergence of linguistic codes. We present realistic simulations of phonological system emergence showing that COSMO is able to predict the main regularities in vowel, stop consonant and syllable systems in human languages.

This work is currently under consideration as a target article for a special issue in an international journal. Pierre-Yves Oudeyer joined this process as a member of the editing committee.

6.1.3. Recognizing speech in a novel accent: the Motor Theory of Speech Perception reframed

Participants: Clément Moulin-Frier, Michael Arbib.

Clément Moulin-Frier engaged this work with Michael Arbib during his 6-month visit in 2009 at the USC Brain Project, University of Southern California, Los Angeles, USA, during his PhD thesis at Gipsa-Lab, Grenoble. He continues to write a journal article during his post-doc in the Flowers team in 2012-2013. This paper has been published recently in Biological Cybernetics [29], in which we offer a novel computational model of foreign-accented speech adaptation, together with a thorough analysis of its implications with respect to the motor theory of speech perception.

6.2. Life-Long Robot Learning and Development of Motor and Social Skills

6.2.1. Active Learning and Intrinsic Motivation

6.2.1.1. Active Learning of Inverse Models with Goal Babbling

Participants: Adrien Baranes, Pierre-Yves Oudeyer.

We have continued to elaborate and study our Self-Adaptive Goal Generation - Robust Intelligent Adaptive Curiosity (SAGG-RIAC) architecture as an intrinsically motivated goal exploration mechanism which allows active learning of inverse models in high-dimensional redundant robots. Based on active goal babbling, this allows a robot to efficiently and actively learn distributions of parameterized motor skills/policies that solve a corresponding distribution of parameterized tasks/goals. The architecture makes the robot sample actively novel parameterized tasks in the task space, based on a measure of competence progress, each of which triggers low-level goal-directed learning of the motor policy parameters that allow to solve it. For both learning and generalization, the system leverages regression techniques which allow to infer the motor policy parameters corresponding to a given novel parameterized task, and based on the previously learnt correspondences between policy and task parameters.

We have conducted experiments with high-dimensional continuous sensorimotor spaces in three different robotic setups: 1) learning the inverse kinematics in a highly-redundant robotic arm, 2) learning omnidirectional locomotion with motor primitives in a quadruped robot 2324 , 3) an arm learning to control a fishing rod with a flexible wire. We show that 1) exploration in the task space can be a lot faster than exploration in the actuator space for learning inverse models in redundant robots; 2) selecting goals maximizing competence progress creates developmental trajectories driving the robot to progressively focus on tasks of increasing complexity and is statistically significantly more efficient than selecting tasks randomly, as well as more efficient than different standard active motor babbling methods; 3) this architecture allows the robot to actively discover which parts of its task space it can learn to reach and which part it cannot.

This work was published in the journal Robotics and Autonomous Systems [25].

6.2.1.2. Learning Exploration Strategies in Model-based Reinforcement Learning

Participants: Manuel Lopes, Todd Hester, Peter Stone, Pierre-Yves Oudeyer.

We studied how different exploration algorithms can be combine and selected at runtime. Typically the user must hand-tune exploration parameters for each different domain and/or algorithm that they are using. We introduced an algorithm called leo for learning to select among different exploration strategies on-line. This algorithm makes use of bandit-type algorithms to adaptively select exploration strategies based on the rewards received when following them. We show empirically that this method performs well across a set of five domains In contrast, for a given algorithm, no set of parameters is best across all domains. Our results demonstrate that the leo algorithm successfully learns the best exploration strategies on-line, increasing the received reward over static parameterizations of exploration and reducing the need for hand-tuning exploration parameters [46].

6.2.1.3. Active Inverse Reinforcement Learning through Generalized Binary Search

Participants: Manuel Lopes, Francisco Melo.

We contributed the first aggressive active learning algorithm for nonseparable multi-class classification. We generalize an existing active learning algorithm for binary classification [116] to the multi-class setting, and identify mild conditions under which the proposed method provably retains the main properties of the original algorithm, namely consistency and sample complexity. In particular, we show that, in the binary case, our method reduces to the original algorithm of [116]. We then contribute an extension of our method to multi-label settings, identify its main properties and discuss richer querying strategies. We conclude the paper with two illustrative application examples. The first application features a standard text-classification problem. The second application scenario features a learning from demonstration setting. In both cases we demonstrate the advantage of our active sampling approach against random sampling. We also discuss the performance of the proposed approach in terms of the derived theoretical bounds.

../../../../projets/flowers/IMG/AIBOVECTEN.png

Figure 23. Experimenting SAGG-RIAC for learning an inverse model for omnidirectional locomotion of a quadruped robot. The quadruped robot is controlled using 24 dimensional motor synergies parameterized with 24 continuous values : 12 for the amplitudes and 12 others for the phases of a sinusoid tracked by each motor. Experiments consider a task space u, v, α which corresponds to the 2D position and orientation of the quadruped.

../../../../projets/flowers/IMG/FigureEvolutionQuad.png

Figure 24. Evolution of the quality of the learnt inverse model for the quadruped robot experiment, depending on various exploration strategies (measured as mean error over a set of uniformly distributed goals generated independantly from learning trials).

6.2.1.4. Exploration strategies in developmental robotics: a unified probabilistic framework

Participants: Clément Moulin-Frier, Pierre-Yves Oudeyer.

We present a probabilistic framework unifying two important families of exploration mechanisms recently shown to be efficient to learn complex non-linear redundant sensorimotor mappings. These two explorations mechanisms are: 1) goal babbling, 2) active learning driven by the maximization of empirically measured learning progress. We show how this generic framework allows to model several recent algorithmic architectures for autonomous exploration. Then, we propose a particular implementation using Gaussian Mixture Models, which at the same time provides an original empirical measure of the competence progress. Finally, we perform computer simulations on two simulated setups: the control of the end effector of a 7-DoF arm and the control of the formants produced by an articulatory synthesizer. We are able to reproduce previous results from [25] with the advantages of a clean and compact probabilistic framework to efficiently express, implement and compare various exploration strategies on developmental robotics setups.

This work was published in three international conferences [54], [56], [55].

6.2.1.5. Autonomous Reuse of Motor Exploration Trajectories

Participants: Fabien Benureau, Pierre-Yves Oudeyer.

We developped an algorithm for transferring exploration strategies between tasks that share a common motor space in the context of lifelong autonomous learning in robotics. In such context sampling is costly, and exploration can take a long time before finding interesting, learnable data about a task. Our algorithm shows that we can significantly reduce sampling by reusing past data of other learned tasks, with no need of external knowledge or specific task structure. The algorithm does not transfer observations, or make assumptions about how the learning is conducted. Instead, only selected motor commands are transferred between tasks, chosen autonomously according to an empirical measure of learning progress. We show that on a wide variety of variations from a source task, such as changing the object the robot is interacting with or altering the morphology of the robot, this simple and flexible transfer method increases early performance significantly in the new task. We also investigate the limitation of this algorithm on specific situations.

This work has been published at ICDL, in Osaka [40].

6.2.2. Learning and optimization of motor policies

6.2.2.1. Off-Policy Actor-Critic

Participants: Thomas Degris, Martha White, Richard Sutton.

Actor–critic architectures are an interesting candidate for learning with robots: they can represent complex stochastic policies suitable for robots, they can learn online and incrementally and their per-time-step complexity scales linearly with the number of learned weights. Moreover, interesting connections have been identified in the existing literature with neuroscience. Until recently, however, practical actor–critic methods have been restricted to the on-policy setting, in which the agent learns only about the policy it is executing.

In an off-policy setting, on the other hand, an agent learns about a policy or policies different from the one it is executing. Off-policy methods have a wider range of applications and learning possibilities. Unlike on-policy methods, off-policy methods are able to, for example, learn about an optimal policy while executing an exploratory policy, learn from demonstration, and learn multiple tasks in parallel from a single sensory-motor interaction with an environment. Because of this generality, off-policy methods are of great interest in many application domains.

We have presented the first actor-critic algorithm for off-policy reinforcement learning. Our algorithm is online and incremental, and its per-time-step complexity scales linearly with the number of learned weights. We have derived an incremental, linear time and space complexity algorithm that includes eligibility traces and empirically show better or comparable performance to existing algorithms on standard reinforcement-learning benchmark problems. This work was was reproduced independently by Saminda Abeyruwan from the University of Miami.

6.2.2.2. Auto-Actor Critic

Participant: Thomas Degris.

As mentioned above, actor–critic architectures are an interesting candidate for robots to learn new skills in unknown and changing environments. However, existing actor–critic architectures, as many machine learning algorithms, require manual tuning of different parameters to work in the real world. To be able to systematize and scale-up skill learning on a robot, learning algorithms need to be robust to their parameters. The Flowers team has been working on making existing actor–critic algorithms more robust to make them suitable to a robotic setting. Results on standard reinforcement learning benchmarks are encouraging. This work will be submitted to international conference related with reinforcement learning. Interestingly, the methods developed in this work also offer a new formalism to think about different existing themes of Flowers research such as curiosity and maturational constraints.

6.2.2.3. Deterministic Policy Gradient Algorithms

Thomas Degris and colleagues from UCL and Deepming have consider deterministic policy gradient algorithms for reinforcement learning with continuous actions. The deterministic pol- icy gradient has a particularly appealing form: it is the expected gradient of the action-value func- tion. This simple form means that the deter- ministic policy gradient can be estimated much more efficiently than the usual stochastic pol- icy gradient. To ensure adequate exploration, we introduce an off-policy actor-critic algorithm that learns a deterministic target policy from an exploratory behaviour policy. We demonstrate that deterministic policy gradient algorithms can significantly outperform their stochastic counter- parts in high-dimensional action spaces. [58]

6.2.2.4. Relationship between Black-Box Optimization and Reinforcement Learning

Participant: Freek Stulp.

Policy improvement methods seek to optimize the parameters of a policy with respect to a utility function. There are two main approaches to performing this optimization: reinforcement learning (RL) and black-box optimization (BBO). In recent years, benchmark comparisons between RL and BBO have been made, and there has been several attempts to specify which approach works best for which types of problem classes.

We have made several contributions to this line of research by: 1) Defining four algorithmic properties that further clarify the relationship between RL and BBO. 2) Showing how the derivation of ever more powerful RL algorithms displays a trend towards BBO. 3) Continuing this trend by applying two modifications to the state-of-the-art PI^2 algorithm, which yields an algorithm we denote PI^{BB}. We show that PI^{BB} is a BBO algorithm, and, more specifically, that it is a special case of the state-of-the-art CMAES algorithm. 4) Demonstrating that the simpler PI^{BB} achieves similar or better performance than PI^2 on several evaluation tasks. 5) Analyzing why BBO outperforms RL on these tasks. These contributions have been published on HAL [129], and in Paladyn: Journal of Behavioral Robotics [36].

6.2.2.5. Probabilistic optimal control: a quasimetric approach

Participants: Steve N'Guyen, Clément Moulin-Frier, Jacques Droulez.

During his previous post-doc at the Laboratoire de Physiologie de la Perception et de l'Action (College de France, Paris), Clément Moulin-Frier joined Jacques Droulez and Steve N'Guyen to work on an alternative and original approach of probabilistic optimal control called *quasimetric*. A journal paper was published in PLoS ONE in December 2013 [31], where the authors propose a new approach dealing with decision making under uncertainty.

6.2.3. Social learning and intrinsic motivation

6.2.3.1. Socially Guided Intrinsic Motivation for Skill Learning

Participants: Sao Mai Nguyen, Pierre-Yves Oudeyer.

We have explored how social interaction can bootstrap the learning of a robot for motor learning. We first studied how simple demonstrations by teachers could have a bootstrapping effect on autonomous exploration with intrinsic motivation by building a learner who uses both imitation learning and SAGG-RIAC algorithm [25], and thus designed the SGIM-D (Socially Guided Intrinsic Motivation by Demonstration) algorithm [32], [24], [32] [114], [111]. We then investigated on the reasons of this bootstrapping effect [113], to show that demonstrations by teachers can both enhance more tasks to be explored, as well as favor more easily generalized actions to be used. This analysis is generalizable for all algorithms using social guidance and goal-oriented exploration. We then proposed to build a strategic learner who can learn multiple tasks and with multiple strategies. An overview and theoretical study of multi-task, multi-strategy Strategic Learning is presented in [99]. We also forsook to build a learning algorithm for more natural interaction with the human users. We first designed the SGIM-IM algorithm so that it can determine itself when it should ask for help from the teacher while trying to explore autonomously as long as possible so as to use as little of the teacher's time as possible [112]. After tackling with the problem of how and when to learn, we also investigated an active learner who can determine who to ask for help: in the case of two teachers available, SGIM-IM can determine which strategy to adopt between autonomous exploration and learning by demonstration, and which teacher enhances most learning progress for the learner [115], and ask him for help.

../../../../projets/flowers/IMG/AirHockeyTable.png

Figure 25. Illustration of SGIM-D and SGIM-IM algorithms

While the above results have been shown in simulation environments: of a simple deterministic air hockey game (fig. 25), and a stochastic fishing experiment with a real-time physical simulator (fig. 26), we are now building the experimental setup of the fishing experiment in order to carry out the experiments with naive users.

6.2.3.2. Adaptive task execution for implicit human-robot coordination

../../../../projets/flowers/IMG/FishingRod.png

Figure 26. Illustration of SGIM-D and SGIM-IM algorithms

../../../../projets/flowers/IMG/iCub.jpg

Figure 27. Illustration of SGIM-D and SGIM-IM algorithms

Participants: Ievgen Perederieiev, Manuel Lopes, Freek Stulp.

We began a project which goal is to study how computational models of multi-agent systems can be applied in situations where one agent is a human. We aim at applications where robots collaborate with humans for achieving complex tasks..

A very important capability for efficient collaborative work is the mutual agreement of a task and the ability to predict the behavior of others. We address such aspect by studying methods that increase the predictability of the robot actions. An efficient motor execution becomes the one that not just optimize speed and minimizes energy but also the one that improves the reliability of the team behavior. We are studying policy gradient methods and working on policy improvement algorithms (PI^2, CEM and $CMAES$). A feasibility study will consider a simple task between a robot and a person where the goal is to coordinate the way a set of three colored buttons is pressed.

6.2.4. Unsupervised learning of motor primitives

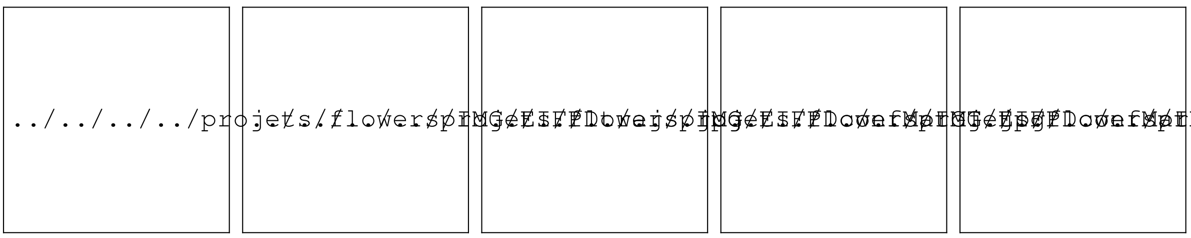
6.2.4.1. Clustering activities

Participants: Manuel Lopes, Luis Montesano, Javier Almingol.

Learning behaviors from data has applications in surveillance and monitoring systems, virtual agents and robotics among others. In our approach, ww assume that in a given unlabeled dataset of multiple behaviors, it is possible to find a latent representation in a controller space that allows to generate the different behaviors. Therefore, a natural way to group these behaviors is to search a common control system that generate them accurately.

Clustering behaviors in a latent controller space has two major challenges. First, it is necessary to select the control space that generate behaviors. This space will be parameterized by a set of features that will change for different behaviors. Usually, each controller will minimize a cost function with respect to several task features. The latent representation is in turn defined by the selected features and their corresponding weight. Second, an unknown number of such controllers is required to generate different behaviors and the grouping must be based on the ability of the controller to generate the demonstrations using a compact set of controllers.

We propose a Dirichlet Process based algorithm to cluster behaviors in a latent controller space which encodes the dynamical system generating the observed trajectories. The controller uses a potential function generated as a linear combination of features. To enforce sparsity and automatically select features for each cluster independently, we impose a conditional Laplace prior over the controller parameters. Based on this models, we derive a sparse Dirichlet Process Mixture Model (DPMM) algorithm that estimates the number of behaviors and a sparse latent controller for each of them based on a large set of features [38].



(a) (b) (c) (d) (e)

Figure 28. EIFPD dataset. (a) Trajectories of the EIFPD to be clustered (color is non-informative). (b-d) correspondence matrix for the 474 trajectories for the labeled ground truth, the KMeans in measurement space and the DPMM, respectively. (e) Reconstructed trajectories from the initial point using the estimated parameters of the DPMM algorithm. Due to the large number of clusters (37), colors are repeated for different clusters.

6.2.4.2. Learning the Combinatorial Structure of Demonstrated Behaviors with Inverse Reinforcement Control

Participants: Olivier Mangin, Pierre-Yves Oudeyer.

We have elaborated and illustrated a novel approach to learning motor skills from demonstration. This approach combines ideas from inverse reinforcement learning, in which actions are assumed to solve a task, and dictionary learning. In this work we introduced a new algorithm that is able to learn behaviors by assuming that the observed complex motions can be represented in a smaller dictionary of concurrent tasks. We developed an optimization formalism and show how we can learn simultaneously the dictionary and the mixture coefficients that represent each demonstration. We presented results on a toy problem and shown that our algorithm finds an efficient set of primitive tasks where naive approaches such as PCA and using a dictionary built from random examples fail to achieve the same tasks. These results that where presented as [60], extend the ones from [103].

6.2.4.3. Interaction of Maturation and Intrinsic Motivation for Developmental Learning of Motor Skills in Robots

Participants: Adrien Baranes, Pierre-Yves Oudeyer.

We have introduced an algorithmic architecture that couples adaptively models of intrinsic motivation and physiological maturation for autonomous robot learning of new motor skills. Intrinsic motivation, also called curiosity-driven learning, is a mechanism for driving exploration in active learning. Maturation denotes here mechanisms that control the evolution of certain properties of the body during development, such as the number and the spatio-temporal resolution of available sensorimotor channels. We argue that it is useful to introduce and conceptualize complex bidirectional interactions among these two mechanisms, allowing to actively control the growth of complexity in motor development in order to guide efficiently exploration and learning. We introduced a model of maturational processes, taking some functional inspiration from the myelination process in humans, and show how it can be coupled in an original and adaptive manner with the intrinsic motivation architecture SAGG-RIAC (Self-Adaptive Goal Generation - Robust Intelligent Adaptive Curiosity algorithm), creating a new system, called McSAGG-RIAC. We then conducted experiments to evaluate both qualitative and quantitative properties of these systems when applied to learning to control a high-dimensional robotic arm, as well as to learning omnidirectional locomotion in a quadruped robot equipped with motor synergies. We showed that the combination of active and maturational learning can allow to gain orders of magnitude in learning speed as well as reach better generalization performances. A journal article is currently being written.

6.3. Autonomous and Social Perceptual Learning

6.3.1. The Impact of Human-Robot Interfaces on the Learning of Visual Objects

Participants: Pierre Rouanet, Pierre-Yves Oudeyer, Fabien Danieau, David Filliat.

We have continued and finalized a large-scale study of the impact of interfaces allowing non-expert users to efficiently and intuitively teach a robot to recognize new visual objects. We identified challenges that need to be addressed for real-world deployment of robots capable of learning new visual objects in interaction with everyday users. We argue that in addition to robust machine learning and computer vision methods, well-designed interfaces are crucial for learning efficiency. In particular, we argue that interfaces can be key in helping non-expert users to collect good learning examples and thus improve the performance of the overall learning system. Then, we have designed four alternative human-robot interfaces: three are based on the use of a mediating artifact (smartphone, wiimote, wiimote and laser), and one is based on natural human gestures (with a Wizard-of-Oz recognition system). These interfaces mainly vary in the kind of feedback provided to the user, allowing him to understand more or less easily what the robot is perceiving, and thus guide his way of providing training examples differently. We then evaluated the impact of these interfaces, in terms of learning efficiency, usability and user's experience, through a real world and large scale user study. In this experiment, we asked participants to teach a robot twelve different new visual objects in the context of a robotic game. This game happens in a home-like environment and was designed to motivate and engage users in an interaction where using the system was meaningful. We then analyzed results that show significant differences among

interfaces. In particular, we showed that interfaces such as the smartphone interface allows non-expert users to intuitively provide much better training examples to the robot, almost as good as expert users who are trained for this task and aware of the different visual perception and machine learning issues. We also showed that artifact-mediated teaching is significantly more efficient for robot learning, and equally good in terms of usability and user's experience, than teaching thanks to a gesture-based human-like interaction.
This work was published in the IEEE Transactions on Robotics [34].

../../../../projets/flowers/IMG/bd-iphone-1.png

Figure 29. Smartphone Interface. To make the robot collect a new learning example, users have to first draw the robot's attention toward the object they want to teach through simple gestures. Once the robot sees the object, they touch the head of the robot to trigger the capture. Then, they directly encircle the area of the image that represents the object on the screen. The selected area is then used as the new learning example. The combination of the video stream and the gestures facilitate the achievement of joint attention.

6.3.2. Developmental object learning through manipulation and human demonstration

Participants: Natalia Lyubova, David Filliat.

../../../../projets/flowers/IMG/bd-wiilaser-1.png

Figure 30. Wiimote + laser pointer interface. With this interface users can draw the robot's attention with a laser pointer toward an object. The laser spot is automatically tracked by the robot. They can ensure that the robot detects the spot thanks to haptic feedback on the Wiimote. Then, they can touch the head of the robot to trigger the capture of a new learning example. Finally, they encircle the object with the laser pointer to delimit its area which will be defined as the new learning example.

../../../../projets/flowers/IMG/cluttered.jpg

Figure 31. The real world environment designed to reproduce a typical living room. Many objects were added in the scene in order to make the environment cluttered.

../../../../projets/flowers/IMG/iCub.jpg

Figure 32. iCub performing curiosity-driven exploration and active recognition of visual objects in 3D

The goal of this work is to design a visual system for a humanoid robot. We used a developmental approach that allows a humanoid robot to continuously and incrementally learn entities through interaction with a human partner in a first stage before categorizing these entities into objects, humans or robot parts and using this knowledge to improve objects models by manipulation in a second stage. This approach does not require prior knowledge about the appearance of the robot, the human or the objects. The proposed perceptual system segments the visual space into proto-objects, analyses their appearance, and associates them with physical entities. Entities are then classified based on the mutual information with proprioception and on motion statistics. The ability to discriminate between the robot's parts and a manipulated object then allows to update the object model with newly observed object views during manipulation. We evaluate our system on an iCub robot, showing the independence of the self-identification method on the robot's hands appearances by wearing different colored gloves. The interactive object learning using self-identification shows an improvement in the objects recognition accuracy with respect to learning through observation only [52], [51].

6.3.3. A Comparison of Geometric and Energy-Based Point Cloud Semantic Segmentation Methods

Participants: Mathieu Dubois, Alexander Gepperth, David Filliat.

The software we developped for object segmentation and recognition rely on a geometric segmentation of the space. We tested alternative methods for this semantic segmentation task in which the goal is to find some relevant classes for navigation such as wall, ground, objects, etc. Several effective solutions have been proposed, mainly based on the recursive decomposition of the point cloud into planes. We compare such a solution to a non-associative MRF method inspired by some recent work in computer vision.

The results [42] shows that the geometric method gives superior results for the task of semantic segmentation in particular for the object class. This can be explained by the fact that it incorporates a lot of domain knowledge (namely that indoor environments are made of planes and that objects lie on top of them). However, MRF segmentation gives interesting results and has several advantages. First most of it's components can be used for other purpose or in other, less constrained, environments where domain knowledge is not available. For instance we could try to recognize more precisely the objects. Second it requires less tuning since most parameters are learned from the database. Third, it uses the appearance information which could help to identify different types of ground or wall (this was one of the goal in the CAROTTE challenge). Last but not least, as it gives a probabilistic output, it allows the robot to draw hypothesis on the environment and adapt its behavior. Therefore we think it is interesting to investigate improvements to improve the exploitation of the structure of the point clouds.

6.3.4. Efficient online bootstrapping of sensory representations

Participant: Alexander Gepperth.

This work [86] is a simulation-based investigation exploring a novel approach to the open-ended formation of multimodal representations in autonomous agents. In particular, we addressed here the issue of transferring (bootstrapping) features selectivities between two modalities, from a previously learned or innate reference representation to a new induced representation. We demonstrated the potential of this algorithm by several experiments with synthetic inputs modeled after a robotics scenario where multimodal object representations are bootstrapped from a (reference) representation of object affordances, focusing particularly on typical challenges in autonomous agents: absence of human supervision, changing environment statistics and limited computing power. We proposed an autonomous and local neural learning algorithm termed PROPRE (projection-prediction) that updates induced representations based on predictability: competitive advantages are given to those feature-sensitive elements that are inferable from activities in the reference representation, the key ingredient being an efficient online measure of predictability controlling learning. We verified that the proposed method is computationally efficient and stable, and that the multimodal transfer of feature selectivity is successful and robust under resource constraints. Furthermore, we successfully demonstrated robustness to noisy reference representations, non-stationary input statistics and uninformative inputs.

6.3.5. Simultaneous concept formation driven by predictability

Participants: Alexander Gepperth, Louis-Charles Caron.

This work [83] was conducted in the context of developmental learning in embodied agents who have multiple data sources (sensors) at their disposal. We developed an online learning method that simultaneously discovers meaningful concepts in the associated processing streams, extending methods such as PCA, SOM or sparse coding to the multimodal case. In addition to the avoidance of redundancies in the concepts derived from single modalities, we claim that meaningful concepts are those who have statistical relations across modalities. This is a reasonable claim because measurements by different sensors often have common cause in the external world and therefore carry correlated information. To capture such cross-modal relations while avoiding redundancy of concepts, we propose a set of interacting self-organization processes which are modulated by local predictability. To validate the fundamental applicability of the method, we conducted a plausible simulation experiment with synthetic data and found that those concepts which are predictable from other modalities successively ”grow”, i.e., become overrepresented, whereas concepts that are not predictable become systematically under-represented. We additionally explored the applicability of the developed method to real-world robotics scenarios.

6.3.6. The contribution of context: a case study of object recognition in an intelligent car

Participants: Alexander Gepperth, Michael Garcia Ortiz.

In this work [84], we explored the potential contribution of multimodal context information to object detection in an ”intelligent car”. The used car platform incorporates subsystems for the detection of objects from local visual patterns, as well as for the estimation of global scene properties (sometimes denoted scene context or just context) such as the shape of the road area or the 3D position of the ground plane. Annotated data recorded on this platform is publicly available as the a ”HRI RoadTraffic” vehicle video dataset, which formed the basis for the investigation. In order to quantify the contribution of context information, we investigated whether it can be used to infer object identity with little or no reference to local patterns of visual appearance. Using a challenging vehicle detection task based on the ”HRI RoadTraffic” dataset, we trained selected algorithms (context models) to estimate object identity from context information alone. In the course of our performance evaluations, we also analyzed the effect of typical real-world conditions (noise, high input dimensionality, environmental variation) on context model performance. As a principal result, we showed that the learning of context models is feasible with all tested algorithms, and that object identity can be estimated from context information with similar accuracy as by relying on local pattern recognition methods. We also found that the use of basis function representations [1] (also known as ”population codes” allows the simplest (and therefore most efficient) learning methods to perform best in the benchmark, suggesting that the use of context is feasible even in systems operating under strong performance constraints.

6.3.7. Co-training of context models for real-time object detection

Participant: Alexander Gepperth.

In this work [85], we developed a simple way to reduce the amount of required training data in context-based models of real- time object detection and demonstrated the feasibility of our approach in a very challenging vehicle detection scenario comprising multiple weather, environment and light conditions such as rain, snow and darkness (night). The investigation is based on a real-time detection system effectively composed of two trainable components: an exhaustive multiscale object detector (signal-driven detection), as well as a module for generating object-specific visual attention (context models) controlling the signal-driven detection process. Both parts of the system require a significant amount of ground-truth data which need to be generated by human annotation in a time-consuming and costly process. Assuming sufficient training examples for signal-based detection, we showed that a co-training step can eliminate the need for separate ground-truth data to train context models. This is achieved by directly training context models with the results of signal-driven detection. We demonstrated that this process is feasible for different qualities of signal-driven detection, and maintains the performance gains from context models. As it is by now widely accepted that signal-driven object detection can be significantly improved by context models, our method allows to train strongly improved detection systems without additional labor, and above all, cost.

6.4. Robot Multimodal Learning of Language and Action

6.4.1. *Learning semantic components from sub-symbolic multi-modal perception*

Participants: Olivier Mangin, Caio Tomazelli Da Silva Oliveira, Pierre-Yves Oudeyer.

Perceptual systems often include sensors from several modalities. However, existing robots do not yet sufficiently discover patterns that are spread over the flow of multimodal data they receive. In this work we establish a framework to learns multimodal components from perception. We use a nonnegative matrix factorization algorithm to learn a dictionary of components that represent meaningful elements present in the multimodal perception, without providing the system with a symbolic representation of the semantics. In [53] we illustrate this framework by showing how a learner discovers word-like components from observation of gestures made by a human together with spoken descriptions of the gestures, and how it captures the semantic association between the two. These experiments were further extended during the internship of Caio Tomazelli Da Silva Oliveira. Importantly these experiments provide an example of language grounding into perception, and feature global understanding of a linguistic task without requiring its compositional understanding. The code of the experiments from [53] as well as the motion dataset have been made publicly available to improve the reproducibility of the experiments.

6.4.2. *Curiosity-driven exploration and interactive learning of visual objects with the ICub robot*

Participants: Mai Nguyen, Natalia Lyubova, Damien Gerardeaux-Viret, David Filliat, Pierre-Yves Oudeyer.

We studied how various mechanisms for cognition and learning, such as curiosity, action selection, imitation, visual learning and interaction monitoring, can be integrated in a single embodied cognitive architecture. We have conducted an experiment with the iCub robot for active recognition of objects in 3D through curiosity-driven exploration, in which the robot can manipulate the robot or ask a human user to manipulate objects to gain information and recognise better objects (fig. 27). For this experiment carried out within the MACSi project, we address the problem of learning to recognise objects in a developmental robotics scenario. In a life-long learning perspective, a humanoid robot should be capable of improving its knowledge of objects with active perception. Our approach stems from the cognitive development of infants, exploiting active curiosity-driven manipulation to improve perceptual learning of objects. These functionalities are implemented as perception, control and active exploration modules as part of the Cognitive Architecture of the MACSi project. We integrated a bottom-up vision system based on swift feature points and motor-primitive based robot control with the SGIM-ACTS algorithm (Socially Guided Intrinsic Motivation with Active Choice of Task and Strategy as the active exploration module. SGIM-ACTS is a strategic learner who actively chooses which task to concentrate on, and which strategy is better according to this task. It thus monitors the learning progress for each strategy on all kinds of tasks, and actively interacts with the human teacher. We obtained an active object recognition approach, which exploits curiosity to guide exploration and manipulation, such that the robot can improve its knowledge of objects in an autonomous and efficient way. Experimental results show the effectiveness of our approach: the humanoid iCub is now capable of deciding autonomously which actions must be performed on objects in order to improve its knowledge, requiring a minimal assistance from its caregiver. This work constitutes the base for forthcoming research in autonomous learning of affordances. This work have been published in a conference [57] and in a journal paper [28].

6.4.3. *Imitation Learning and Language*

Participants: Thomas Cederborg, Pierre-Yves Oudeyer.

We have studied how context-dependant imitation learning of new skills and language learning could be seen as special cases of the same mechanism. We argue that imitation learning of context-dependent skills implies complex inferences to solve what we call the ”motor Gavagai problem”, which can be viewed as a generalization of the so-called ”language Gavagai problem”. In a full symbolic framework where percepts and actions are continuous, this allows us to articulate that language may be acquired out of generic sensorimotor imitation learning mechanisms primarily dedicated at solving this motor Gavagai problem. Through the use of a computational model, we illustrate how non-linguistic and linguistic skills can be learnt concurrently, seamlessly, and without the need for symbols. We also show that there is no need to actually represent the distinction between linguistic and non-linguistic tasks, which rather appears to be in the eye of the observer of the system. This computational model leverages advanced statistical methods for imitation learning, where closed-loop motor policies are learnt from human demonstrations of behaviours that are dynamical responses to a multimodal context. A novelty here is that the multimodal context, which defines what motor policy to achieve, includes, in addition to physical objects, a human interactant which can produce acoustic waves (speech) or hand gestures (sign language). This was published in [26].

6.4.4. Learning to Interpret the Meaning of Teaching Signals in Socially Guided Robot Learning

Participants: Manuel Lopes, Jonathan Grizou, Thomas Cederborg, Pierre-Yves Oudeyer.

We elaborated an algorithm to bootstrap shared understanding in a human-robot interaction scenario where the user teaches a robot a new task using teaching instructions yet unknown to it. In such cases, the robot needs to estimate simultaneously what the task is and the associated meaning of instructions received from the user. For this work, we consider a scenario where a human teacher uses initially unknown spoken words, whose associated unknown meaning is either a feedback (good/bad) or a guidance (go left, right, ...). We present computational results, within an inverse reinforcement learning framework, showing that a) it is possible to learn the meaning of unknown and noisy teaching instructions, as well as a new task at the same time, b) it is possible to reuse the acquired knowledge about instructions for learning new tasks, and c) even if the robot initially knows some of the instructions' meanings, the use of extra unknown teaching instructions improves learning efficiency. Published articles: [43], [45].

An extension to the use of brain signals has been made [44]. Do we need to explicitly calibrate Brain Machine Interfaces (BMIs)? Can we start controlling a device without telling this device how to interpret brain signals? Can we learn how to communicate with a human user through practical interaction? It sounds like an ill posed problem, how can we control a device if such device does not know what our signals mean? This paper argues and present empirical results showing that, under specific but realistic conditions, this problem can be solved. We show that a signal decoder can be learnt automatically and online by the system under the assumption that both, human and machine, share the same a priori on the possible signals' meanings and the possible tasks the user may want the device to achieve. We present results from online experiments on a Brain Computer Interface (BCI) and a Human Robot Interaction (HRI) scenario.

6.4.5. Active Learning for Teaching a Robot Grounded Relational Symbols

Participants: Johannes Kulick, Tobias Lang, Marc Toussaint, Manuel Lopes.

The present work investigates an interactive teaching scenario, where a human aims to teach the robot symbols that abstract geometric (relational) features of objects. There are multiple motivations for this scenario: First, state-of-the-art methods for relational Reinforcement Learning demonstrated that we can successfully learn abstracting and well-generalizing probabilistic relational models and use them for goal-directed object manipulation. However, these methods rely on given grounded action and state symbols and raise the classical question Where do the symbols come from? Second, existing research on learning from human-robot interaction has focused mostly on the motion level (e.g., imitation learning). However, if the goal of teaching is to enable the robot to autonomously solve sequential manipulation tasks in a goal-directed manner, the human should have the possibility to teach the relevant abstractions to describe the task and let the robot eventually leverage powerful relational RL methods (see Figure 33). We formalize human-robot teaching of grounded symbols as an Active Learning problem, where the robot actively generates geometric situations that maximize his information gain about the symbol to be learnt. We demonstrate that the learned symbols can be used in a relational RL framework for the robot to learn probabilistic relational rules and use them to solve object manipulation tasks in a goal-directed manner. [47].

../../../../projets/flowers/IMG/activesymbollearn.png

Figure 33. Active learning of symbol descriptions on a real world robot.

6.5. Robot Design and Morphological Computation

6.5.1. The Poppy Humanoid Robot: Leg Design for Biped Locomotion

Participants: Matthieu Lapeyre, Pierre Rouanet, Pierre-Yves Oudeyer.

In this paper introduced for Poppy as a novel humanoid robotic platform designed to jointly address three central goals of humanoid robotics: 1) study the role of morphology in biped locomotion; 2) study full-body compliant physical human-robot interaction; 3) be robust while easy and fast to duplicate to facilitate experimentation. The taken approach relies on functional modeling of certain aspects of human morphology, optimizing materials and geometry, as well as on the use of 3D printing techniques. In this article, we have focused on the presentation of the design of specific morphological parts related to biped locomotion: the hip, the thigh, the limb mesh and the knee. We also presented an initial experiments showing properties of the robot when walking with the physical guidance of a human. [50].

6.5.2. Poppy Humanoid Platform: Experimental Evaluation of the Role of a Bio-inspired Thigh Shape

Participants: Matthieu Lapeyre, Pierre Rouanet, Pierre-Yves Oudeyer.

In this paper, we present an experimental evaluation of the role of the morphology in the Poppy humanoid platform. More precisely, we have investigated the impact of the bio-inspired thigh, bended of 6^{o}, on the balance and biped locomotion. We compare this design with a more traditional straight thigh. We describe both the theoretical model and real experiments showing that the bio-inspired thigh allows the reduction of falling speed by almost 60% (single support phase) and the decrease of the lateral motion needed for the mass transfer from one foot to the other by 30% (double support phase). We also present an experiment where the robot walks on a treadmill thanks to the social and physical guidance of expert users and we show that the bended thigh reduces the upper body motion by about 45% indicating a more stable walk.[48].

6.5.3. Morphological computation and body intelligence

6.5.3.1. Comparative Study of the Role of Trunk in Human and Robot Balance Control

Participants: Matthieu Lapeyre [correspondant], Christophe Halgand, Jean-Renã© Cazalet, Etienne Guilllaud, Pierre-Yves Oudeyer.

../../../../projets/flowers/IMG/marche_bended.jpg ../../../../projets/flowers/IMG/marche_straig

Figure 34. Five pictures have been taken while Poppy was walking and were stacked to obtain a qualitative view of the difference in the walking behavior in function of the morphology of the thigh.

Numerous studies in the field of functional motor rehabilitation were devoted to understanding the functioning of members but few are interested in the coordination of the trunk muscles and the relationship between axial and appendicular motricity which is essential in maintaining balance during travel. Acquiring new knowledge on this subject is a prerequisite in the development of new therapeutic strategies to restore motor function to the overall development of robotic orthosis that would assist the movement. Many robotic orthosis using EMG signals were unfortunately using few joints [82] and a system for controlling a multi articulated spine has not yet been developed. We propose here to use a multidisciplinary approach to define the neuro-mechanical principles where an axial system is operating in synergy with human and robot limbs.

To bring us a theoretical framework, we chose to study the reactions of the Acroban humanoid robot. Including 5 joints in the trunk, Acroban can reproduce in part the fluid movements of the human body [101] and especially to test its behavior when its trunk is held fixed or his arms are no longer used for rebalance. To disrupt postural balance in humans and robots, we have developed a low cost mobile platform (see Figure 35). This platform is made up of a broad stable support $(0.8x5m)$ mounted on a skateboard having a power of 800W. The substitution of the initial order of skate by an embedded microcontroller allows us to generate mono-axial perturbations precise intensity and duration to ensure repeatability of the disturbance. We capture movements (Optitrack 250Hz) and record the acceleration of the platform (accelerometer embedded 2kHz), the center of pressure (WiiBalanceBoard 60Hz), and electromyography (EMG).

The experimental device (mobile platform and synchronized recordings) is operational. Preliminary experiments have allowed us to refine the profiles of disturbance on the robot Acroban. The analysis of preliminary results is in progress. Following this study, we hope to improve the modeling of the motor system in humans and robotic simulation as a basis for the development of robotic orthosis axial system. Second, the results provide a basis for improved balancing of Acroban primitives but also the development of future humanoid robots.

6.6. Educational Technologie

6.6.1. KidLearn: Adaptive Personalization of Educational Content with Machine Learning

Kidlearn is a research project studying how machine learning can be applied to intelligent tutoring systems. It aims at developing methodologies and software which adaptively personalize sequences of learning activities

Figure 35. Experimental setup for comparative study of the role of the trunk in human and robot balance control

to the particularities of each individual student. Our systems aim at proposing to the student the right activity at the right time, maximizing concurrently his learning progress and its motivation. In addition to contributing to the efficiency of learning and motivation, the approach is also made to reduce the time needed to design ITS systems.

Intelligent Tutoring System (ITS) are computer environments designed to guide students in their learning. Through the proposal of different activities, it provides teaching experience, guidance and feedback to improve learning. The FLOWERS team has developed several computational models of artificial curiosity and intrinsic motivation based on research on psychology that might have a great impact for ITS. Results showed that activities with intermediate levels of complexity, neither too easy nor too difficult but just a little more difficult that the current level, provide better teaching experiences. The system is based on the combination of three approaches. First, it leverages Flowers team's recent models of computational models of artificial curiosity and intrinsic motivation based on research in psychology and neuroscience. One overview can be be found in [27]. Second, it uses state-of-the-art Multi-Arm Bandit (MAB) techniques to efficiently manage the exploration/exploitation challenge of this optimization process. Third, it leverages expert knowledge to constrain and bootstrap initial exploration of the MAB, while requiring only coarse guidance information of the expert and allowing the system to deal with didactic gaps in its knowledge. In 2013, we have run a first pilot experiment in elementary schools of Région Aquitaine, where 7-8 year old kids could learn elements of mathematics thanks to an educational software that presented the right exercises at the right time to maximize learning progress. A report is available at: http://arxiv.org/pdf/1310.3174v1.pdf.

6.7. Other applications

6.7.1. Real-time Reaction-Diffusion Simulation: a Machine Learning Technique

Participants: Thomas Degris, Nejib Zemzemi.

Carmen is an Inria team working on modeling the electrical activity of the human heart. Their models are mainly based on reaction-diffusion equations. These methods are expansive in terms of computational costs which limits their use in practice. More specifically, some recent chirurgical intervention techniques on the heart (atrial ablation) requires to identify the source of the electrical wave. Finding such sources requires an optimization procedure. Using classical methods, this procedure is very heavy computationally.

In this project, our goal is to reduce the computational cost using supervised learning techniques. The idea is to replace the incremental resolution of partial differential equations by more suitable data structures for real-time running. Starting from data generated by simulating different excitations scenari on a human atria, this data is afterwords used as a training data set for machine learning algorithms. This approach will allow a faster optimization procedure.

This work is in collaboration with Nejib Zemzemi from the Inria Carmen team. This project is in preliminary steps.

6.7.2. Appearance-based segmentation of indoors/outdoors sequences of spherical views

Participant: David Filliat.

In collaboration with Patrick Rives and Alexandre Chapoulie from the Arobas team at Inria Sophia-Antipolis, we developped a method for environment segmentation based on spherical views [41]. Navigating in large scale, complex and dynamic environments requires reliable representations able to capture metric, topological and semantic aspects of the scene for supporting path planing and real time motion control. In a previous work, we addressed metric and topological representations thanks to a multi-cameras system which allows building of dense visual maps of large scale 3D environments. The map is a set of locally accurate spherical panoramas related by 6dof poses graph. The work presented here is a further step toward a semantic representation. We aim at detecting the changes in the structural properties of the scene during navigation. Structural properties are estimated online using a global descriptor relying on spherical harmonics which are particularly well-fitted to capture properties in spherical views. A change-point detection algorithm based on a statistical Neyman-Pearson test allows us to find optimal transitions between topological places. Results are presented and discussed both for indoors and outdoors experiments.

6.7.3. Modelling Stop Intersection Approaches using Gaussian Processes

Participant: David Filliat.

In collaboration with Javier-Ibanez Guzman and Alexandre Armand from Renault, we developped an approach toward the development of an electronic co-pilot adapted to the driver behavior [39]. Indeed, each driver reacts differently to the same traffic conditions, however, most Advanced Driving Assistant Systems (ADAS) assume that all drivers are the same. This work proposes a method to learn and to model the velocity profile that the driver follows as the vehicle decelerates towards a stop intersection. Gaussian Processes (GP), a machine learning method for non-linear regressions are used to model the velocity profiles. It is shown that GP are well adapted for such an application, using data recorded in real traffic conditions. GP allow the generation of a normally distributed speed, given a position on the road. By comparison with generic velocity profiles, benefits of using individual driver patterns for ADAS issues are presented.

FLUMINANCE Project-Team

6. New Results

6.1. Fluid motion estimation

6.1.1. Stochastic uncertainty models for motion estimation

Participants: Sébastien Beyou, Etienne Mémin, Emmanuel Saunier.

In this study we have proposed a stochastic formulation of the brightness consistency used principally in motion estimation problems. In this formalization the image luminance is modeled as a continuous function transported by a flow known only up to some uncertainties. Stochastic calculus then enables to built conservation principles which take into account the motion uncertainties. These uncertainties defined either from isotropic or anisotropic models can be estimated jointly to the motion estimates. Such a formulation besides providing estimates of the velocity field and of its associated uncertainties allows us to naturally define a linear multiresolution scale-space framework. The corresponding estimator, implemented within a local least squares approach, has shown to improve significantly the results of the corresponding deterministic estimator (Lucas and Kanade estimator). This fast local motion estimator provides results that are of the same order of accuracy than state-of-the-art dense fluid flow motion estimator for particle images. The uncertainties estimated supply a useful piece of information in the context of data assimilation. This ability has been exploited to define multiscale incremental data assimilation filtering schemes. This work has been recently published in Numerical Mathematics: Theory, Methods and Applications [14]. It is also described in Sébastien Beyou's PhD dissertation [11]. The development of an efficient GPU based version of this estimator recently started through the Inria ADT project FLUMILAB

6.1.2. 3D flows reconstruction from image data

Participants: Ioana Barbu, Kai Berger, Cédric Herzet, Etienne Mémin.

Our work focuses on the design of new tools for the estimation of 3D turbulent flow motion in the experimental setup of Tomo-PIV. This task includes both the study of physically-sound models on the observations and the fluid motion, and the design of low-complexity and accurate estimation algorithms. On the one hand, we investigate state-of-the-art methodologies such as “sparse representations" for the characterization of the observation and fluid motion models. On the other hand, we place the estimation problem into a probabilistic Bayesian framework and use state-of- the-art inference tools to effectively exploit the strong time-dependence on the fluid motion. In our previous work, we have focussed on the problem of reconstructing the particle positions from several two-dimensional images. Our approach was based on the exploitation of a particular family of sparse representation algorithms, leading to a good trade-off between performance and complexity. Moreover, we also tackled the problem of estimating the 3D velocity field of the fluid flow from two instances of reconstructed volumes of particles. Our approach was based on a generalization of the well-known Lucas-Kanade's motion estimator to 3D problems. A potential strength of the proposed approach is the possibility to consider a fully parallelized (and therefore very fast) hardware implementation. This year, we have focused on the design of new methodologies to jointly estimate the volume of particles and the velocity field from the received image data. Our approach is based on the minimization (with respect to both the position of the particles and the velocity field) of a cost function penalizing both the discrepancies with respect to a conservation equation and some prior estimates of particle positions. This work has led to one publication in an international conference (PIV13) [27] and one publication in a national conference (Fluvisu13) [31].

Since October 2013, with our new postdoctoral fellow Kai Berger, we have started a new direction of research targeting the volume reconstruction problem. In particular, we address the question of devising effective reconstruction procedures taking into account the limited computational budget available in practice. Our approach is based on the design of simple thresholding operators, allowing to reduce the dimension of the initial problem and amenable to fast parallel implementations.

6.1.3. Motion estimation techniques for turbulent fluid flows

Participants: Patrick Héas, Dominique Heitz, Cédric Herzet, Etienne Mémin.

In this study we have devised smoothing functional adapted to the multiscale structure of homogeneous turbulent flows. These regularization constraints ensue from a classical phenomenological description of turbulence. The smoothing is in practice achieved by imposing some scale-invariance principles between histograms of motion increments computed at different scales. Relying on a Bayesian formulation, an inference technique, based on likelihood maximization and marginalization of the motion variable, has been proposed to jointly estimate the fluid motion, the regularization parameters and a proper physical models. The performance of the proposed Bayesian estimator has been assessed on several image sequences depicting synthetic and real turbulent fluid flows. The results obtained in the context of fully developed turbulence show that an improvement in terms of small-scale motion estimation can be achieved as compared to classical motion estimator. This work, performed within a collaboration with Pablo Mininni from the university of Buenos Aires, have been published in the IEEE Transactions on Pattern Analysis And Machine Learning [22].

6.1.4. Wavelet basis for multiscale motion estimation

Participants: Patrick Héas, Cédric Herzet, Etienne Mémin.

In this study we focused on the implementation of a simple wavelet-based optical-flow motion estimator dedicated to the recovery of fluid motions. The wavelet representation of the unknown velocity field is considered. This scale-space representation, associated to a simple gradient-based optimization algorithm, sets up a natural multiscale/multigrid optimization framework for the optical flow estimation that can be combined to more traditional incremental multiresolution approaches. Moreover, a very simple closure mechanism, approximating locally the solution by high-order polynomials, is provided by truncating the wavelet basis at intermediate scales. This offers a very interesting alternative to traditional Particle Image Velocimetry techniques. As another alternative to this medium-scale estimator, we explored strategies to define estimation at finer scales. These strategies rely on the encoding of high-order smoothing functional on divergence free wavelet basis. This study has been published in the journal of Numerical Mathematics: Theory, Methods and Applications [19] and in the international Journal of Computer Vision [23]. This work has strongly benefited from a collaboration with Souleyman Kadri-Harouna (University of La Rochelle and who was formerly on a post-doctoral position in our team). The divergence free wavelets basis proposed in [24] constitutes the building blocks on which we have elaborated our wavelet based motion estimation solutions. We have otherwise pursue our collaboration with Chico university through the post-doc of Pierre Dérian on the GPU implementation of such motion estimator for Lidar data.

6.1.5. Sparse-representation algorithms

Participant: Cédric Herzet.

The paradigm of sparse representations is a rather new concept which turns out to be central in many domains of signal processing. In particular, in the field of fluid motion estimation, sparse representation appears to be potentially useful at several levels: i) it provides a relevant model for the characterization of the velocity field in some scenarios; ii) it plays a crucial role in the recovery of volumes of particles in the 3D Tomo-PIV problem.

Unfortunately, the standard sparse representation problem is known to be NP hard. Therefore, heuristic procedures have to be devised to try and access to the solution of this problem. Among the popular methods available in the literature, one can mention orthogonal matching pursuit, orthogonal least squares and the family of procedures based on the minimization of ℓ_p norms. In order to assess and improve the performance of these algorithms, theoretical works have been undertaken in order to understand under which conditions these procedures can succeed in recovering the "true" sparse vector.

This year, we have contributed to this research axis by deriving conditions of success for the algorithms mentioned above when some partial information is available about the position of the nonzero coefficients in the sparse vector. This paradigm is of interest in the Tomographic-PIV volume reconstruction problem: one can indeed expect volumes of particles at two successive instants to be pretty similar; any estimate of the position of the particles at one given instant can therefore serve as a prior estimate about their position at the next instant. The conditions of success of such procedure have been rigorously formalized in two publications in the IEEE Transactions on Information Theory [21], [26] and one publication in an international conference (SPARS13) [28].

6.2. Tracking, Data assimilation and model-data coupling

6.2.1. Stochastic filtering for fluid motion tracking

Participants: Sébastien Beyou, Anne Cuzol, Etienne Mémin.

Within the PhD thesis of Sébastien Beyou [11], we investigated the study of a recursive Bayesian filter for tracking velocity fields of fluid flows. We resort in this study to Monte-Carlo approximations based on the particle filtering paradigm. In particular, we investigated the use of the so-called ensemble Kalman filtering for fluid tracking problems. This kind of filters introduced for the analysis of geophysical fluids is based on the Kalman filter update equations. Nevertheless, unlike traditional Kalman filtering setting, the covariances of the estimation errors, required to compute the so-called Kalman gain, relies on an ensemble of forecasts. Such a process gives rise to a Monte-Carlo approximation for a family of non-linear stochastic filters enabling to handle state spaces of large dimension. The method we proposed can be seen as an extension of this technique that combines sequential importance sampling and the propagation law of an ensemble Kalman filter. This technique leads to an ensemble Kalman filter with an improved efficiency. Within this type of scheme, we have in particular investigated the introduction of a nonlinear direct image measurement operator. This modification of the filter provides very good results on 2D numerical and experimental flows even in the presence of strong noises. We assessed successfully its application to oceanic satellite images for the recovering of ocean streams. We have also studied the impact on the stochastic dynamics of auto-similar Gaussian noise mimicking statistical properties of turbulence and the introduction within an incremental ensemble analysis scheme of multiscale motion measurements. This work has been published in the Tellus A journal [15].

6.2.2. Stochastic filtering technique for the tracking of closed curves

Participant: Etienne Mémin.

We have studied a stochastic filtering technique for the tracking of closed curves along an image sequence. In that aim, we designed a continuous-time stochastic dynamics that allows us to infer inter-frame deformations. The curve is defined by an implicit level-set representation and the stochastic dynamics is expressed properly on the level-set function. It takes the form of a stochastic partial differential equation with a Brownian motion of low dimension. The evolution model we proposed combines local photometric information, deformations induced by the curve displacement and an uncertainty modeling of the dynamics. Specific choices of noise models and drift terms lead to an evolution law based on mean curvature as in classic level set methods, while other choices yield new evolution laws. The approach we propose is implemented through a particle filter, which includes color measurements characterizing the target and the background photometric probability densities respectively. The merit of this parameter free filter is demonstrated on various satellite image sequences depicting the evolution of complex geophysical flows. This work has been recently published in the Journal of Mathematical Imaging and Vision [13]. Let us note the method provides an empirical dynamical model learned recursively from a data flow. Its short time forecasting skills have been used in the context of weather-watch radar images within a fruitful collaboration with MeteoFrance.

6.2.3. Sequential smoothing for fluid motion

Participants: Anne Cuzol, Etienne Mémin.

In parallel to the construction of stochastic filtering techniques for fluid motions, we have proposed a new sequential smoothing method within a Monte-Carlo framework. This smoothing aims at reducing the temporal discontinuities induced by the sequential assimilation of discrete time data into continuous time dynamical models. The time step between observations can indeed be long in environmental applications for instance, and much longer than the time step used to discretize the model equations. While the filtering aims at estimating the state of the system at observations times in an optimal way, the objective of the smoothing is to improve the estimation of the hidden state between observation times. The method is based on a Monte-Carlo approximation of the filtering and smoothing distributions, and relies on a simulation technique of conditional diffusions. The proposed smoother can be applied to general non linear and multidimensional models. It has been applied to a turbulent flow in a high-dimensional context, in order to smooth the filtering results obtained from a particle filter with a proposal density built from an Ensemble Kalman procedure. This conditional simulation framework can also be used for filtering problem with low measurement noise. This has been explored through a collaboration with Jean-Louis Marchand (ENS Bretagne) in the context of vorticity tracking from image data.

6.2.4. Stochastic fluid flow dynamics under uncertainty

Participants: Etienne Mémin, Valentin Resseguier.

In this research axis we aim at devising Eulerian expressions for the description of fluid flow evolution laws under uncertainties. Such an uncertainty is modeled through the introduction of a random term that allows taking into account large-scale approximations or truncation effects performed within the dynamics analytical constitution steps. This includes for instance the modeling of unresolved scales interaction in large eddies simulation (LES) or in Reynolds average numerical simulation (RANS), but also uncertainties attached to non-uniform grid discretization. This model is mainly based on a stochastic version of the Reynolds transport theorem. Within this framework various simple expressions of the drift component can be exhibited for different models of the random field carrying the uncertainties we have on the flow. We aim at using such a formalization within image-based data assimilation framework and to derive appropriate stochastic versions of geophysical flow dynamical modeling. This formalization has been published in the journal Geophysical and Astrophysical Fluid Dynamics [25]. Numerical simulation on divergence free wavelets basis of 3D viscous Taylor-Green vortex and Crow instability have been performed within a collaboration with Souleymane Kadri-Harouna. First promising results have been published in the TSFP8 conference [30]. Besides, we explore in the context of Valentin Resseguier's PhD the extension of such framework to oceanic models and to satellite image data assimilation. This PhD thesis takes place within a fruitful collaboration with Bertrand Chapron (CERSAT/IFREMER).

6.2.5. Free surface flows reconstruction and tracking

Participants: Dominique Heitz, Etienne Mémin, Cordelia Robinson, Yin Yang.

Characterizing a free-surface flow (space and time-dependent velocity and geometry) given observations/measures at successive times is an ubiquitous problem in fluid mechanic and in hydrology. Observations can consist of e.g. measurements of velocity, or like in this work of measurements of the geometry of the free-surface. Indeed, recently developed depth/range sensors allow to capture directly a rough 3D geometry of surfaces with high space and time resolution. We have investigated the performance of the Kinect and have shown that it is likely to capture temporal sequences of depth observations of wave-like surfaces with wavelengths and amplitudes sufficiently small to characterize medium/large scale flows. Several data assimilation methods have been experimented and compared to estimate both time dependent geometry and displacement field associated to a free-surface flow from a temporal sequence of Kinect data. This study has been conducted on synthetic and real-world data. Finally, we explored the application of such techniques to hydrological applications. These results have been recently submitted to Journal of Computational Physics.

6.2.6. Variationnal ensemble methods for data assimilation

Participants: Dominique Heitz, Etienne Mémin, Cordelia Robinson, Yin Yang.

In this work, we aim at studying an ensemble based optimal control strategy for data assimilation. Such a formulation nicely combines the ingredients of ensemble Kalman filters and variational assimilation. In the same way as standard variational assimilation, it is formulated as the minimization of an objective function. However, similarly to ensemble filters, it introduces in its objective function an empirical ensemble-based background-error covariance and works in an off-line smoothing mode rather than sequentially like filtering approaches in a sequential filter. These techniques have the great advantage to avoid the introduction of tangent linear and adjoint models, which are necessary for standard incremental variational techniques. As the background error covariance matrix plays a key role in the variational process, our study particularly focuses on the generation of the analysis ensemble state with localization techniques. We compared the performances of both methods in different cases in which the system's component are fully observed or only partially. The comparisons have been leaded on the basis of a Shallow Water model.

6.2.7. Optimal control techniques for the coupling of large scale dynamical systems and image data

Participants: Dominique Heitz, Etienne Mémin, Cordelia Robinson.

This work aims at investigating the use of optimal control techniques for the coupling of Large Eddies Simulation (LES) techniques and 2D image data. The objective is to reconstruct a 3D flow from a set of simultaneous time resolved 2D image sequences visualizing the flow on a set of 2D plans enlightened with laser sheets. This approach will be experimented on shear layer flows and on wake flows generated on the wind tunnel of Irstea Rennes. Within this study we wish also to explore techniques to enrich large-scale dynamical models by the introduction of uncertainty terms or through the definition of subgrid models from the image data. This research theme is related to the issue of turbulence characterization from image sequences. Instead of predefined turbulence models, we aim here at tuning from the data the value of coefficients involved in traditional LES subgrid models or in longer-term goal to learn empirical subgrid models directly from image data. An accurate modeling of this term is essential for Large Eddies Simulation as it models all the non resolved motion scales and their interactions with the large scales.

We have pursued the first investigations on a 4DVar assimilation technique, integrating PIV data and Direct Numerical Simulation (DNS), to reconstruct two-dimensional turbulent flows. The problem we are dealing with consists in recovering a flow obeying Navier-Stokes equations, given some noisy and possibly incomplete PIV measurements of the flow. By modifying the initial and inflow conditions of the system, the proposed method reconstructs the flow on the basis of a DNS model and noisy measurements. The technique has been evaluated in the wake of a circular cylinder. It denoises the measurements and increases the spatiotemporal resolution of PIV time series. These results have been recently published in the Journal of Computational Physics [20]. A paper has been also recently published on the denoising aspect in the (PIV13) international conference [29]. Along the same line of studies the 3D case is ongoing. The goal consists here to reconstruct a 3D flow from a set of simultaneous time resolved 2D images of planar sections of the 3D volume. This work is mainly conducted within the PhD of Cordelia Robinson. The development of the variational assimilation code has been initiated within a collaboration with A. Gronskis, S. Laizé (lecturer, Imperial College, UK) and Eric Lamballais (institut P' Poitiers). A High Reynolds number simulation of the wake behind a cylinder has been recently performed within this collaboration.

6.2.8. Variational assimilation of images for large scale fluid flow dynamics with uncertainty

Participant: Etienne Mémin.

In this work we explore the assimilation of a large scale representation of the flow dynamics with image data provided at a finer resolution. The velocity field at large scales is described as a regular smooth components whereas the complement component is a highly oscillating random velocity field defined on the image grid but living at all the scales. Following this route we have started to assess the performance of a variational assimilation technique with direct image data observation. Preliminary encouraging results obtained for a wavelet-based 2D Navier Stokes implementation and images of a passive scalar transported by the flow have been obtained. Large-scale simulation under uncertainty for the 3D viscous Taylor-Green vortex flow have been carried out and show promising results of the approach.

6.2.9. Reduced-order models for flows representation from image data

Participants: Cédric Herzet, Etienne Mémin, Véronique Souchaud.

One of the possibilities to neglect the influence of some degrees of freedom over the main characteristics of a flow consists in representing it as a sum of K orthonormal spatial basis functions weighted with temporal coefficients. To determine the basis function of this expansion, one of the usual approaches relies on the Karhunen-Loeve decomposition (refered to as proper orthogonal decomposition – POD – in the fluid mechanics domain). In practice, the spatial basis functions, also called modes, are the eigenvectors of an empirical auto-correlation matrix which is built from “snapshots" of the considered physical process.

In this axis of work we focus on the case where one does not have a direct access to snapshots of the considered physical process. Instead, the POD has to be built from the partial and noisy observation of the physical phenomenon of interest. Instances of such scenarios include situations where real instantaneous vector-field snapshots are estimated from a sequence of images. We have been working on several approaches dealing with such a new paradigm. A first approach consists in extending standard penalized motion-estimation algorithms to the case where the sought velocity field is constrained to span a low-dimensional subspace. In particular, we have considered scenarios where the standard optical flow constraint (OFC) is no longer satisfied and one has therefore to resort to a Discrete Finite Difference (DFD) model. The non-linearity of the latter leads to several practical issues that we have addressed this year.

Within a collaboration with the University of Buenos Aires, we have also explored, a method that combines Proper Orthogonal Decomposition with a spectral technique to analyze and extract reduced order models of flows from time resolved data of velocity fields. This methodology, relying on the eigenfunctions of the Koopman operator, is specifically adapted to flows with quasi periodic orbits in the phase space. The technique is particularly suited to cases requiring a discretization with a high spatial and temporal resolution. The proposed analysis enables to decompose the flow dynamics into modes that oscillate at a single frequency. For each modes an energy content and a spatial structure can be put in correspondence. This approach has been assessed for a wake flow behind a cylinder at Reynolds number 3900 and has been recently published to the journal of Theoretical and Computational Fluid Dynamics [16]. The assessment of this method on oceanic model simulation data is on going.

6.3. Analysis and modeling of turbulent flows

6.3.1. Hot-wire anemometry at low velocities

Participant: Dominique Heitz.

A new dynamical calibration technique has been developed for hot-wire probes. The technique permits, in a short time range, the combined calibration of velocity, temperature and direction calibration of single and multiple hot-wire probes. The calibration and measurements uncertainties were modeled, simulated and controlled, in order to reduce their estimated values. Based on a market study the french patent application has been extended this year to a Patent Cooperation Treaty (PCT) application.

6.3.2. Numerical and experimental image and flow database

Participant: Dominique Heitz.

The goal was to design a database for the evaluation of the different techniques developed in the Fluminance group. The main challenge was to enlarge a database mainly based on two-dimensional flows, with three-dimensional turbulent flows. New synthetic image sequences based on homogeneous isotropic turbulence and on circular cylinder wake have been provided. These images have been completed with real image sequences based on wake and mixing layers flows. This new database provides different realistic conditions to analyse the performance of the methods: time steps between images, level of noise, Reynolds number, large-scale images. A Wake flow at high Reynolds number has been also simulated on one of the IDRISS super computer. This simulation, whose results analysis is on going, has been performed within a collaboration with Sylvain Laizet (Imperial College).

6.4. Visual servoing approach for fluid flow control

6.4.1. Minimization of the kinetic energy density in the 2D plane Poiseuille flow

Participants: Christophe Collewet, Xuan Quy Dao.

This works concerns the PhD thesis of Xuan-Quy Dao. This year we have focused on a way to ensure a strict decreasing of the kinetic energy density. In that purpose, we have first proposed an approach to increase the controlled degrees of freedom. Indeed, the classical way to model this flow leads to only two degrees of freedom. With so few degrees of freedom it is obviously impossible to reach high desired performances as the strict minimization of the kinetic energy density. This way to proceed leads to a better minimization of the kinetic energy density. We have also proposed on approach based on a local decoupling of the controlled degree of freedom of the system so that an exponential decoupled decrease of each components of the state vector is locally obtained. This work has been presented at the CFM conference (Congrés Français de Mécanique) [32].

6.4.2. Control behind a backward-facing step

Participant: Christophe Collewet.

This work is performed in the context of the PhD thesis of Nicolas Gautier from ESCPI in collaboration with J.L. Aider. The separated flow downstream a backward-facing step is studied using visual information for feedback. More precisely, flow velocity fields are computed from a real time optical flow algorithm. The control law we used is a simple PID controller. Even a better control law could be used, this study validates that visual servo control is an effective approach to control a flow.

6.4.3. Control of systems described by partial differential equations

Participants: Tudor-Bogdan Airimitoaie, Christophe Collewet.

This work concerns principally the post-doctoral research of Tudor-Bogdan Airimiţoaie. It aims at controlling continuously evolving systems described by partial differential equations (PDEs). This is relevant in the context of the Fluminance team because fluid flows are infinite dimensional systems and can be rigorously described only through PDEs. In spite of this, practical approaches of flow control are based on low order numerical implementation relying on space and time discretization of the continuous system. This implies to setup strategies for model reduction that must be then in return properly understood with respect to the convergence of the control law. For finite dimensional implementations, one of the research directions pursued concerns the study on the benefit of increasing the controlled degrees of freedom (see the work of Xuan-Quy Dao). Another research direction, started recently, consists in improving control by using real-time estimation of a finite number of parameters related to the original infinite dimensional system. Indeed, this opens the possibility of improving performances by using more advanced robust linear parametric varying (LPV) control techniques existing in the literature. Two conference papers on these works have been submitted at the "7th AIAA Flow Control Conference".

FOCUS Project-Team

6. New Results

6.1. Service-oriented computing

Participants: Mario Bravetti, Ivan Lanese, Fabrizio Montesi, Gianluigi Zavattaro.

6.1.1. Primitives

We have studied primitives used in the context of service-oriented computing, at different levels of abstraction and in different contexts. At the abstract level, we considered both standard web services and Internet of Things, where computational and communication capabilities are attached to real-world objects such as smartphones or alarm clocks. For web services, we defined SSCC (Stream-based Service-Centered Calculus) [17], a calculus allowing to describe both service composition (orchestration) via streams and the protocols that services follow when invoked (conversation). We assessed the expressive power of SSCC by modeling van der Aalst's workflow patterns and an automotive case study from the European project Sensoria. For analysis, we presented a simple type system ensuring compatibility of client and service protocols. We also studied the behavioral theory of the calculus, highlighting some axioms that capture the behavior of the different primitives. As a final application of the theory, we defined and proved correct some program transformations. For Internet of Things, a main contribution [37] has been the definition of a calculus and of an equivalence allowing to capture the behavior of the system as seen by the human end-user. Since this equivalence is not compositional we defined also a finer equivalence which is compositional. We showed how our equivalences can be applied to reason on simple Internet of Things examples.

At a more concrete level, we have continued to study and extend the Jolie language. In [44] we present a detailed description of the Jolie language. We put our emphasis on how Jolie can deal with heterogeneous services. On the one hand, Jolie combines computation and composition primitives in an intuitive and concise syntax. On the other hand, the behavior and deployment of a Jolie program are orthogonal: they can be independently defined and recombined as long as they have compatible typing. In [42] we extended Jolie to model process-aware web information systems. Our major contribution is to offer a unifying approach for the programming of distributed architectures based on HTTP that support typical features of the process-oriented paradigm, such as structured communication flows and multiparty sessions.

6.1.2. Choreographies

Choreographies are high-level descriptions of distributed interacting systems featuring as basic unit a communication between two participants. A main feature of choreographies is that they are deadlock-free by construction. From a choreography one can automatically derive the behavior of each participant using a notion of projection. Under suitable conditions on the structure of a choreography, the correctness of its projection can be established in terms of a trace-based semantics. In [24] we have proposed a purely-global programming model. The idea is to define abstract choreographies – called protocol specifications – and use them to type a more concrete choreography. This more concrete choreography is used to generate executable code for the different participants. The approach is based on a novel interpretation of asynchrony and parallelism. We evaluated the approach by providing a prototype implementation for a concrete programming language and by applying it to some examples from multicore and service-oriented programming [49]. In [43] we tackled one of the main limitations of choreographies, namely the fact that they model closed systems. To this end we proposed a notion of composable choreographies. The key of our approach is the introduction of partial choreographies, which can mix global descriptions with communications among external peers. We prove that if two choreographies are composable, then the endpoints independently generated from each choreography are also composable, preserving their typability and deadlock-freedom. In [39] we showed how to transform choreographies which do not satisfy the conditions for their projection into choreographies that satisfy them preserving their behavior and enabling a correct projection.

6.2. Models for reliability

Participants: Ivan Lanese, Michael Lienhardt, Gianluigi Zavattaro.

6.2.1. Reversibility

We have continued the study of reversibility started in the past years, aimed at developing programming abstractions for reliable distributed systems. In [38] we present croll-pi, a concurrent calculus extending roll-pi – an higher-order pi-calculus featuring a rollback operator – allowing the specification of alternatives to a computation to be used upon rollback. Alternatives in croll-pi are attached to messages. We show the robustness of this mechanism by encoding more complex idioms for specifying alternatives. We illustrate the expressiveness of our approach by encoding a calculus of communicating transactions and by modeling the 8-queens problem. We also formally prove that croll-pi is strictly more expressive than roll-pi.

6.2.2. Compensations

We have continued the study of the expressive power of primitives for specifying compensations in long running transactions. Dynamic compensation installation allows for easier specification of fault handling in complex interactive systems since it enables to update the compensation policies according to run-time information. In [40] we show that in a simple pi-like calculus with static compensations the termination of a process is decidable, but it is undecidable in one with dynamic compensations. We then consider three commonly used patterns for dynamic compensations: parallel compensations, where new compensation items can only be added in parallel, replacing compensations, where old compensations are replaced, and nested compensations, where old compensations can be used (linearly) to build new ones. We show that termination is decidable in the first two cases and undecidable in the last one.

6.3. Cloud Computing

Participants: Elena Giachino, Michael Lienhardt, Tudor Alexandru Lascu, Jacopo Mauro, Gianluigi Zavattaro.

6.3.1. Languages for cloud applications

To foster the industrial adoption of virtualized services, it is necessary to address two important problems: (1) the efficient analysis, dynamic composition of services with qualitative and quantitative service levels and (2) the dynamic control of resources such as storage and processing capacities according to the internal policies of the services. Current technologies for cloud computing, addresses these problems at deployment and run time. The ENVISAGE project and the position paper [20] proposes, on the contrary, to overcome these problems by leveraging service-level agreements into software models and resource management into early phases of service design.

6.3.2. Models for cloud application deployment

Cloud computing offers the possibility to build sophisticated software systems on virtualized infrastructures at a fraction of the cost necessary just few years ago, but deploying/maintaining/reconfiguring such software systems is a serious challenge. The AEOLUS project, aims to tackle the scientific problems that need to be solved in order to ease the problem of efficient and cost-effective deployment and administration of the complex distributed architectures which are at the heart of cloud applications [25]. In particular, it is necessary to define appropriate models for the representation of the interdependencies among the software components of a cloud application as well as declarative languages for the specification of the desired application configuration. We have proposed [31] a model for the representation of the component lifecycle and of its dependencies/conflicts with the other components. Based on such model, we have defined a sound and complete algorithm that efficiently computes a deployment plan (i.e. a sequence of low-level component deployment actions) capable of reaching a final configuration including at least some predefined basic components [48] and we have realized a prototypical implementation of such algorithm which was proved to be effective on case-studies of realistic size (i.e. hundreds of components) [41].

6.4. Resource Control

Participants: Michele Alberti, Alberto Cappai, Ugo Dal Lago, Marco Gaboardi, Simone Martini, Paolo Parisen Toldin, Giulio Pellitta, Davide Sangiorgi, Marco Solieri, Valeria Vignudelli.

6.4.1. Expressive type systems for complexity analysis

Along 2013, our work on expressive methodologies for complexity analysis of higher-order languages has proceeded. In particular, we have focused our attention on extending linear dependent types to languages with control operators in the style of `callcc` [27]. This has taken the form of a generalization of bounded linear logic towards Laurent's polarized linear logic, which is then turned into a type system for the lambda-mu-calculus (in which the aforementioned control operator can indeed be implemented). In the introduced type system, all typable terms can be reduced in polynomial time. We also worked on the linear dependent type inference and on its implementation (though the work has not yet been transferred onto the *Lideal* tool implementing type inference for dependently linear type systems, see http://lideal.cs.unibo.it/); more specifically, we showed that type inference can in this context be reduced to a form of constraint amenable to be solved by SMT solvers [28]. Finally, a call-by-value version of $\mathsf{d}\ell\mathsf{PCF}$ has been defined and proved sound but also relatively complete as a tool for complexity analysis of programs [16].

6.4.2. Complexity analysis and process algebras

Complexity analysis methodologies drawn from linear logic have been adapted to higher-order process algebras, obtaining linear versions of the higher-order π-calculus in which reduction sequences are guaranteed to have a length bounded by a polynomial [14]. This is done by following the exponential discipline Lafont's Soft Linear logic suggests.

6.4.3. Characterizing probabilistic complexity classes

We have also been looking [10] (papers extracted from the thesis should appear soon) at probabilistic computation and at whether probabilistic complexity classes like **BPP**, **ZPP** and **PP** can be characterized by logics and λ-calculi. We encountered some problems in doing the above for **BPP** and **ZPP**, which are semantic classes and which, as a consequence, cannot be easily enumerated (and captured by ICC systems). On the other hand, probabilistic classes like **PP** can indeed be characterized by λ-calculi, as shown by our recent work on RSLR, a system derived from Hofmann's SLR that captures the (deterministic) polytime computable functions.

6.4.4. Ensuring differential privacy

Differential privacy offers a strong guaranteed bound on loss of private information due to release of query results, even under worst-case assumptions. One of the challenges in proving queries differentially private is to prove an upper bound on the query's sensitivity, i.e., the maximum change in the query's output that can result from changing the data of a single individual. Reed and Pierce have recently proposed a type analysis using numerical annotations in types to describe bounds on the sensitivity of the queries. A first delicate aspect of this approach is that in order to verify if a program is typable or not one needs to come up with numerical annotations and verify their consistency. Finding a "small" annotation is crucial, since the privacy depends on it. For this reason we designed a sensitivity inference tool [26] that combined with the Z3 SMT solver is able to verify and infer a minimal sensitivity bound in an automatic and efficient way. Another delicate aspect of this approach is the *expressivity* of the type analysis. Reed and Pierce's type system offers only a very limited form of numerical annotations. These numerical annotations are not enough to provide a bound for programs whose sensitivity depends on data available only at runtime. To recover this problem we introduced *Dfuzz* [32], a language combining linear types and lightweight dependent types.

6.5. Verification of extensional properties

Participants: Ornela Dardha, Elena Giachino, Michael Lienhardt, Cosimo Laneve, Fabrizio Montesi.

Extensional refers to properties that have to do with behavioral descriptions of a system (i.e., how a system looks like from the outside). Examples of such properties include classical functional correctness and deadlock freedom. Most work carried out this year has to do with type systems for concurrent objects and components ensuring safe and reliable interactions, and on deadlock analysis for systems of concurrent objects or within process sessions.

6.5.1. Type systems for objects and components

In previous work, we had developed an integration of session types, for specifying and validating structured communication sequences (sessions) into a class-based core object language for building network applications. We have defined [12] a constraint-based type system that reconstructs the appropriate session types of session declarations instead of assuming that session types are explicitly given by the programmer, and used static analysis via types to ensure that, once a session has started, computation cannot get stuck on a communication deadlock.

In previous papers, we had proposed a component layer for object-oriented language ABS (studied in the EU project Hats), that allows one to perform updates on objects by means of communication ports and their rebinding. We have now [29] introduced a type system for this component model that statically enforces that no object will attempt illegal rebinding.

6.5.2. Deadlock analysis

Deadlock represents an insidious and recurring threat when systems also exhibit a high degree of resource and data sharing. We address deadlock analysis of two such systems: (1) concurrent object-oriented languages; (2) protocol specifications.

For (1), we have developed a framework for statically detecting deadlocks in a concurrent object-oriented language with asynchronous method calls and cooperative scheduling of method activations. Since this language features recursion and dynamic resource creation, deadlock detection is extremely complex and state-of-the-art solutions either give imprecise answers or do not scale. In order to augment precision and scalability we propose a modular framework that allows several techniques to be combined. The basic component of the framework is a front-end inference algorithm that extracts abstract behavioral descriptions of methods, called contracts, which retain resource dependency information [33]. This component is integrated with a number of possible different back-ends that analyze contracts and derive deadlock information. As a proof-of-concept, we discuss two such back-ends: (i) an evaluator that computes a fixpoint semantics [33] and (ii) an evaluator using abstract model checking [34].

For (2), in [24], we develop a typing discipline that verifies choreographies against protocol specifications, based on multiparty sessions. Exploiting the nature of global descriptions, our type system defines a new class of deadlock-free concurrent systems (deadlock-freedom-by-design), provides type inference, and supports session mobility. We give a notion of Endpoint Projection (EPP) which generates correct entity code (as pi-calculus terms) from a choreography. Finally, we evaluate our approach by providing a prototype implementation for a concrete programming language and by applying it to some examples from multicore and service-oriented programming.

Finally, en passant we remind [23], that studies deadlock analysis of concurrent object-oriented languages via encoding into Petri nets, which had already been discussed in last year's report.

6.6. Expressiveness of computational models

Participants: Roberto Amadini, Ornela Dardha, Maurizio Gabbrielli, Daniel Hirschkoff, Jean-Marie Madiot, Jacopo Mauro, Davide Sangiorgi, Gianluigi Zavattaro.

Expressiveness refers to the study of the descriptive power of computational models.

The fusion calculi are a simplification of the π-calculus in which input and output are symmetric and restriction is the only binder. We show [35] a major difference between these calculi and the π-calculus from the point of view of types, proving some impossibility results for subtyping in fusion calculi. We propose a modification of fusion calculi in which the name equivalences produced by fusions are replaced by name preorders, so to be able to import subtype systems, and related results, from the π-calculus. We have studied the consequences of the modification on behavioural equivalence and expressiveness.

In Focus we use notions of constraint to define in a succinct way models of computation and current constraint solving technologies to solve problems modeled using constraints. For this reason we have studied the expressive power of various computational models involving constraints and their practical impact in terms of solving/execution performances. In [18] we have investigated how the notion of constraint augments the expressive power of a concurrent language if priorities are introduced. The chosen language is Constraint Handling Rules, a committed-choice declarative language originally designed for writing constraint solvers and that is nowadays a general purpose language. The result has been otbained by first formalising the meaning of language encodings and language embedding, widely used in concurrency theory. Different ways to model and define disaster scenarios are analyzed and compared in [11], where we study a model expressive enough to define a disaster scenario that, at the same time, can be used to find plans to save the victims of a disaster using modern constraint solving technology. Similarly, different computation models are considered in [22] where we study how machine learning techniques can be used to boost the performances of constraint solvers. A technique dubbed "portfolio approach" is used to combine the different performances of constraint solvers to obtain a globally better solver using, as a starting point, a simple low-level constraint language.

In [30] we propose an integration of structural sub-typing with boolean connectives and semantic sub-typing to define a Java-like programming language that exploits the benefits of both techniques. The resulting language has a more expressive set of types that comes from the use of boolean constructs, negation types, and the integration of structural and nominal sub-typing in an object-oriented setting. By implementing traditional Java-language constructs we show that the proposed language is also expressive enough w.r.t. the Java language.

FORMES Team

6. New Results

6.1. Type and rewriting theory

Participants: Frédéric Blanqui, Jean-Pierre Jouannaud, Jianqi Li, Qian Wang.

Qian Wang and Bruno Barras have proved the strong normalization property of CoqMTU in presence of strong elimination, a major step towards the full certification of CoqMTU [16].

Jouannaud and Li have developped a new framework, Normal Abstract Rewriting Systems (NARS), that captures all known Church-Rosser results in presence of a termination assumption allowing to reduce equality of terms to a simpler equality on their normal forms. This result applies to the paticular case of higher-order rewriting for which it solved long-standing open problems [10].

Jouannaud and Liu have continued their investigation of Church-Rosser properties of non-terminating rewrite systems [10], showing recently first, that many results found in the litterature could be captured, and generalized, by using van Oostrom's decreasing diagram technique (accepted at Symposium on Algebraic Specifications, Kanazawa, Japan, April 2014). The next step, which has been recently completed, is a powerful result generalizing Knuth and Bendix confluence test to non terminating rewrite system (submitted).

Frédéric Blanqui, Jean-Pierre Jouannaud and Albert Rubio (Technical University of Catalonia) have developed a method aiming at carrying out termination proof for higher-order calculi. CPO appears to be the ultimate improvement of the higher-order recursive path ordering (HORPO) [25] in the sense that this definition captures the essence of computability arguments *à la* Tait and Girard, therefore explaining the name of the improved ordering. It has been shown that CPO allows to consider higher-order rewrite rules in a simple type discipline with inductive types, that most of the guards present in the recursive calls of its core definition cannot be relaxed in any natural way without losing well-foundedness, and that the precedence on function symbols cannot be made more liberal anymore. This result is submitted to journal, and has been concurrently generalized to higher-order calculi with dependent types by Jouannaud and Li (submitted).

Frédéric Blanqui worked on the formalization in the Coq proof assistant of various definitions of the notion of α-equivalence on pure λ-terms. In particular, he formalized and formally proved equivalent the definitions of Church (1932), Curry and Feys (1958), Krivine (1993), and Gabbay and Pitts (1999). This work is freely available from the CoLoR library released on December 13th.

Frédéric Blanqui worked with John Steinberger (Tsinghua University) on the formal verification in Coq of proofs of theorems on coset arrays and non-negative integer linear combinations.

6.2. Automated theorem proving

Participant: Kim-Quyen Ly.

Kim-Quyen Ly extended her formally-proved (in Coq) automated termination-certificate (for first-order rewrite systems) verifier Rainbow for dealing with certificates using arguments filtering [22] and other termination techniques.

6.3. Simulation

Participants: Vania Joloboff, Antoine Rouquette, Shenpeng Wang.

There exists very fast Loosely Timed simulators such as **SimSoC** that can run the application software to validate its functionality and possibly test real time software using timers. But such simulators do not provide good enough timings to evaluate the software performance. The idea of “Approximately Timed” simulation is to provide a fast simulation that can be used by software developers, and yet provide performance estimate. The goal of approximately timed simulation is to provide estimates that are within a small margin error from the real hardware performance, but at a simulation speed that is an order of magnitude faster than a cycle accurate one.

Modern processors have complex architectures. They can execute a certain number of instructions per clock cycle. There are however several cases where the instruction flow is disrupted, introducing delays in the computation. In order to make an Approximately Timed simulator, our idea is to simulate enough of the processes causing the delays, not simulating the exact hardware processes of the caches and pipe line and I/Os, but using a model with wich the delays can be computed with a reasonable approximation while maintaining fast simulation. Delays may also be related to bus arbitration and interconnect access. These delays are beyond the scope of our work, but can be captured by TLM (timed) transactions. In our work, we are considering only the processor model and we rely upon TLM interface to the interconnect for peripheral access to provides us with timing delays.

We have started to investigate a new approach to provide a fast Approximately Timed ISS, that does not simulate fully the hardware, yet provides good precision estimates, and does not use stastistical methods. Our approach consists in developing a higher abstraction model of the processor (than the CA models) that still executes instructions using fast SystemC/TLM code, but in parallel maintains some architecture state to measure the delays introduces by cache misses and pipe line stalls, although the pipe line is not really simulated.

6.4. Certification of a Simulator

Participants: Vania Joloboff, Jean-François Monin, Xiaomu Shi.

We have developed a correctness proof of a part of the hardware simulator **SimSoC**. This is not only an attempt to certify a simulator, but also a new experiment on the certification of non-trivial programs written in C. We have provided a formalized representation of the ARM instruction set and addressing modes in Coq. We also constructed a Coq representation of the ARM simulator in C, using the abstract syntax defined in **CompCert**.

From these two Coq representations, we have developed Coq proofs to prove the correctness of the C code, using the operational semantics of C provided by **CompCert**.

During this work, we have also improved the technology available in Coq for performing *inversions*, a kind of proof steps which heavily occurs in this context.

All of this work has been described in Xiaomu SHI PhD thesis dissertation, presented at University of Grenoble in July 2013, and at ITP 2013 conference[15].

FUN Project-Team

5. New Results

5.1. Routing in FUN

Participants: Thierry Delot, Tony Ducrocq, Nicolas Gouvy, Nathalie Mitton, Enrico Natalizio, David Simplot-Ryl, Tahiry Razafindralambo, Dimitrios Zormpas.

Wireless sensor and actuator/robot networks need some routing mechanisms to ensure that data travel the network to the sink with some guarantees. The FUN research group has investigated different geographic routing paradigms. Georouting assumes that every node is aware of its location, the one of its neighbors and of the destination(s).

In this context, we first propose the first k-anycating georouting protocol, ie in which a node wishes to send a message to k sinks in the network [13]. Then, we tried to relax some of the assumptions. For instance in [12], we introduce HECTOR which is the first position based routing protocol which relies on virtual positions, is energy-aware and guarantees the data delivery.

In [46], [21], we assume that only a part of nodes is aware of its position and proposes a hybrid approach between position-based greedy approach and traditional on-demand routing. Indeed, geographic routing protocols show good properties for WSNs. They are stateless, local and scalable. However they require that each node of the network is aware of its own position. While it may be possible to equip each node with GPS receiver, even if it is costly, there are some issues and receiving a usable GPS signal may be difficult in some situations. For these reasons, we propose a geographic routing algorithm, called HGA, able to take advantages of position informations of nodes when available but also able to continue the routing in a more traditional way if position information is not available. We show with simulations that our algorithm offers an alternative solution to classical routing algorithm (non-geographic) and offers better performances for network with a density above 25 and more than 5% of nodes are aware of their position. [46] analyses the impact of nodes topology on network performances. We show that different topologies can lead to a difference of up to 25% on delivery ratio and average route length and more than 100% on overall cost of transmissions.

In [24], [25], [26], [3], we consider that nodes are able to move by themselves and we try to take advantage of this feature to improve the network performance. In sensor networks, there is often more than one sensor which reports an event to the sink in WSN. In existing solutions, this leads to oscillation of nodes which belong to different routes and their premature death. Experiments show that the need of a routing path merge solution is high. As a response, [24], [25] introduce the first routing protocol which locates and uses paths crossing to adapt the topology to the network traffic in a fully localized way while still optimizing energy efficiency. Furthermore the protocol makes the intersection to move away from the destination, getting closer to the sources, allowing higher data aggregation and energy saving. Our approach outperforms existing solutions and extends network lifetime up to 37%.

Using nodes location, position-based routing protocols generally apply a greedy routing that makes a sensor forward data to route to one of its neighbors in the forwarding direction of the destination. If this greedy step fails, the routing protocol triggers a recovery mechanism. Such recovery mechanisms are mainly based on graph planarization and face traversal or on a tree construction. Nevertheless real-world network planarization is very difficult due to the dynamic nature of wireless links and trees are not so robust in such dynamic environments. Recovery steps generally provoke huge energy overhead with possibly long inefficient paths. In [26], we propose to take advantage of the introduction of controlled mobility to reduce the triggering of a recovery process. We propose Greedy Routing Recovery (GRR) routing protocol. GRR enhances greedy routing energy efficiency as it adapts network topology to the network activity. Furthermore GRR uses controlled mobility to relocate nodes in order to restore greedy and reduce energy consuming recovery step triggering. Simulations demonstrate that GRR successfully bypasses topology holes in more than 72% of network topologies avoid- ing calling to expensive recovery steps and reducing energy consumption while preserving network connectivity.

[31] relaxes the assumption that nodes are aware of their neighbors and considers that dynamic energy sources could be available. It introduces MEGAN (Mobility assisted Energy efficient Georouting in energy harvesting Actuator and sensor Networks), a beacon-less protocol that uses controlled mobility, and takes account of the energy consumption and the energy harvesting to select next hop. MEGAN aims at prolonging the overall network lifetime rather than reducing the energy consumption over a single path. When node s needs to send a message to the sink d, it first computes the ideal position of the forwarder node based on available and needed energy, and then broadcasts this data. Every node within the transmission range of s in the forward direction toward d will start a backoff timer. The backoff time is based on its available energy and on its distance from the ideal position. The first node whose backoff timer goes o is the forwarder node. This node informs its neighbor- hood and then moves toward the ideal position. If, on its route, it finds a good spot for energy harvesting, it will actually stop its movement and forward the original message by using MEGAN, which will run on all the intermediate nodes until the destination is reached. Simulations show that MEGAN reduces energy consumption up to 50% compared to algorithms where mobility and harvesting capabilities are not exploited.

Additionally, according to a wide range of studies, (Informatics Technologies) IT should become a key facilitator in establishing primary education, reducing mortality and supporting commercial initiatives in Least Developed Countries (LDCs). The main barrier to the development of IT services in these regions is not only the lack of communication facilities, but also the lack of consistent information systems, security procedures, economic and legal support, as well as political commitment. In [18], we propose the vision of an infrastructureless data platform well suited for the development of innovative IT services in LDCs. We propose a participatory approach, where each individual implements a small subset of a complete information system thanks to highly secure, portable and low-cost personal devices as well as opportunistic networking, without the need of any form of infrastructure. [18] reviews the technical challenges that are specific to this approach.

5.2. Self-organization

Participants: Tony Ducrocq, Nathalie Mitton, David Simplot-Ryl, Isabelle Simplot-Ryl.

Self-organization encompasses several mechanisms. This year, the FUN research group contributes to some of them such as neighbor discovery, localization, clustering and topology control in FUN.

5.2.1. Neighbor discovery

HELLO protocol or neighborhood discovery is essential in wireless ad hoc networks. It makes the rules for nodes to claim their existence/aliveness. In the presence of node mobility, no x optimal HELLO frequency and optimal transmission range exist to maintain accurate neighborhood tables while reducing the energy consumption and bandwidth occupation. Thus a Turnover based Frequency and transmission Power Adaptation algorithm (TFPA) is presented in [27]. The method enables nodes in mobile networks to dynamically adjust both their HELLO frequency and transmission range depending on the relative speed. In TFPA, each node monitors its neighborhood table to count new neighbors and calculate the turnover ratio. The relationship between relative speed and turnover ratio is formulated and optimal transmission range is derived according to battery consumption model to minimize the overall transmission energy. By taking advantage of the theoretical analysis, the HELLO frequency is adapted dynamically in conjunction with the transmission range to maintain accurate neighborhood table and to allow important energy savings. The algorithm is simulated and compared to other state-of-the-art algorithms. The experimental results demonstrate that the TFPA algorithm obtains high neighborhood accuracy with low HELLO frequency (at least 11% average reduction) and with the lowest energy consumption. Besides, the TFPA algorithm does not require any additional GPS-like device to estimate the relative speed for each node, hence the hardware cost is reduced.

5.2.2. Topology control

Topology control is a tool for self-organizing wireless networks locally. It allows a node to consider only a subset of links/neighbors in order to later reduce computing and memory complexity. Topology control in wireless sensor networks is an important issue for scalability and energy efficiency. It is often based on graph reduction performed through the use of Gabriel Graph or Relative Neighborhood Graph. This graph reduction is usually based on geometric values.

In [11], we propose a radically new family of geometric graphs, i.e., Hypocomb, Reduced Hypocomb and Local Hypocomb for topology control. The first two are extracted from a complete graph; the last is extracted from a Unit Disk Graph (UDG). We analytically study their properties including connectivity, planarity and degree bound. All these graphs are connected (provided the original graph is connected) planar. Hypocomb has unbounded degree while Reduced Hypocomb and Local Hypocomb have maximum degree 6 and 8, respectively. To the best of our knowledge, Local Hypocomb is the first strictly-localized, degree-bounded planar graph computed using merely 1-hop neighbor position information. We present a construction algorithm for these graphs and analyze its time complexity. Hypocomb family graphs are promising for wireless ad hoc networking. We report our numerical results on their average degree and their impact on FACE [49] routing. We discuss their potential applications and some open problems.

5.2.3. Clustering

Clustering in wireless sensor networks is an efficient way to structure and organize the network. It aims to identify a subset of nodes within the network and bind it a leader (i.e. cluster-head). This latter becomes in charge of specific additional tasks like gathering data from all nodes in its cluster and sending them by using a longer range communication to a sink.

As a consequence, a cluster-head exhausts its battery more quickly than regular nodes. In [8], [22], [1], we present BLAC, a novel Battery-Level Aware Clustering family of schemes. BLAC considers the battery-level combined with another metric to elect the cluster-head. It comes in four variants. The cluster-head role is taken alternately by each node to balance energy consumption. Due to the local nature of the algorithms, keeping the network stable is easier. BLAC aims to maximize the time with all nodes alive to satisfy application requirements. Simulation results show that BLAC improves the full network lifetime 3-time more than traditional clustering schemes by balancing energy consumption over nodes and still delivering high data percentage.

On another approach, [34] considers the Slepian-Wolf coding based data aggregation problem and the corresponding dependable clustering problem in WSN. A dependable Slepian-Wolf coding based clustering (DSWC) algorithm is proposed to provide dependable clustering against cluster-head failures. The proposed D-SWC algorithm attempts to elect a primary cluster head and a backup cluster head for each cluster member during clustering so that once a failure occurs to the primary cluster head the cluster members within the failed cluster can promptly switchover to the backup cluster head and thus recover the connectivity of the failed cluster to the data sink without waiting for the next-round clustering to be performed. Simulation results show that the DSWC algorithm can effectively increase the amount of data transmitted to the data sink as compared with an existing nondependable clustering algorithm for Slepian-Wolf coding based data aggregation in WSNs.

5.3. Controlled mobility

Participants: Milan Erdelj, Valeria Loscri, Kalypso Magklara, Karen Miranda, Enrico Natalizio, Jean Razafimandimby Anjalalaina, Tahiry Razafindralambo, David Simplot-Ryl, Dimitrios Zormpas.

Controlled mobility [5] is a new paradigm that leads to a set of great new challenges.

5.3.1. Target coverage

One of the main operations in wireless sensor networks is the surveillance of a set of events (targets) that occur in the field. In practice, a node monitors an event accurately when it is located closer to it, while the opposite happens when the node is moving away from the target. This detection accuracy can be represented by a probabilistic distribution. Since the network nodes are usually randomly deployed, some of the events are monitored by a few nodes and others by many nodes. In applications where there is a need of a full coverage and of a minimum allowed detection accuracy, a single node may not be able to sufficiently cover an event by itself. In this case, two or more nodes are needed to collaborate and to cover a single target. Moreover, all the nodes must be connected with a base station that collects the monitoring data.

In [15], we describe the problem of the minimum sampling quality, where an event must be sufficiently detected by the maximum possible amount of time. Since the probability of detecting a single target using randomly deployed static nodes is quite low, we present a localized algorithm based on mobile nodes. Our algorithm sacrifices a part of the energy of the nodes by moving them to a new location in order to satisfy the desired detection accuracy. It divides the monitoring process in rounds to extend the network lifetime, while it ensures connectivity with the base station. Furthermore, since the network lifetime is strongly related to the number of rounds, we propose two redeployment schemes that enhance the performance of our approach by balancing the number of sensors between densely covered areas and areas that are poorly covered. Finally, our evaluation results show an over 10 times improvement on the network lifetime compared to the case where the sensors are static. Our approaches, also, outperform a virtual forces algorithm when connectivity with the base station is required. The redeployment schemes present a good balance between network lifetime and convergence time.

[47], [28] assume that these targets to cover are dynamic. We assume that no knowledge about either event position or duration is given a priori. Nonetheless, the events need to be monitored and covered thanks to mobile wireless sensors. Thus, mobile sensors have to discover the events and move towards a new Zone of Interest (ZoI) when the previous monitored event is over. An efficient, distributed and localized solution of this problem would be immediately exploitable by several applications domains, such as environmental, civil, etc. We propose two novel approaches to deal with dynamic event coverage. The first one is a modified version of the PSO, where particles (mobile sensors, nodes or devices in the following) update their velocity by using only local information coming from their neighbors. In practice, the velocity update is performed by considering neighbors' sensed events. Our distributed version of PSO is integrated with a distributed version of the Virtual Force Algorithm (VFA). Virtual Force approach has the ability to "position" nodes with no overlap, by using attractive and repulsive forces based on the distance between nodes. The other proposed algorithm is a distributed implementation of the VFA by itself. Both techniques are able to reach high levels of coverage and show a satisfying reactivity when the ZoI changes. This output parameter is measured as the capability for the sensors to "follow" a sequence of events happening in different ZoIs. The effectiveness of our techniques is shown through a series of simulations and comparisons with the classical centralized VFA.

On another approach consists in using flying drone to cover this set of targets. [39] focuses on the energy efficiency problem where camera equipped flying drones are able to detect and follow mobile events that happen on the ground. We give a mathematical formulation of the problem of minimizing the total energy consumption of a fleet of drones when coverage of all events is required. Due to the extremely high complexity of the binary optimization problem, the optimum solution cannot be obtained even for small instances. On the contrary, we present LAS, a localized solution for the aforementioned problem which takes into account the ability of the drones to fly at lower altitudes in order to conserve energy. We simulate LAS and we compare its performance to a centralized algorithm and to an approach that uses static drones to cover all the terrain. Our findings show that LAS performs similar to the centralized algorithm, while it outperforms the static approach by up to 150% in terms of consumed energy. Finally, the simulation results show that LAS is very sustainable in presence of communication errors.

5.3.2. Multiple Point of Interest coverage

The coverage of Points of Interest (PoI) is a classical requirement in mobile wireless sensor applications. Optimizing the sensors self-deployment over a PoI while maintaining the connectivity between the sensors and the base station is thus a fundamental issue.

The problems of multiple PoI discovery, coverage and data report are still solved separately and there are no works that combine the aforementioned problems into a single deployment scheme. In [9], [2], we present a novel approach for mobile sensor deployment, where we combine multiple PoI discovery and coverage with network connectivity preservation in order to capture the dynamics of the monitored area. Furthermore, we derive analytical expressions for circular movement parameters and examine the performance of our approach through extensive simulation campaigns.

[10] addresses the problem of autonomous deployment of mobile sensors that need to cover a predefined PoI with a connectivity constraint. In our algorithm, each sensor moves toward a PoI but has also to maintain the connectivity with a subset of its neighboring sensors that are part of the Relative Neighborhood Graph (RNG). The Relative Neighborhood Graph reduction is chosen so that global connectivity can be provided locally. Our deployment scheme minimizes the number of sensors used for connectivity thus increasing the number of monitoring sensors. Analytical results, simulation results and practical implementation are provided to show the efficiency of our algorithm.

5.3.3. Robot cooperation

The concept of autonomous mobile agents gets a lot of attention in the domain of WSN or wireless sensor and actuator networks (WSAN). Multiple robots that coordinate or cooperate with other sensors, robots or human operator, allow the WSN/WSAN to perform tasks that are far beyond the scope of single robot unit. In[23], we describe the robot middleware architecture that allows networked multi-robot control and data acquisition in the context of wireless sensor networks. Furthermore, we present three examples of robot network deployment and illustrate the proposed architecture usability: the robotic network deployment with the goal of covering the Point of Interest, adaptable multi-hop video transmission scenario, and the case of obtaining the energy consumption during the deployment.

5.3.4. Substitution networks

A substitution network [4] is a temporary network that will be deployed to support a base network in trouble and help it to provide the best service.

WSN are widely deployed nowadays on a large variety of applications. The major goal of a WSN is to collect information about a set of phenomena. Such process is non trivial since batteries' life is limited and thus wireless transmissions as well as computing operations must be minimized. A common task in WSNs is to estimate the sensed data and to spread the estimated samples over the network. Thus, time series estimation mechanisms are vital on this type of processes so as to reduce data transmission. In [30], we assume a single-hop clustering mechanism in which sensor nodes are grouped into clusters and communicate with a sink through a single hop. We propose a couple of autoregressive mechanisms to predict local sensed samples in order to reduce wireless data communication. We compare our proposal with a model called EEE that has been previously proposed in the literature. We prove the efficiency of our algorithms with real samples publicly available and show that they outperform the EEE mechanism.

In [32], we propose an algorithm to efficiently (re)-deploy the wireless mobile routers of a substitution network by considering the energy consumption, a fast deployment scheme and a mix of the network metric. We consider a scenario where we have two routers in a fixed network and where connectivity must be restored between those two routers with a wireless mobile router. The main objective of the wireless mobile router is to increase the communication performance such as the throughput by acting as relay node between the two routers of the fixed network. We present a fast, adaptive and localized approach which takes into account different network metrics such as Received Signal Strength (RSS), Round-Trip Time (RTT) and the Transmission Rate, between the wireless mobile router and the two routers of the fixed network. Our method ameliorates the performance of our previous approach from the literature by shortening the deployment time, increasing the throughput, and consuming less energy in some specific cases.

5.4. Security

Participants: Nathalie Mitton, Enrico Natalizio.

[19] deals with the energy efficient issue of cryptographic mechanisms used for secure communication between devices in wireless sensor networks. Since these devices are mainly targeted for low power consumption appliances, there is an effort for optimization of any aspects needed for regular sensor operation. On a basis of utilization of hardware cryptographic accelerators integrated in microcontrollers, this article provides the comparison between software and hardware solutions. Proposed work examines the problems and solutions for implementation of security algorithms for WSN devices. Because the speed of hardware accelerator should

be much higher than the software implementation, there are examination tests of energy consumption and validation of performance of this feature. Main contribution of the article is real testbed evaluation of the time latency and energy requirements needed for securing the communication. In addition, global evaluation for all important network communication parameters like throughput, delay and delivery ratio are also provided.

The Internet of Things (IoT) will enable objects to become active participants of everyday activities. Introducing objects into the control processes of complex systems makes IoT security very difficult to address. Indeed, the Internet of Things is a complex paradigm in which people interact with the technological ecosystem based on smart objects through complex processes. The interactions of these four IoT components, person, intelligent object, technological ecosystem, and process highlight a systemic and cognitive dimension within security of the IoT. The interaction of people with the technological ecosystem requires the protection of their privacy. Similarly, their interaction with control processes requires the guarantee of their safety. Processes must ensure their reliability and realize the objectives for which they are designed. We believe that the move towards a greater autonomy for objects will bring the security of technologies and processes and the privacy of individuals into sharper focus. Furthermore, in parallel with the increasing autonomy of objects to perceive and act on the environment, IoT security should move towards a greater autonomy in perceiving threats and reacting to attacks, based on a cognitive and systemic approach. In [33], we will analyze the role of each of the mentioned actors in IoT security and their relationships, in order to highlight the research challenges and present our approach to these issues based on a holistic vision of IoT security.

5.5. RFID

Participants: Ibrahim Amadou, Nathalie Mitton.

Mitigating reader-to-reader collisions is one of the principal challenges in a large-scale dynamic RFID system with a number of readers deployed in order to maximize the system performance (i.e., throughput, fairness and latency). In prior works, contention-based and activity scheduling medium access control (MAC) protocols are commonly used approaches to reduce such problems. Existing protocols typically perform worse in a large-scale RFID dynamic system and require more additional components or are based on unrealistic assumptions. So far, many research efforts have been made to improve the performance or the reliability of Carrier Sense Multiple Access (CSMA) techniques for Mobile Ad-Hoc Networks (MANETs) by using an adaptive Backoff schemes. In [17], we look at these well known solutions that proved their efficiency in high congestion wireless networks. We evaluate the performance and characterize these solutions when they are used to reserve the wireless channel through broadcasting message for reader-to-tag communication. Based on the application requirements, we study their capacity to mitigate collisions, the channel access latency, the average number of successful requests sent per reader and the fairness index in the context of RFID networks.

5.6. Data collection and aggregation

Participant: Nathalie Mitton.

Named Data Networking (NDN) is a new promising paradigm for content retrieval and distribution in the future Internet. NDN communication is driven by data consumers that broadcast Interest packets to require named contents. The requests are forwarded towards the source(s) by directly using content names (instead of IP addresses), while in-network caching is used to improve delivery performance. NDN shows many similarities with data-centric models defined for wireless sensor networks (WSNs), e.g., directed diffusion. In addition, NDN defines a new complete communication framework with innovative naming and security schemes and novel routing and transport strategies. This clearly opens new perspectives in the design and development of sensor networks, which can benefit of the NDN framework to better support different kinds of applications and services. In [16], we explore the potentialities of NDN applied to WSNs and propose enhanced delivery strategies inspired by the directed diffusion scheme to be deployed in the NDN framework. Performance of a plain NDN scheme and of our enhanced solution is evaluated through the ndnSIM simulator. Achieved results confirm the viability of a NDN-like approach over WSNs and the better efficiency and effectiveness of the proposed solution compared to a plain NDN.

[38] considers the Slepian-Wolf coding based energy-minimization rate allocation problem in a WSN and propose a distributed rate allocation algorithm to solve the problem. The proposed distributed algorithm is based on an existing centralized rate allocation algorithm which has a high computational complexity. To reduce the computational complexity of the centralized algorithm and make the rate allocation performable in a distributed manner, we make necessary modifications to the centralized algorithm by reducing the number of sets in calculating the average energy consumption cost and limiting the number of conditional nodes that a set can use. Simulation results show that the proposed distributed algorithm can significantly reduce the computational time when compared with the existing centralized algorithm at the cost of the overall energy consumption for data transmission and the total amount of data transmitted in the network.

5.7. VANET

Participant: Nathalie Mitton.

Routing is a critical issue in vehicular ad hoc networks (VANETs). This paper considers the routing issue in both vehicle to vehicle (V2V) and vehicle to infrastructure (V2I) communications in VANETs, and proposes a Moving dirEction and DestinAtion Location based routing (MEDAL) algorithm for supporting V2V and V2I communications. MEDAL [36] takes advantage of both the moving directions of vehicles and the destination location to select a neighbor vehicle as the next hop for forwarding data. Unlike most existing routing algorithms, it only uses a HELLO message to obtain or update routing information without using other control messages, which largely reduces the number of control messages used in routing. Simulation results show that MEDAL can significantly improve the packet delivery ratio of the network as compared with the well-known Ad hoc On-demand Distance Vector Routing (AODV) algorithm.

5.8. Industrial Applications

Participants: Milan Erdelj, Nathalie Mitton, Enrico Natalizio.

The collaborative nature of industrial wireless sensor networks (IWSNs) brings several advantages over traditional wired industrial monitoring and control systems, including self-organization, rapid deployment, flexibility, and inherent intelligent processing. In this regard, IWSNs play a vital role in creating more reliable, efficient, and productive industrial systems, thus improving companies' competitiveness in the marketplace. Industrial Wireless Sensor Networks: Applications, Protocols, and Standards [42] examines the current state of the art in industrial wireless sensor networks and outlines future directions for research.

GALAAD Project-Team

6. New Results

6.1. Algebraic representations for geometric modeling

6.1.1. Fitting ideals and multiple-points of surface parameterizations

Participant: Laurent Busé.

Parameterized algebraic surfaces are ubiquitous in geometric modeling and the determination of their singular loci is an important problem. Given a birational parameterization ϕ from $\mathbb{P}^2$ to $\mathbb{P}^3$ of a rational algebraic surface $\mathcal{S}$, the purpose of this work is to investigate the sets of points on $\mathcal{S}$ whose preimage consists in k or more points, counting multiplicity. In collaboration with Nicolas Botbol (University of Buenos Aires) and Marc Chardin (UMPC), we prove that they can be described in terms of Fitting ideals of some graded parts of the symmetric algebra associated to the parameterization ϕ. More precisely, we show that the drop of rank of a certain elimination matrix $M(\phi)$ at a given point $P \in \mathbb{P}^3$ is in relation with the fiber of the graph of ϕ over P. Thus, the Fitting ideals attached to $M(\phi)$ provide a filtration of the surface which is in correspondence with the degree and the dimension of the fibers of the graph of the parameterization ϕ. This property is linked with the double-point formulas that have been extensively studied in the field of intersection theory for finite maps.

This work has been accepted for presentation at the international conference MEGA 2013 and is submitted for publication [33].

6.1.2. Discriminant of a homogeneous and symmetric polynomial

Participant: Laurent Busé.

Polynomial algebra offers a standard approach to handle several problems in geometric modeling. A key tool is the discriminant of a well-constrained system of polynomial equations, which expresses the existence of a multiple root. In this work the factorization of a single homogeneous and symmetric polynomial is investigated. Indeed, in this setting the discriminant possesses a lot of symmetries and all of these symmetries produce an independent factor of the global discriminant. The two difficult points here are to prove that each of these factors are irreducible over a nice base ring and to determine its multiplicity in the factorization of the discriminant. This work, in collaboration with Anna Karasoulou (University of Athens) is still under progress.

6.1.3. On the cactus rank of cubic forms

In [14], we prove that the smallest degree of an apolar 0-dimensional scheme of a general cubic form in $n+1$ variables is at most $2n+2$, when $n \geq 8$, and therefore smaller than the rank of the form. For the general reducible cubic form the smallest degree of an apolar subscheme is $n+2$, while the rank is at least $2n$.

This is a work done by Alessandra Bernardi when she was post-doctorate for DECONSTRUCT IEF project, in collaboration with Kristian Ranestad (University of Oslo).

6.1.4. Grassmann secants and linear systems of tensors

For any irreducible non-degenerate variety $X \subset \mathbb{P}^r$, we relate in [11] the dimension of the s-th secant varieties of the Segre embedding of $\mathbb{P}^k \times X$ to the dimension of the (k,s)-Grassmann secant variety $GS_X(k,s)$ of X. We also give a criterion for the s-identifiability of X.

This is a work done by Alessandra Bernardi when she was post-doctorate for DECONSTRUCT IEF project, in collaboration with Edoardo Ballico (University of Trento), Maria Virgina Catalisano (DIPTEM, Genova), Luca Chiantini (University of Sienna).

6.1.5. Optimal analysis-aware parameterization of computational domain in 3D isogeometric analysis

Participants: André Galligo, Bernard Mourrain.

In isogeometric analysis framework, computational domain is exactly described using the same representation as that employed in the CAD process. For a CAD object, we can construct various computational domain with same shape but with different parameterization. One basic requirement is that the resulting parameterization should have no self-intersections. In [27], a linear and easy-to-check sufficient condition for injectivity of trivariate B-spline parameterization is proposed. By an example of 3D thermal conduction problem, we show that different parameterization of computational domain has different impact on the simulation result and efficiency in isogeometric analysis. For problems with exact solutions, we propose a shape optimization method to obtain optimal parameterization of computational domain. The proposed injective condition is used to check the injectivity of initial trivariate B-spline parameterization constructed by discrete Coons volume method, which is the generalization of discrete Coons patch method. Several examples and comparisons are presented to show the effectiveness of the proposed method. Compared with the initial parameterization during refinement, the optimal parameterization can achieve the same accuracy but with less degrees of freedom.

This is a joint work with Régis Duvigneau (Inria, EPI OPALE) and Xu Gang (College of computer - Hangzhou Dianzi University, China).

6.1.6. Constructing analysis-suitable parameterization of computational domain from CAD boundary by variational harmonic method

Participants: André Galligo, Bernard Mourrain.

In isogeometric analysis, parameterization of computational domain has great effects as mesh generation in finite element analysis. In the paper [26], based on the concept of harmonic mapping from the computational domain to parametric domain, a variational harmonic approach is proposed to construct analysis-suitable parameterization of computational domain from CAD boundary for 2D and 3D isogeometric applications. Different from the previous elliptic mesh generation method in finite element analysis, the proposed method focuses on isogeometric version, and converts the elliptic PDE into a nonlinear optimization problem, in which a regular term is integrated into the optimization formulation to achieve more uniform and orthogonal iso-parametric structure near convex (concave) parts of the boundary. Several examples are presented to show the efficiency of the proposed method in 2D and 3D isogeometric analysis.

This is a joint work with Régis Duvigneau (Inria, EPI OPALE) and Xu Gang (College of computer - Hangzhou Dianzi University, China).

6.1.7. Spline Spaces over Quadrangle Meshes with Complex Topologies

Participants: Meng Wu, André Galligo, Bernard Mourrain.

We study a new type of spline functions defined over a rectangular mesh equipped with an equivalence relation, in such a way that physical spaces with a complex topology can be represented as an homomorphic image of such meshes. We provide general definitions, a dimension formula for a subclass of these spline spaces, an explicit construction of their bases and also a process for local refinement. These developments, motivated by plane curvilinear mesh constructions are illustrated on several parametrization problems. Our main target in these constructions is to approximate isobaric lines of magnetic fields encountered in MHD (Magnetohydrodynamics) simulation for Tokamaks. Their particularity is that one of the isobaric curve has a node singularity.

This work is done in collaboration with Boniface Nkonga (Inria, EPI CASTOR and University of Nice).

6.1.8. Lagrangian Curves in Affine Symplectic 4-space

Participant: Evelyne Hubert.

Lagrangian curves in 4-space entertain intriguing relationships with second order deformation of plane curves under the special affine group and null curves in a 3-dimensional Lorentzian space form. In [39] we provide a natural affine symplectic frame for Lagrangian curves. It allows us to classify Lagrangrian curves with constant symplectic curvatures, to construct a class of Lagrangian tori and determine Lagrangian geodesics.

This is joint work with Emilio Musso, Dipartimento di Scienze Matematiche, Politecnico de Turino (Italy).

6.2. Algebraic algorithms for geometric computing

6.2.1. Implicit matrix representations of rational Bézier curves and surfaces

Participant: Laurent Busé.

In this work, we introduce and study a new implicit representation of rational Bézier curves and surfaces in the 3-dimensional space. Given such a curve or surface, this representation consists of a matrix whose entries depend on the space variables and whose rank drops exactly on this curve or surface. Our approach can be seen as an extension of the moving lines implicitization method introduced by Sederberg, from non-singular matrices to the more general context of singular matrices. First, we describe the construction of these new implicit matrix representations and their main geometric properties, in particular their ability to solve efficiently the inversion problem. Second, we show that these implicitization matrices adapt geometric problems, such as intersection problems, to the powerful tools of numerical linear algebra, in particular to one of the most important: the singular value decomposition. So, from the singular values of a given implicit matrix representation, we introduce a real evaluation function. We show that the variation of this function is qualitatively comparable to the Euclidean distance function. As an interesting consequence, we obtain a new determinantal formula for implicitizing a rational space curve or surface over the field of real numbers. Then, we show that implicit matrix representations can be used with numerical computations, in particular there is no need for symbolic computations to use them. We give some rigorous results explaining the numerical stability that we have observed in our experiments. We end the paper with a short illustration on ray tracing of parameterized surfaces.

This work has been accepted for presentation and publication at the SIAM conference on Geometric and Physical Modeling 2013 (Denver, USA, Nov. 11-14) [15]. It has been awarded the best paper price, 1st place.

6.2.2. Superfast solution of Toeplitz systems based on syzygy reduction

Participant: Bernard Mourrain.

In [22], we present a new superfast algorithm for solving Toeplitz systems. This algorithm is based on a relation between the solution of such problems and syzygies of polynomials or moving lines. We show an explicit connection between the generators of a Toeplitz matrix and the generators of the corresponding module of syzygies. We show that this module is generated by two elements and the solution of a Toeplitz system $Tu = g$ can be reinterpreted as the remainder of a vector depending on g, by these two generators. We obtain these generators and this remainder with computational complexity $O(n \log^2 n)$ for a Toeplitz matrix of size $n \times n$.

This is a joint work with Houssam Khalil (Université Claude Bernard - Lyon I) and Michelle Schatzman (Institut Camille Jordan, Lyon).

6.2.3. Budan Tables of Real Univariate Polynomials

Participant: André Galligo.

The Budan table of f collects the signs of the iterated derivative of f. We revisit the classical Budan-Fourier theorem for a univariate real polynomial f and establish a new connexity property of its Budan table. In [18], we use this property to characterize the virtual roots of f, (introduced by Gonzales-Vega, Lombardi, Mahe in 1998); they are continuous functions of the coefficients of f. We also consider a property (P) of a polynomial f, which is generically satisfied, it eases the topological-combinatorial description and study of the Budan tables. A natural extension of the information collected by the virtual roots provides alternative representations of (P)-polynomials; while an attached tree structure allows a strati fication of the space of (P)-polynomials.

6.2.4. A polynomial approach for extracting the extrema of a spherical function and its application in diffusion MRI

Participant: Bernard Mourrain.

Antipodally symmetric spherical functions play a pivotal role in diffusion MRI (Magnetic Resonance Imaging) in representing sub-voxel-resolution microstructural information of the underlying tissue. This information is described by the geometry of the spherical function. In [20], we propose a method to automatically compute all the extrema of a spherical function. We then classify the extrema as maxima, minima and saddle-points to identify the maxima. We take advantage of the fact that a spherical function can be described equivalently in the spherical harmonic (SH) basis, in the symmetric tensor (ST) basis constrained to the sphere, and in the homogeneous polynomial (HP) basis constrained to the sphere. We extract the extrema of the spherical function by computing the stationary points of its constrained HP representation. Instead of using traditional optimization approaches, which are inherently local and require exhaustive search or re-initializations to locate multiple extrema, we use a novel polynomial system solver which analytically brackets all the extrema and refines them numerically, thus missing none and achieving high precision. To illustrate our approach we consider the Orientation Distribution Function (ODF). In diffusion MRI the ODF is a spherical function which represents a state-of-the-art reconstruction algorithm whose maxima are aligned with the dominant fiber bundles. It is, therefore, vital to correctly compute these maxima to detect the fiber bundle directions. To demonstrate the potential of the proposed polynomial approach we compute the extrema of the ODF to extract all its maxima. This polynomial approach is, however, not dependent on the ODF and the framework presented in this paper can be applied to any spherical function described in either the SH basis, ST basis or the HP basis.

This is a joint work with Aurobrata Ghosh (Inria, EPI ATHENA), Elias Tsigaridas (Inria, EPI POLSYS), Rachid Deriche (Inria, EPI ATHENA).

6.2.5. The geometry of sound-source localization using non-coplanar microphone arrays

The paper [29] addresses the task of sound-source localization from time delay estimates using arbitrarily shaped non-coplanar microphone arrays. We fully exploit the direct path propagation model and our contribution is threefold: we provide a necessary and sufficient condition for a set of time delays to correspond to a sound source position, a proof of the uniqueness of this position, and a localization mapping to retrieve it. The time delay estimation task is casted into a non-linear multivariate optimization problem constrained by necessary and sufficient conditions on time delays. Two global optimization techniques to estimate time delays and localize the sound source are investigated. We report an extensive set of experiments and comparisons with state-of-the-art methods on simulated and real data in the presence of noise and reverberations.

This is a joint work with Xavier Alameda-Pineda (Inria, EPI PERCEPTION) and Radu Horaud (Inria, EPI PERCEPTION).

6.2.6. Rational Invariants of a Group Action

Participant: Evelyne Hubert.

The article [28] is based on introductory lectures delivered at the Journées Nationales de Calcul Formel that took place at the Centre International de Recherche en Mathématiques (2013) in Marseille. We introduce basic notions on algebraic group actions and their invariants. Based on geometric considerations, we present algebraic constructions for a generating set of rational invariants. In particular the use of sections and quasi-sections contribute to increased efficiency and reduced output size. The notion of sections is refined compared to the cross-sections used in [9].

6.2.7. Rational Invariants of Finite Abelian groups

Participant: Evelyne Hubert.

In [36] we investigate the field of rational invariants of the linear action of a finite abelian group in the non modular case. By diagonalization, the group is accurately described by an integer matrix of exponents. We make use of linear algebra to compute a minimal generating set of invariants and the substitution to rewrite any invariant in terms of this generating set. We show that the generating set can be chosen to consist of polynomial invariants. As an application, we provide a symmetry reduction scheme for polynomial systems the solution set of which are invariant by the group action.

This is joint work with George Labahn, University of Waterloo, Ontario (Canada).

6.2.8. Exact relaxation for polynomial optimization on semi-algebraic sets

Participants: Marta Abril Bucero, Bernard Mourrain.

In [31], we study the problem of computing by relaxation hierarchies the infimum of a real polynomial function f on a closed basic semialgebraic set S and the points where this infimum is reached, if they exist. We show that when the infimum is reached, a relaxation hierarchy constructed from the Karush-Kuhn-Tucker ideal is always exact and that the vanishing ideal of the KKT minimizer points is generated by the kernel of the associated moment matrix in that degree, even if this ideal is not zero-dimensional. We also show that this relaxation allows to detect when there is no KKT minimizer. We prove that the exactness of the relaxation depends only on the real points which satisfy these constraints. This exploits representations of positive polynomials as elements of the preordering modulo the KKT ideal, which only involves polynomials in the initial set of variables. The approach provides a uniform treatment of different optimization problems considered previously. Applications to global optimization, optimization on semialgebraic sets defined by regular sets of constraints, optimization on finite semialgebraic sets, real radical computation are given.

6.3. Symbolic-Numeric Analysis

6.3.1. Numerical Reconstruction of Convex Polytopes from Directional Moments

Participants: Mathieu Collowald, Evelyne Hubert.

In [35] we address the reconstruction of convex polytopes, in any dimension n, from the knowledge of a finite set of directional moments of the shape. Starting with the formula relating the projection of the vertices to the directional moments, we employ established numerical algorithms for generalized eigenvalues and interval interpolation to recover the coordinates of the vertices. We perform the reconstruction of a diamond cut using our novel method.

This is joint work with Annie Cuyt, Wen-Shin Lee and Oliver Salazar Celis from University of Antwerp (Belgium).

6.3.2. Bulbous Bow Shape Optimization

Participant: Bernard Mourrain.

The aim of the work [30] is to prove the usefulness of a bulbous bow for a fishing vessel, in terms of drag reduction, using an automated shape optimization procedure including hydrodynamic simulations. A bulbous bow is an appendage that is known to reduce the drag, thanks to its influence on the bow wave system. However, the definition of the geometrical parameters of the bulb, such as its length and thickness, is not intuitive, as both parameters are coupled with regards to their influence on the final drag. Therefore, we propose to use an automated shape optimization procedure, based on a high-fidelity flow solver, a surrogate model-based optimizer and a CAD-based geometrical model, to derive the characteristics of the bow geometry allowing to maximize the achievable drag reduction. The numerical tools are first presented, and then applied to the optimization of a bow shape for a real fishing vessel, in order to determine the optimal length and thickness of the bow for drag reduction purpose.

This is a joint work with Louis Blanchard and Régis Duvigneau (Inria, EPI OPALE), Elisa Berrini (MyCFD), Yann Roux (K-Epsilon) Eric Jean (Jean & Frasca Design).

GALEN Project-Team

6. New Results

6.1. Shape, Grouping and Recognition

6.1.1. Descriptors

Participants: Eduard Trulls, Iasonas Kokkinos.

In [30] we have extended our prior work on dense scale- and rotation- invariant image descriptors to take into account soft segmentation information. This allows us to discard measurements stemming from background structures, and as such renders our descriptors invariant to background changes and occlusions. This has allowed us to obtain state-of-the-art results on tasks such as large-displacement optical flow and wide-baseline stereo. We have made the implementation of these descriptors publicly available.

6.1.2. 3D structure detection

Participants: Haithem Boussaid, Iasonas Kokkinos.

In [22] we have started exploring the potential of combinatorial optimization in the medical imaging realm. We cast the problem of finding a 3D structure (a brain tumor) as that of finding the mode of a nonparametric distribution, constructed through Kernel Density Estimation. Current techniques for doing this (e.g. Mean Shift mode-seeking, Fast Gauss Transforms, etc.) are either iterative, or linear in the number of pixels, with a typically large constant. Instead, we develop a scheme that involves a very low-constant linear-time preprocessing step, and then uses Branch-and-Bound for fast mode estimation. As such it is scalable to large volumes, and serves as a rapid initialization of a region segmentation algorithm.

6.1.3. Facade parsing

Participants: Olivier Teboul, Iasonas Kokkinos, Loic Simon, Panagiotis Katsourakis, Nikos Paragios.

In [17] we pursue a Reinforcement Learning-based approach to couple image observations with a grammar-based method to partitioning a building facade. For this we expressed 2D grammar- based image parsing as a Markov decision process where an agent has to take actions in an environment so as to maximize some notion of cumulative reward (reflecting the segmentation quality). This allowed us to accelerate previous stochastic hill-climbing approaches to image parsing by more than an order of magnitude.

6.1.4. Fast object detection

Participant: Iasonas Kokkinos.

In [27] we extended our previous work on fast object detection by developing a sparse-coding method for the efficient sharing of computation among multiple object models. In particular the first processing step of 'part score' computation was originally performed separate per object category; instead, we propose to do it 'in batch mode', so as to exploit the commonalities that exist among object parts. Building on recent developments in sparse coding we have managed to construct a compact basis for this task, which in the end gave us a two-fold acceleration over our previous fastest algorithms.

6.2. Machine Learning

6.2.1. Discriminative Parameter Estimation for Random Walks Segmentation

Participants: Pierre-Yves Baudin, Puneet Kumar, Noura Azzabou, Pierre Carlier, Nikos Paragios, M. Pawan Kumar Blaschko

In [19], we proposed a a novel discriminative learning framework that estimates the parameters of a random walks segmentation framework using a training dataset. The main challenge we face is that the training samples are not fully supervised. Specifically, they provide a hard segmentation of the medical images, instead of a probabilistic segmentation. We overcome this challenge by treating the optimal probabilistic segmentation that is compatible with the given hard segmentation as a latent variable. This allows us to employ the latent support vector machine (LSVM) formulation for parameter estimation.

6.2.2. Structured Sparsity & Applications

Participants: Katerina Gkirtzou, Wojciech Zaremba, Matthew Blaschko, M. Pawan Kumar, Nikos Paragios

We developed several machine learning applications to fMRI data, including graph representations [25] and structured sparsity regularization [26], [44]. A similar structured sparsity approach was applied in the development of a novel learning algorithm, the k-support regularized SVM, with applications to neuromuscular disease classification from diffusion tensor imaging [24]. Efficient training applications for taxonomic classification were developed in [21], while a fine grained taxonomic image classification task was introduced in [45]. The role of non-maximal suppression in accurate and efficient object detection cascades was elucidated in [20]. A fast, consistent two-sample test based on kernelized statistics was developed in [33].

6.2.3. Learning from M/EEG Data with Variable Brain Activation Delays

Participants: Wojciech Zaremba, Alexander Gramfort, M. Pawan Kumar, Matthew Blaschko

In [34], propose to address the misalignment of M/EEG samples by explicitly modeling time shifts of different brain responses in a classification setup. To this end, we use the LSVM formulation, where the latent shifts are inferred while learning the classifier parameters. The inferred shifts are further used to improve the signal-to-noise ratio of the M/EEG data, and to infer the chronometry and the sequence of activations across the brain regions that are involved in the experimental task.

6.3. Biomedical Image Analysis

6.3.1. Reconstruction

Participants: Helen Langet, Nikos Paragios

In [38] an overview of the methodological foundations of biomedical image analysis as well as their use to provide answers to a variety of clinical problems are presented. The problem of volumes of rotational angiography using non-linear sparsity constraints was studied in [28] where a novel method able to handle highly under-sampled acquisitions was introduced.

6.3.2. Graphical models and Image Segmentation

Participants: Bo Xiang, Nikos Paragios

[18] presents an overview of the use of graphical models in artificial vision where both inference, learning as well as applications are discussed. In [32] a max-margin dual decomposition method was used towards learning the compact, pose invariant shape representation using higher order graphs acting both on the connectivity of the graph as well its potentials. Graphical model was used as prior in [13] under a "curve" propagation principle for generic prior-constrained organ segmentation in 2D images. Similar inspiration driven from a higher order pose invariant graphical model learned according to [32] was considered in [31] where a novel segmentation method was proposed coupling model-based and pixel-based concepts while being pose invariant. The underlying idea was to consider a two-layer interconnected graphical model acting on pixel and on control points where segmentation consistency was imposed throgh penalties on label discrepancies of the different layers. Higher order graphical models were also employed in [14] for spine segmentation using an articulated graphical model where a non-linear approach/embedding towards reducing the complexity of the inference step was considered at training.

6.3.3. Deformable Registration and Fusion

Participants: Enzo Ferrante, Sarah Parisot, Nikos Paragios

In [16] a comprehensive survey of deformable registration was presented. It was organized in three sections: the first was studying the deformation model, the second the similarity criterion while the last section discussed the different optimization strategies. The problem of atlas-based segmentation/registration in the presence of brain tumors was studied in [29] an adaptive uncertainty-driven sampling strategy was proposed coupling segmentation and registration. Both sampling spaces (quantization of the search space, deformation grid) were determined according to the observed optimization min-marginals. The challenging problem of image to slice registration was proposed in [23] where an over-parameterized low rank graphical model acting both on the plan selection as well the in-plane deformations was introduced. The main strength of the method was its ability to simultaneously recover both the plane and the organ deformation.

GALLIUM Project-Team

6. New Results

6.1. Formal verification of compilers and static analyzers

6.1.1. The CompCert formally-verified compiler

Participants: Xavier Leroy, Jacques-Henri Jourdan, Robbert Krebbers.

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 [6]. This compiler comprises a back-end part, translating the Cminor intermediate language to PowerPC assembly and reusable for source languages other than C [5], and a front-end translating the CompCert C subset of C to Cminor. The compiler is mostly written within the specification language of the Coq proof assistant, from which Coq's extraction facility generates executable Caml code. The compiler comes with a 50000-line, machine-checked Coq proof of semantic preservation establishing that the generated assembly code executes exactly as prescribed by the semantics of the source C program.

This year we released three versions of CompCert. Version 1.13, released in March, improves conformance with the ISO C standard by defining the semantics of comparisons involving pointers "one past" the end of an array. Such comparisons used to be undefined behaviors in earlier versions of CompCert. Robbert Krebbers formalized a reasonable interpretation of the ISO C rules concerning pointers "one past" and adapted CompCert's proofs accordingly. CompCert 1.13 also features minor performance improvements for the ARM and PowerPC back-ends, notably for parameter passing via stack locations.

Version 2.0 of CompCert, released in June, re-architects the compiler back-end around the new register allocator described in section 6.1.2 . Besides improving the performance of generated code, this new allocator made it possible to add support for 64-bit integers, that is, the `long long` and `unsigned long long` data types of ISO C99. Most arithmetic operations over 64-bit integers are expanded in-line and proved correct, but a few complex operations (division, modulus, and conversions to and from floating-point numbers) are implemented as calls into library functions.

Moreover, conformance with Application Binary Interfaces was improved, especially concerning the passing of function parameters and results of type `float` (single-precision FP numbers).

Finally, CompCert 2.0 features preliminary support for debugging information. The `-g` compiler flag causes DWARF debugging information to be generated for line numbers and call stack structure. However, no information is generated yet for C type definitions and variable declarations.

Version 2.1, released in October, addresses several shortcomings of CompCert for embedded system codes, as identified by Airbus during their experimental evaluation of CompCert. In particular, CompCert 2.1 features the `_Alignas` modifier introduced in ISO C2011, to support precise control of alignment of global variables and structure fields, and uses this modifier to implement packed structures in a more robust fashion than in earlier releases. Xavier Leroy also implemented and proved correct the optimization of integer divisions by constants introduced by Granlund and Montgomery [40].

6.1.2. Register allocation with validation a posteriori

Participant: Xavier Leroy.

Register allocation (the placement of program variables in processor registers) has a tremendous impact on the performance of compiled code. However, advanced register allocation techniques are difficult to prove correct, as they involve complex algorithms and data structures. Since the beginning of the CompCert project, we chose to avoid some of these difficult proofs by performing validation *a posteriori* for part of register allocation: the IRC graph coloring algorithm invoked during register allocation is not proved correct; instead, its results are verified at every compiler run to be a correct coloring of the given interference graph, using a simple validator proved sound in Coq.

In CompCert 2.0, we push this validation-based approach further. The whole register allocator is now subject to validation a posteriori and no longer needs to be proved correct. The validator follows the algorithm invented by Rideau and Leroy [50] and further developed by Tassarotti and Leroy. It proceeds by backward dataflow analysis of symbolic equations between program variables, registers, and stack locations.

Consequently, the new register allocator for CompCert 2.0 is much more aggressive than that of CompCert 1: it features a number of optimizations that could not be proved correct in CompCert, including live-range splitting, better handling of two-address operations and other irregularities of the x86 instruction set, an improved spilling strategy, and iterating register allocation to place temporaries introduced by spilling. Moreover, the new register allocator can handle program variables of 64-bit integer types, allocating them to pairs of 32-bit registers or stack locations. The new register allocator improves the performance of generated x86 code by up to 10% on our benchmarks.

6.1.3. Formal verification of static analyzers based on abstract interpretation

Participants: Sandrine Blazy [EPI Celtique], Vincent Laporte [EPI Celtique], Jacques-Henri Jourdan, Xavier Leroy, David Pichardie [EPI Celtique].

In the context of the ANR Verasco project, we are investigating the formal specification and verification in Coq of a realistic static analyzer based on abstract interpretation. This static analyzer should be able to handle the same large subset of the C language as the CompCert compiler; support a combination of abstract domains, including relational domains; and produce usable alarms. The long-term goal is to obtain a static analyzer that can be used to prove safety properties of real-world embedded C codes.

This year, Jacques-Henri Jourdan worked on numerical abstract domains for the static analyzer. First, he designed, programmed and proved correct an abstraction layer that transforms any relational abstract domain for mathematical, arbitrary-precision integers into a relational abstract domain for finite-precision machine integers, taking overflow and “wrap-around” behaviors into account. This domain transformer makes it possible to design numerical domains without taking into account the finiteness of machine integers. Then, he implemented and proved sound non-relational abstract domains for intervals of integers and of floating-point numbers, supporting almost all CompCert arithmetic operations.

In collaboration with team Celtique, we studied which intermediate languages of the CompCert C compiler are suitable as source language for the static analyzer. Early work by Blazy, Laporte, Maroneze and Pichardie [36] performs abstract interpretation over the RTL intermediate language, a simple language with unstructured control (control-flow graph). However, this language is too low-level to support reporting alarms at the level of the source C program.

Later this year, we decided to use the C#minor intermediate language of CompCert as source language for analysis. This language has mostly structured control (if/then/else, C loops, and `goto`), and is much closer to the source C program. Then, Jacques-Henri Jourdan, Xavier Leroy and David Pichardie designed a generic abstract interpreter for the C#minor language, parameterized by an abstract domain of execution states, using structured fixpoint iteration for loops and a function-global iteration for `goto`. Jacques-Henri Jourdan is in the process of proving the soundness of this abstract interpreter in Coq.

6.1.4. Formalization of floating-point arithmetic

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

Last year, we replaced the axiomatization of floating-point numbers and arithmetic operations used in early versions of CompCert by a fully-formal Coq development, building on the Coq formalization of IEEE-754 arithmetic provided by the Flocq library of Sylvie Boldo and Guillaume Melquiond. A paper describing this work was presented at the ARITH 2013 conference [15].

This year, we extended this formalization of floating-point arithmetic with a more precise modeling of "Not a Number" special numbers, reflecting the signs and payloads of these numbers into their bit-level, in-memory representation. We also proved correct more algebraic identities over FP computations, such as $x/2^n = x \times 2^{-n}$ if $|n| < 1023$, as well as nontrivial implementation schemes for conversions between integer and FP numbers, whose correctness rely on subtle properties of the "round to odd" rounding mode. These extensions are described in a draft journal paper under submission [29], and integrated in version 2.1 of CompCert.

6.1.5. Formal verification of hardware synthesis

Participants: Thomas Braibant, Adam Chlipala [MIT].

Verification of hardware designs has been thoroughly investigated. 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.

Continuing work started during a visit at MIT under the supervision of Adam Chlipala, Thomas Braibant worked on the implementation and proof of correctness of a prototype hardware compiler. This compiler produces descriptions of circuits in RTL style from a high-level description language inspired by BlueSpec. Formal verification of hardware designs of mild complexity was conducted at the source level, making it possible to obtain fully certified RTL designs. A paper describing this compiler and two examples of certified designs was presented at the CAV 2013 conference [16].

6.2. Language design and type systems

6.2.1. The Mezzo programming language

Participants: Jonathan Protzenko, François Pottier, Thibaut Balabonski, Armaël Guéneau, Cyprien Mangin.

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 aims to bring new answers to these questions.

This year, we:

- made significant progress on the proof of soundness, by rewriting it in a more modular fashion;
- improved the implementation, by formalizing the algorithms and rewriting significant parts of the type-checker;
- hosted two interns who explored arithmetic reasoning and modeling of the iterator protocol, respectively;
- formalized libraries for concurrent programming in Mezzo;
- wrote both an interpreter and a compiler for the language.

A paper on Mezzo appeared in the ICFP 2013 conference [21].

During the previous year (2012), François Pottier wrote a formal definition of Mezzo, and proved that Mezzo is type-safe: that is, well-typed programs cannot crash. The proof was machine-checked using Coq. This year, Thibaut Balabonski and François Pottier extended this formalization with support for concurrency and dynamically-allocated locks, and proved that well-typed programs not only cannot crash, but also are data-race free.

The structure of the proof was re-worked so as to make it more modular. A paper, which emphasizes this modularity, has been submitted for presentation at a conference.

The new concurrent features have been integrated in the core library of Mezzo by Thibaut Balabonski. Further concurrent libraries have been included to provide more communication primitives, such as channels for message passing.

Jonathan Protzenko worked on formalizing the type-checking algorithms currently used in the Mezzo prototype compiler. This led to practical results in the form of improvements to the type-checker: we now type-check more programs, and the success of the type-checker is more predictable as well. Some soundness bugs have been identified and fixed. The design of some of the language's features has been improved as well.

The formalization of the type-checker was presented at the IFL 2013 conference, and is to appear in the post-symposium proceedings in 2014.

We set out to promote Mezzo in the wild. Protzenko packaged the software to make it available widely via OPAM, wrote a tutorial for end-users [34], communicated through blog posts about the language, and released the source code online for others to contribute.

We also spread the word about Mezzo through various seminar talks and discussions with other teams (Carnegie-Mellon university, Cambridge Computer Lab, Aarhus University, Brasilia University), and by communicating in international conferences (ICFP'13, FSFMA'13).

This year, two interns worked with us on Mezzo. Armaël Guéneau (L3; June-July 2013) and Cyprien Mangin (M1; April-July 2013) explored several experimental aspects of the language. In particular, Armaël worked on an encoding of iterators in an object-oriented style, which involves transfers of ownership and typestate changes; while Cyprien improved the treatment of arrays and implemented an experimental extension of Mezzo with arithmetic assertions. Armaël presented his work at the workshop HOPE 2013. This work is also described in a short unpublished paper [33].

6.2.2. System F with coercion constraints

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, we extended System F with a richer type-level language and a proposition language. Propositions contain a first-order logic, a coinduction mechanism, coherence assertions and coercion assertions. Types are classified by kinds and extended in order to handle lists of types. We introduce a particular kind restricting a previous kind to its types satisfying a proposition. Abstracting over such a kind means abstracting over arbitrary propositions, and thus enables coercion abstraction. Type abstraction must be coherent: the kind of the abstract type has to be inhabited by a witness type. This language, called Fcc, extends our previous language parametric F-iota and additionally subsumes Constraint ML.

We also extended Fcc with incoherent polymorphism in order to model GADTs. Unlike coercions and thus coherent polymorphism, incoherent polymorphism is not erasable. But in counterpart, incoherent abstraction does not require the kind to be inhabited. Since abstracting over incoherent types permits to write unsound terms, incoherent abstraction has to block the reduction of terms.

This work is part of Julien Cretin's Ph.D. dissertation [11], which will be defended in January 2014.

6.2.3. Type inference for GADTs

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

Type inference for generalized algebraic data types (GADTs) is inherently non monotone: assuming more specific types for GADTs may ensure more invariants, which may result in more general types. This is problematic for type inference and some amount of type annotations is required.

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 create an ambiguity when leaving the scope of the equation: which element should be used for representing the equivalent forms? Idealy, one should use a type disjunction, but this is not allowed—for good reasons. Hence, to avoid arbitrary choices, these situations must be rejected as ambiguous, forcing the user to write more annotations to resolve the 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*, which are a way to trace unavoidable uses of equations within types themselves. We then redefined ambiguities so that only ambivalent types become ambiguous and should be rejected or resolved by a programmer annotation. Interestingly, this solution is fully compatible with unification-based type inference algorithms used in ML dialects.

This work was presented at the APLAS 2013 conference [20]. It is also implemented in the OCaml language since version 4.00.

6.2.4. 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. A paper describing our formalization was presented at the ESOP 2013 conference [23].

6.2.5. Singleton types for code inference

Participants: Gabriel Scherer, Didier Rémy.

We continued working on 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. As this is still work in progress, there was no publication on this topic this year, but we presented our directions on three occasions: at the PLUME team in ENS Lyon, at the LIX team in École Polytechnique (whose proof-search research is highly relevant to our work), and at the Dependently Typed Programming workshop (satellite of the International Conference on Functional Programming) in Boston.

6.2.6. Open closure types

Participants: Gabriel Scherer, Jan Hoffmann [Yale University, FLINT group].

During a visit to Yale, Gabriel Scherer worked with Jan Hoffmann on a type system for program analysis of higher-order functional languages. Open closure types are a novel typing construct that lets the type system statically reason about closure variables present in the lexical context. This allows fine-grained analysis (e.g., for resource consumption or information-flow control) of functional programming patterns such as function currying. This work was presented at the LPAR 2013 conference [22] (Logic for Programming, Artificial Intelligence, and Reasoning) in October.

6.3. Shared-memory parallelism

6.3.1. Algorithms and data structures for parallel computing

Participants: Umut Acar, Arthur Charguéraud [EPI Toccata], Mike Rainey.

The ERC Deepsea project, with principal investigator Umut Acar, started in June and is hosted by the Gallium team. This project aims at developing techniques for parallel and self-adjusting computations in the context of shared-memory multiprocessors (i.e., multicore platforms). The project is continuing work that began at Max Planck Institute for Software Systems in the previous three years. As part of this project, we are developing a C++ library, called PASL, for programming parallel computations at a high level of abstraction. We use this library to evaluate new algorithms and data structures. We have recently been pursuing two main lines of work.

We have been developing an algorithm that is able to perform dynamic load balancing in the style of work stealing but without requiring atomic read-modify-write operations. These operations may scale poorly with the number of cores due to synchronization bottlenecks. We have designed the algorithm, proved it correct using a new technique for the x86-TSO weak memory model. We have evaluated our algorithm on a modern multicore machine. Although we use no synchronization operations, we achieve performance that is no more than a few percent slower than the industrial-strengh algorithm, even though the industrial-strength algorithm takes full advantage of synchronization operations. We have a soon-to-be-submitted research article describing our contributions [25].

The design of efficient parallel graph algorithms requires a sequence data structure that supports logarithmic-time split and concatenation operations in addition to push and pop operations with excellent constant factors. We have designed such a data structure by building on a recently introduced data structure called Finger Tree and by integrating a "chunking" technique. Our chunking technique is based on instantiating the leaves of the Finger Tree with chunks of contiguous memory. Unlike previous chunked data structures, we are able to prove efficient constant factors even in worst-case scenarios. Moreover, we implemented our data structure in C++ and OCaml and showed it to be competitive with state-of-the-art sequence data structures that do not support split and concatenation operations. We are currently writing a report on our results.

6.3.2. Weak memory models

Participants: Luc Maranget, Jacques-Pascal Deplaix, Jade Alglave [University College London].

Modern multicore and multiprocessor computers do not follow the intuitive "Sequential Consistency" model that would define a concurrent execution as the interleaving of the execution of its constituting threads and that would command instantaneous writes to the shared memory. This situation is due both to in-core optimisations such as speculative and out-of-order execution of instruction and to the presence of sophisticated (and cooperating) caching devices between processors and memory.

In the last few years, Luc Maranget took part in an international research effort to define the semantics of the computers of the multi-core era. This research effort relies both on formal methods for defining the models and on intensive experiments for validating the models. Joint work with, amongst others, Jade Alglave (now at University College London) and Peter Sewell (University of Cambridge) achieved several significant results, including two semantics for the IBM Power and ARM memory models: one of the operational kind [52] and the other of the axiomatic kind [46]. In particular, Luc Maranget is the main developer of the **diy** tool suite (see section 5.3). Luc Maranget also performs most of the experiments involved.

In 2013, Luc Maranget pursued this collaboration. He mainly worked with Jade Alglave to produce a new model for Power/ARM. The new model is simpler than the previous ones, in the sense that it is based on fewer mathematical objects and can be simulated more efficiently than the previous models. The new model is at the core of a journal submission which is now at the second stage of reviewing. The submitted work contains in-depth testing of ARM devices which led to the discovery of anomalous behaviours acknowledged as such by our ARM contact, and of legitimate features now included in the model. The new model also impacted our **diy** tool suite that now includes a generic memory model simulator built by following the principles exposed in the submitted article. At the moment the new simulator is available as an experimental release (http://diy.inria.fr/herd). It will be include in future releases of the tool suite.

In the same research theme, Luc Maranget supervises the internship of Jacques-Pascal Deplaix (EPITECH), from Oct. 2013 to May 2014. The internship aims at extending **litmus**, our tool to to run tests on hardware: at the moment **litmus** accepts test written in assembler; Jacques-Pascal is extending **litmus** so that it accepts tests written in C. The general objective is to achieve conformance testing of C compilers and machines with respect to the new C11/C++11 standard.

6.4. The OCaml language and system

6.4.1. The OCaml system

Participants: Damien Doligez, Alain Frisch [Lexifi SAS], Jacques Garrigue [University of Nagoya], Fabrice Le Fessant, Xavier Leroy, Gabriel Scherer.

This year, we released version 4.01.0 of the OCaml system. This is a major release that fixes about 140 bugs and introduces 44 new features suggested by users. Damien Doligez acted as release manager for this version.

The major innovations in OCaml 4.01 are:

- The overloading of variant constructors and record field labels, resolved using typing information. Before this, programmers had to use globally unique field labels across all record types. The new typechecking algorithm enables programmers to use more natural names for fields in their data structures. The algorithm is carefully engineered to preserve principality of inferred types.
- New warnings give the programmer the option of applying very strict checking of problematic constructs in the source code.

Other features of this release include:

- Suggestion of possible typos in case of “unbound identifier” error.
- New infix application operators in the standard library.
- Options to reduce the verbosity (and enhance the readability) of error messages.
- Many internal improvements, especially in compiler performance.

In parallel, we designed and experimented with several new features that are candidate for inclusion in the next major release of OCaml in 2014:

- Module aliases: a more efficient way of typechecking and compiling module declarations of the form `module M = ModuleName`, providing a lighter, more practical alternative to packed modules and reducing the need for name spaces.
- Extension points and preprocessing by rewriting abstract syntax trees: this approach provides an alternative to Camlp4 for macro processing and automatic code generation.
- A native code generator for the new ARM 64 bit instruction set (also known as AArch64).
- Several ongoing experiments to improve the performance of OCaml-compiled code: more aggressive function inlining and constant propagation; more unboxing of numbers; and a pass of common subexpression elimination.

6.4.2. Run-time types for the OCaml language

Participants: Grégoire Henry, Jacques Garrigue [University of Nagoya], Fabrice Le Fessant.

With the addition of GADTs to OCaml in version 4.00, it is now possible to provide a clean implementation of run-time types in the language, thus allowing the definition of polytypic function, a.k.a. generic function defined by case analysis on the structure of its argument’s type. However, when integrating this mechanism into the language, its interaction with other parts of the type-system proved delicate, the main difficulty being the semantic of abstract types.

In collaboration with Jacques Garrigue during a 3 month stay in Japan, Grégoire Henry worked on different semantics for the runtime representation of abstract types. They tried to design a mechanism that preserves abstraction by default, and still allows to propagate type information when requested by the programmer.

6.4.3. Multi-runtime OCaml

Participants: Luca Saiu, Fabrice Le Fessant.

Multicore architectures are now broadly available, and developers expect their programs to be able to benefit from them. In OCaml, there is no portable way to use such architectures, as only one OCaml thread can run at any time.

As part of the ANR project "BWare", Luca Saiu and Fabrice Le Fessant developed a multi-runtime version of OCaml that takes advantage of multicore architectures. In this version, a program can start several runtimes that can run on different cores. As a consequence, OCaml threads running on different runtimes can run concurrently. This implementation required a lot of rewriting of the OCaml runtime system (written in C), to make all global variables context-dependent and all functions reentrant. The compiler was also modified to generate reentrant code and context-dependent variables. The sources of the prototype were released in September 2013, to be tested by users.

Luca Saiu then developed a library based on skeletons to facilitate the development of parallel applications that take advantage of the multi-runtime architecture.

6.4.4. Evaluation strategies and standardization

Participants: Thibaut Balabonski, Flávio de Moura [Universidade de Brasília].

During the past years, Thibaut Balabonski studied evaluation strategies, laziness and optimality for functional programming languages, in particular in relation to pattern matching. These investigations continued this year, with two highlights:

- Publication in the ICFP conference [14] of a theoretical result relating fully lazy evaluation (as can be found in some Haskell compilers) to optimal reduction in the weak λ-calculus.
- Collaboration with Flávio de Moura (Universidade de Brasília) on so-called "standard" evaluation strategies for a calculus with rich pattern matching mechanisms (the *Pure Pattern Calculus* of Jay and Kesner [42]). The challenge here lies in that the calculus does not satisfies the usual stability properties. As a consequences, standard strategies are not unique anymore, and new approaches are needed. A paper is in preparation.

6.5. Software specification and verification

6.5.1. Tools for TLA+

Participants: Damien Doligez, Jael Kriener, Leslie Lamport [Microsoft Research], Stephan Merz [EPI VeriDis], Tomer Libal [Microsoft Research-Inria Joint Centre], Hernán Vanzetto [Microsoft Research-Inria Joint Centre].

Damien Doligez is head of the "Tools for Proofs" team in the Microsoft-Inria Joint Centre. The aim of this team is to extend the TLA+ language with a formal language for hierarchical proofs, formalizing the ideas in [43], and to build tools for writing TLA+ specifications and mechanically checking the corresponding formal proofs.

This year, the TLA+ tools were released as open-source (MIT license), and in September we released a new version of the TLA+ Proof System (TLAPS), an environment for writing and checking TLA+ proofs. This environment is described in [38].

We have implemented a (not yet released) extension of TLAPS to deal with proofs of temporal formulas, using the propositional temporal logic prover LS4 as a back-end. Until now, TLAPS could only be used to prove safety properties (invariants). With this new version, our users will be able to prove liveness properties (absence of deadlock), refinement relations between specifications, etc.

Jael Kriener started a 2-year post-doc contract in December. She is working on theoretical and implementation aspects of TLA+ and TLAPS.

Web sites:
http://research.microsoft.com/users/lamport/tla/tla.html
http://tla.msr-inria.inria.fr/tlaps

6.5.2. The Zenon automatic theorem prover

Participants: Damien Doligez, David Delahaye [CNAM], Pierre Halmagrand [CNAM], Olivier Hermant [Mines ParisTech], Mélanie Jacquel [CNAM].

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

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 [37]. Mélanie Jacquel defended her thesis on this subject in April.

Pierre Halmagrand did an internship and started a thesis on integrating Deduction Modulo in Zenon; some results of this work are described in two papers published at LPAR [19] and IWIL [18].

6.5.3. Implementing hash-consed structures in Coq

Participants: Thomas Braibant, Jacques-Henri Jourdan, David Monniaux [CNRS, VERIMAG].

Hash-consing is a programming technique used to implement maximal sharing of immutable values in memory, keeping a single copy of semantically equivalent objects. Hash-consed data-structures give a unique identifier to each object, allowing fast hashing and comparisons of objects. This may lead to major improvements in execution time by itself, but it also make it possible to do efficient memoization of computations.

Hash-consing and memoization are examples of imperative techniques that are of prime importance for performance, but are not easy to implement and prove correct using the purely functional language of a proof assistant such as Coq. In a joint article at ITP 2013 [17], we described three different implementation techniques for hash-consed data-structures in Coq through the running example of Binary Decision Diagrams (BDDs). BDDs are representations of Boolean functions, and are often used in software and hardware verification tools (e.g., model checkers).

We substantially improved the work described in this ITP 2013 article afterwards. First, we came up with a fourth implementation technique for hash-consed data-structures in Coq. Then, we performed an in-depth comparative study of how our “design patterns” for certified hash-consing fare on two real-scale examples: BDDs and lambda-terms. This work is currently under revision for publication in a journal.

6.5.4. Working with names and binders

Participant: François Pottier.

François Pottier released **dblib**, a Coq library that helps work with de Bruijn indices in a generic and lightweight manner. This library is used in the formalization of Mezzo (see section 6.2.1). It is available at http://gallium.inria.fr/~fpottier/.

6.6. Technology transfer

6.6.1. Analysis of the Scilab Language

Participants: Fabrice Le Fessant, Michael Laporte.

The Scilab language is a scripting language providing easy access to efficient implementations of mathematical operations (on matrices, for example). It suffers from the lack of verifications of an untyped language, together with the performance problems of an interpreted language. As part of the FUI Richelieu project, Fabrice Le Fessant and Michael Laporte have been investigating solutions to these issues.

The first part of the work was to clarify the semantics of the Scilab language. For that, an interpreter was implemented in OCaml, based on the C++ AST provided by the forthcoming version 6 of Scilab. This work exhibited a number of bugs in the new implementation, and proved to be more performant than the C++ implementation, thanks to a better algorithm to manage the dynamic scopes of Scilab.

The second part of the work was to understand how users write Scilab code. For that, a style-checking application, called *Scilint*, has been developed. It implements static checking of some properties of Scilab programs, to be able to detect runtime errors before running the program. Warnings are displayed for suspicious cases. Using Scilint on large sets of Scilab code (from the Scilab forge or the Atom repository) showed that the most erroneous features of Scilab are commonly used and that, to achieve the ultimate goal of partial typing of the language, a subset of the language must be specified that the user should conform to, in order for the code to benefit from the next part of the work, i.e. just-in-time compilation.

GAMMA3 Project-Team

4. New Results

4.1. From discrete to continuous metric fields

Participants: Patrick Laug [correspondant], Houman Borouchaki.

Adaptive computation using adaptive meshes is now recognized as essential for solving complex PDE problems. This computation requires at each step the definition of a continuous metric field to govern the generation of the adapted meshes. In practice, via an appropriate *a posteriori* error estimation, metrics are calculated at the vertices of the computational domain mesh. In order to obtain a continuous metric field, the discrete field is interpolated in the whole domain mesh. In this study, a new method for interpolating discrete metric fields, based on a so-called "natural decomposition" of metrics, is introduced. The proposed method is based on known matrix decompositions and is computationally robust and efficient. Some qualitative comparisons with classical methods are made to show the relevance of this methodology [19].

4.2. Hex-dominant meshing of geologic structures

Participants: Patrick Laug [correspondant], Houman Borouchaki.

Simulation by a finite volume method of the transfer by water of radioactive elements in sites of nuclear waste storage, on large time and space scales, is the only possible way to analyze the safety of disposal. To properly represent the different pathways of radionuclides, surface topography (valleys, reliefs, rivers), geologic layers and simplified storage facilities must be accurately modeled. We propose a new methodology for generating hex-dominant meshes (well suited for a finite volume formulation) of geologic structures complying with these different geometric constraints.

First, a reference 2D domain is obtained by projecting all the line constraints into a horizontal plane. Different size specifications are given for workings, outcrop lines and rivers. Using an adaptive methodology, the size variation is bounded by a specified threshold in order to obtain a high quality quad-dominant mesh. Secondly, a hex-dominant mesh of the geological medium is generated by a vertical extrusion. Depending on the configuration of the surfaces found (interfaces between two layers, top or bottom faces of underground workings), hexahedra, prisms, pyramids and tetrahedra are generated. The generation of volume elements follows a global order established on the whole set of surfaces to ensure the conformity of the resulting mesh. An example of mesh construction of a geologic structure illustrates the suitability of the proposed methodology [22].

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

Participants: Thomas Grosges [correspondant], Dominique Barchiesi, Michael François
Validité du projet: 2009-2013.
Production scientifique: 1 thèse soutenue (M. François, 17/10/2012), 6 articles publiés.
L'utilisation des nombres (pseudo)-aléatoires a pris une dimension importante ces dernières décennies. De nombreuses applications dans le domaine des télécommunications, de la cryptographie, des simulations numériques ou encore des jeux de hasard, ont contribué au développement et à l'usage de ces nombres. Les méthodes utilisées pour la génération de tels nombres (pseudo)-aléatoires proviennent de deux types de processus : physique et algorithmique. Ce projet de recherche a donc pour objectif principal le développement de nouveaux procédés de génération de clés de chiffrement, dits "exotiques", basés sur des processus physiques, multi-échelles, multi-domaines assurant un niveau élevé de sécurité. Deux classes de générateurs basés sur des principes de mesures physiques et des processus mathématiques ont été développé.

La première classe de générateurs exploite la réponse d'un système physique servant de source pour la génération des séquences aléatoires. Cette classe utilise aussi bien des résultats de simulation que des résultats de mesures interférométriques pour produire des séquences de nombres aléatoires. L'application du maillage adaptatif sert au contrôle de l'erreur sur la solution des champs physiques (simulés ou mesurés). A partir de ces cartes physiques, un maillage avec estimateur d'erreur sur l'entropie du système est appliqué. Celui-ci permet de redistribuer les positions spatiales des noeuds. L'étude (locale) de la réduction d'entropie des clés tout au long de la chaîne de création et l'étude (globale) de l'entropie de l'espace des clés générées sont réalisées à partir de tests statistiques.

La seconde classe de générateurs porte sur le développement de méthodes avancées et est basée sur l'exploitation de fonctions chaotiques en utilisant les sorties de ces fonctions comme indice de permutation sur un vecteur initial. Ce projet s'intéresse également aux systèmes de chiffrement pour la protection des données et deux algorithmes de chiffrement d'images utilisant des fonctions chaotiques sont développés et analysés. Ces Algorithmes utilisent un processus de permutation-substitution sur les bits de l'image originale. Une analyse statistique approfondie confirme la pertinence des cryptosystèmes développés.

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

Participants: Thomas Grosges [correspondant], Dominique Barchiesi, Abel Cherouat, Houman Borouchaki, Laurence Giraud-Moreau, Anis Chaari.
Validité du projet: 2011-2014.
Production scientifique: 1 thèse en cours (A. Chaari), 1 articles publiés, 1 conférence (CSMA 2013).
Ce projet de recherche (NANOMORPH) a pour objet principal le développement et la mise au point d'une instrumentation optique pour déterminer la distribution en tailles et le coefficient de forme de nanofils (NF) ou de nanotubes (NT) en suspension dans un écoulement. Au cours de ce projet, deux types de techniques optiques complémentaires sont développées. La première, basée sur la diffusion statique de la lumière, nécessite d'étudier au préalable la physico-chimie de la dispersion, la stabilisation et l'orientation des nanofils dans les milieux d'étude. La seconde méthode, basée sur une méthode opto-photothermique pulsée, nécessite en sus, la modélisation de l'interaction laser/nanofils, ainsi que l'étude des phénomènes multiphysiques induits par ce processus. L'implication de l'équipe-projet GAMMA3 concerne principalement la simulation multiphysique de l'interaction laser-nanofils et l'évolution temporelle des bulles et leurs formations. L'une des principales difficultés de ces problématiques est que la géométrie du domaine est variable (à la fois au sens géométrique et topologique). Ces simulations ne peuvent donc être réalisées que dans un schéma adaptatif de calcul nécessitant le remaillage tridimensionnel mobile, déformable avec topologie variable du domaine (formation et évolution des bulles au cours du temps et de l'espace).

4.5. Applications du maillage à des problèmes multi-physiques, développement de méthodes de résolutions avancées et modélisation électromagnetisme-thermique-mécanique à l'échelle mesoscopique

Participants: Dominique Barchiesi [correspondant], Abel Cherouat, Thomas Grosges, Houman Borouchaki, Laurence Giraud-Moreau, Sameh Kessentini, Anis Chaari, Fadhil Mezghani
Validité du projet: 2009-2015 (thèse de Fadhil Mezghani initiée en 2012 coencadrée par D. Barchiesi et A. Cherouat).
Production scientifique: 1 thèse soutenue (S. Kessentini, 22/10/2012), 9 articles publiés, 4 conférences.

Le contrôle et l'adaptation du maillage lors de la résolution de problèmes couplés ou/et non linéaires reste un problème ouvert et fortement dépendant du type de couplage physique entre les EDP à résoudre. Notre objectif est de développer des modèles stables afin de calculer les dilatations induites par l'absorption d'énergie électromagnétique, par des structures matérielles inférieures au micron. Les structures étudiées sont en particulier des nanoparticules métalliques en condition de résonance plasmon. Dans ce cas, un maximum d'énergie absorbée est attendu, accompagné d'un maximum d'élévation de température et de dilatation. Il faut en particulier développer des modèles permettant de simuler le comportement multiphysique de particules de formes quelconques, pour une gamme de fréquences du laser d'éclairage assez étendue afin d'obtenir une étude spectroscopique de la température et de la dilatation. L'objectif intermédiaire est de pouvoir quantifier la dilatation en fonction de la puissance laser incidente. Le calcul doit donc être dimensionné et permettre finalement des applications dans les domaines des capteurs et de l'ingénierie biomédicale. En effet, ces nanoparticules métalliques sont utilisées à la fois pour le traitement des cancers superficiels par nécrose de tumeur sous éclairage adéquat, dans la fenêtres de transparence cellulaire. Déposées sur un substrat de verre, ces nanoparticules permettent de construire des capteurs utilisant la résonance plasmon pour être plus sensibles (voir projet européen *Nanoantenna* et l'activité génération de nombres aléatoires). Cependant, dans les deux cas, il est nécessaire, en environnement complexe de déterminer la température locale, voire la dilatation de ces nanoparticules, pouvant conduire à un désaccord du capteur, la résonance plasmon étant très sensible aux paramètres géométriques et matériels des nanostructures. Dans ce sens, l'étude permet d'aller plus loin que la << simple >> interaction électromagnétique avec la matière du projet européen *Nanoantenna*.

Le travail de l'année 2013 a constitué en la poursuite de la 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.

4.6. Validity of rational and nonrational Lagrange finite elements of degree 1 and 2

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

A finite element is valid if its jacobian is strictly positive everywhere. The jacobian is the determinant of the jacobian matrix related to the partials of the mapping function which maps the parameter space (reference element) to the current element. Apart when it is constant, the jacobian is a polynomial whose degree is related to the degree of the finite element (but not the same in general). The value of the jacobian varies after the point where it is evaluated. Validating an element relies in finding the sign of this polynomial when one traverses the element.

Various papesr and a synthesis of those reports, shows how to calculating the jacobian of the different usual Lagrange finite elements of degree 1 and 2. To this end, we take the form of this polynomial as obtained in the classical finite element framework (shape functions and nodes) or after reformulating the element by means of a Bezier form (Bernstein polynomials and control points) which makes easier the discussion. We exhibit sufficient (necessary and sufficient in some cases) conditions to ensure the validity of a given element.

4.7. Mesh adaptation for very high-order numerical scheme

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

In the past, we have demonstrated that multi-scale anisotropic mesh adaptation is a powerful tool to accurately simulate compressible flow problems 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.

4.8. Visualization and modification of high-order curved meshes

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

During the partnership between Inria and Distene, a new visualization software has been designed. It address the typical operations that are required to quickly assess the newly algorithm developed in the team. In particular, interactive modifications of high-order curved mesh has been addressed. The software VIZIR is freely available at https://www.rocq.inria.fr/gamma/gamma/vizir/.

4.9. A changing-topology ALE numerical scheme

Participants: Frédéric Alauzet [correspondant], Nicolas Barral.

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.

4.10. Mesh adaptation for Navier-Stokes Equations

Participants: Frédéric Alauzet, Victorien Menier, Adrien Loseille [correspondant].

Adaptive simulations for Navier-Stokes equations require to propose accurate error estimates and design robust mesh adaptation algorithms (for boundary layers).

For error estimates, we design new estimates suited to accurately capture the speed profile in the boundary layers. For mesh adaptation, we design a new method to generate structured boundary layer meshes which are mandatory to accurately compute compressible flows a high Reynolds number (several millions). It couple the specification of the optimal boundary layer from the geometry boundary and moving mesh techniques to extrude the boundary layer in an already existing mesh. The main advantage of this approach is its robustness, *i.e.,* at each step of the algorithm we have always a valid mesh.

4.11. Serial and parallel cavity-based mesh adaptation

Participants: Victorien Menier, Adrien Loseille [correspondant].

A new algorithm to derive adaptive meshes has been introduced through new cavity-based algorithms. It allows to generate anisotropic surface and volume mesh along with adaptive quasi-structured elements. The later point is of main interest when dealing with viscous phenomena where a boundary layer mesh is needed [26].

In addition, a parallel version of the algorithm was designed [27].

GANG Project-Team

5. New Results

5.1. Understanding graph representations

5.1.1. Connected graph searching

5.1.1.1. Computing H-Joins with Application to 2-Modular Decomposition

Participants: Michel Habib, Antoine Mamcarz, Fabien de Montgolfier.

We present in [10], a general framework to design algorithms that compute H-join. For a given bipartite graph H, we say that a graph G admits a H-join decomposition or simply a H-join, if the vertices of G can be partitioned in |H| parts connected as in H. This graph H is a kind of pattern, that we want to discover in G. This framework allows us to present fastest known algorithms for the computation of P 4-join (aka N-join), P 5-join (aka W-join), C 6-join (aka 6-join). We also generalize this method to find a homogeneous pair (also known as 2-module), a pair M 1,M 2 such that for every vertex x∉(M 1∪M 2) and i∈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 in Graphs Comb. 3:127-139, 1987), it is a generalization of splits (a.k.a. 1-joins) and of modules. The algorithmics to compute them appears quite involved. In this paper, we describe an O(mn 2)-time algorithm computing all maximal homogeneous pairs of a graph, which not only improves a previous bound of O(mn 3) for finding only one pair (Everett et al. in Discrete Appl. Math. 72:209-218, 1997), but also uses a nice structural property of homogenous pairs, allowing to compute a canonical decomposition tree for sesquiprime graphs (i.e., graphs G having no module and such that for every vertex v∈G, G-v also has no module).

5.1.1.2. Algorithmic Aspects of Switch Cographs

Participants: Vincent Cohen-Addad, Michel Habib, Fabien de Montgolfier.

The paper [27], introduces the notion of involution module, the first generalization of the modular decomposition of 2-structure which has a unique linear-sized decomposition tree. We derive an $O(n^2)$ decomposition algorithm and we take advantage of the involution modular decomposition tree to state several algorithmic results. Cographs are the graphs that are totally decomposable w.r.t modular decomposition. In a similar way, we introduce the class of switch cographs, the class of graphs that are totally decomposable w.r.t involution modular decomposition. This class generalizes the class of cographs and is exactly the class of (Bull, Gem, Co-Gem, C_5)-free graphs. We use our new decomposition tool to design three practical algorithms for the maximum cut, vertex cover and vertex separator problems. The complexity of these problems was still unknown for this class of graphs. This paper also improves the complexity of the maximum clique, the maximum independant set, the chromatic number and the maximum clique cover problems by giving efficient algorithms, thanks to the decomposition tree. Eventually, we show that this class of graphs has Clique-Width at most 4 and that a Clique-Width expression can be computed in linear time.

5.1.1.3. LDFS-Based Certifying Algorithm for the Minimum Path Cover Problem on Cocomparability Graphs

Participants: Derek Corneil, Dalton Barnaby, Michel Habib.

For graph $G(V, E)$, a minimum path cover (MPC) is a minimum cardinality set of vertex disjoint paths that cover V (i.e., every vertex of G is in exactly one path in the cover). This problem is a natural generalization of the Hamiltonian path problem. Cocomparability graphs (the complements of graphs that have an acyclic transitive orientation of their edge sets) are a well studied subfamily of perfect graphs that includes many popular families of graphs such as interval, permutation, and cographs. Furthermore, for every cocomparability graph G and acyclic transitive orientation of the edges of $\overline{G}$ there is a corresponding poset P_G; it is easy to see that an MPC of G is a linear extension of P_G that minimizes the bump number of P_G. Although there are directly graph-theoretical MPC algorithms (i.e., algorithms that do not rely on poset formulations) for various subfamilies of cocomparability graphs, notably interval graphs, until now all MPC algorithms for cocomparability graphs themselves have been based on the bump number algorithms for posets. In this paper [5], we present the first directly graph-theoretical MPC algorithm for cocomparability graphs; this algorithm is based on two consecutive graph searches followed by a certifying algorithm. Surprisingly, except for a lexicographic depth first search (LDFS) preprocessing step, this algorithm is identical to the corresponding algorithm for interval graphs. The running time of the algorithm is $O(\min(n^2, n + \mathrm{mloglogn}))$, with the nonlinearity coming from LDFS.

5.1.1.4. Easy identification of generalized common and conserved nested intervals

Participants: Fabien de Montgolfier, Mathieu Raffinot, Irena Rusu.

In the paper [28], we explain how to easily compute gene clusters, formalized by classical or generalized nested common or conserved intervals, between a set of K genomes represented as K permutations. A $b-$nested common (resp. conserved) interval I of size $|I|$ is either an interval of size 1 or a common (resp. conserved) interval that contains another b-nested common (resp. conserved) interval of size at least $|I| - b$. When $b = 1$, this corresponds to the classical notion of nested interval. We exhibit two simple algorithms to output all b-nested common or conserved intervals between K permutations in $O(Kn + \mathrm{nocc})$ time, where nocc is the total number of such intervals. We also explain how to count all b-nested intervals in O(Kn) time. New properties of the family of conserved intervals are proposed to do so.

5.1.1.5. On computing the diameter of real-world undirected graphs

Participants: Pierluigi Crescenzi, Roberto Grossi, Michel Habib, Leonardo Lanzi, Andrea Marino.

We propose in [2],a new algorithm for the classical problem of computing the diameter of undirected unweighted graphs, namely, the maximum distance among all the pairs of nodes, where the distance of a pair of nodes is the number of edges contained in the shortest path connecting these two nodes. Although its worst-case complexity is O(nm) time, where n is the number of nodes and m is the number of edges of the graph, we experimentally show that our algorithm works in O(m) time in practice, requiring few breadth-first searches to complete its task on almost 200 real-world graphs.

5.1.1.6. Toward more localized local algorithms: removing assumptions concerning global knowledge

Participants: Amos Korman, Jean-Sébastien Sereni, Laurent Viennot.

Numerous sophisticated local algorithm were suggested in the literature for various fundamental problems. Notable examples are the MIS and $(\Delta + 1)$-coloring algorithms by Barenboim and Elkin, by Kuhn, and by Panconesi and Srinivasan, as well as the $o(\Delta^2)$-coloring algorithm by Linial. Unfortunately, most known local algorithms (including, in particular, the aforementioned algorithms) are *non-uniform*, that is, they assume that all nodes know good estimations of one or more global parameters of the network, e.g., the maximum degree Δ or the number of nodes n. This paper [11], provides a rather general method for transforming a non-uniform local algorithm into a *uniform* one. Furthermore, the resulting algorithm enjoys the same asymptotic running time as the original non-uniform algorithm. Our method applies to a wide family of both deterministic and randomized algorithms. Specifically, it applies to almost all of the state of the art non-uniform algorithms regarding MIS and Maximal Matching, as well as to many results concerning the coloring problem. (In particular, it applies to all aforementioned algorithms.) To obtain our transformations we introduce a new distributed tool called *pruning* algorithms, which we believe may be of independent interest.

5.1.2. Self-organizing Flows in 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 this paper [24], 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.

5.2. Large Scale Networks Performance and Modeling

5.2.1. Can P2P Networks be Super-Scalable?

Participants: François Baccelli, Fabien Mathieu, Ilkka Norros, Rémi Varloot.

We propose in [14], a new model for peer-to-peer networking which takes the network bottlenecks into account beyond the access. This model can cope with key features of P2P networking like degree or locality constraints together with the fact that distant peers often have a smaller rate than nearby peers. Using a network model based on rate functions, we give a closed form expression of peers download performance in the system's fluid limit, as well as approximations for the other cases. Our results show the existence of realistic settings for which the average download time is a decreasing function of the load, a phenomenon that we call super-scalability.

5.2.2. Contenu généré par les utilisateurs : une étude sur DailyMotion

Participants: Yannick Carlinet, The Dang Huynh, Bruno Kauffmann, Fabien Mathieu, Ludovic Noirie, Sébastien Tixeuil.

Actuellement, une large part du trafic Internet vient de sites de "User-Generated Content" (UGC). Comprendre les caractéristiques de ce trafic est important pour les opérateurs (dimensionnement réseau), les fournisseurs (garantie de la qualité de service) et les équipementiers (conception d'équipements adaptés). Dans ce contexte, nous proposons [15], d'analyser et de modéliser des traces d'usage du site DailyMotion.

5.2.3. Rumor Spreading in Random Evolving Graphs

Participants: Andrea Clementi, Pierluigi Crescenzi, Carola Doerr, Pierre Fraigniaud, Isopi Marco, Alessandro Panconesi, Pasquale Francesco, Silvestri Riccardo.

In [13], we aim at analyzing the classical information spreading "push" protocol in *dynamic* networks. We consider the *edge-Markovian* evolving graph model which captures natural temporal dependencies between the structure of the network at time t, and the one at time $t+1$. Precisely, a non-edge appears with probability p, while an existing edge dies with probability q. In order to fit with real-world traces, we mostly concentrate our study on the case where $p=\Omega(\frac{1}{n})$ and q is constant. We prove that, in this realistic scenario, the "push" protocol does perform well, completing information spreading in $O(\log n)$ time steps, w.h.p., even when the network is, w.h.p., disconnected at every time step (e.g., when $p \ll \frac{\log n}{n}$). The bound is tight. We also address other ranges of parameters p and q (e.g., $p+q=1$ with arbitrary p and q, and $p=\Theta\left(\frac{1}{n}\right)$ with arbitrary q). Although they do not precisely fit with the measures performed on real-world traces, they can be of independent interest for other settings. The results in these cases confirm the positive impact of dynamism.

5.3. Complexity issues in distributed graph algorithms

5.3.1. What can be decided locally without identifiers?

Participants: Pierre Fraigniaud, Mika Göös, Amos Korman, Jukka Suomela.

Do unique node identifiers help in deciding whether a network G has a prescribed property P? We study this question in the context of distributed local decision, where the objective is to decide whether $G \in P$ by having each node run a constant-time distributed decision algorithm. If $G \in P$, all the nodes should output yes; if $G \notin P$, at least one node should output no. A recent work (Fraigniaud et al., OPODIS 2012) studied the role of identifiers in local decision and gave several conditions under which identifiers are not needed. In this article [21], we answer their original question. More than that, we do so under all combinations of the following two critical variations on the underlying model of distributed computing: (B): the size of the identifiers is bounded by a function of the size of the input network; as opposed to ($\neg B$): the identifiers are unbounded. (C): the nodes run a computable algorithm; as opposed to ($\neg C$): the nodes can compute any, possibly uncomputable function. While it is easy to see that under ($\neg B, \neg C$) identifiers are not needed, we show that under all other combinations there are properties that can be decided locally if and only if identifiers are present. Our constructions use ideas from classical computability theory.

5.3.2. Local Distributed Decision

Participants: Pierre Fraigniaud, Amos Korman, David Peleg.

A central theme in distributed network algorithms concerns understanding and coping with the issue of locality. Inspired by sequential complexity theory, we focus on a complexity theory for distributed decision problems. In the context of locality, solving a decision problem requires the processors to independently inspect their local neighborhoods and then collectively decide whether a given global input instance belongs to some specified language. Our paper [7], introduces several classes of distributed decision problems, proves separation among them and presents some complete problems. More specifically, we consider the standard LOCAL model of computation and define LD (for local decision) as the class of decision problems that can be solved in constant number of communication rounds. We first study the intriguing question of whether randomization helps in local distributed computing, and to what extent. Specifically, we define the corresponding randomized class BPLD, and ask whether LD=BPLD. We provide a partial answer to this question by showing that in many cases, randomization does not help for deciding hereditary languages. In addition, we define the notion of local many-one reductions, and introduce the (nondeterministic) class NLD of decision problems for which there exists a certificate that can be verified in constant number of communication rounds. We prove that there exists an NLD-complete problem. We also show that there exist problems not in NLD. On the other hand, we prove that the class NLD#n, which is NLD assuming that each processor can access an oracle that provides the number of nodes in the network, contains all (decidable) languages. For this class we provide a natural complete problem as well.

5.3.3. Locality and checkability in wait-free computing

Participants: Pierre Fraigniaud, Sergio Rajsbaum, Travers Corentin.

The paper [9],studies notions of locality that are inherent to the specification of distributed tasks, and independent of the computing model, by identifying fundamental relationships between the various scales of computation, from the individual process to the whole system. A locality property called *projection-closed* is identified. This property completely characterizes tasks that are wait-free *checkable*, where a task $T = (\mathcal{I}, \mathcal{O}, \Delta)$ is said to be checkable if there exists a distributed algorithm that, given $s \in \mathcal{I}$ and $t \in \mathcal{O}$, determines whether $t \in \Delta(s)$, i.e., whether t is a valid output for s according to the specification of T. Projection-closed tasks are proved to form a rich class of tasks. In particular, determining whether a projection-closed task is wait-free solvable is shown to be undecidable. A stronger notion of locality is identified by considering tasks whose outputs "look identical" to the inputs at every process: a task $T = (\mathcal{I}, \mathcal{O}, \Delta)$ is said to be *locality-preserving* if $\mathcal{O}$ is a covering complex of $\mathcal{I}$. We show that this topological property yields obstacles for wait-free solvability different in nature from the classical impossibility results. On the other hand, locality-preserving tasks are projection-closed, and thus they are wait-free checkable. A classification of locality-preserving tasks in term of their relative computational power is provided. This is achieved by defining a correspondence between subgroups of the *edgepath* group of an input complex and locality-preserving tasks. This correspondence enables to demonstrate the existence of hierarchies of locality-preserving tasks, each one containing, at the top, the universal task (induced by the universal covering complex), and, at the bottom, the trivial identity task.

5.3.4. Delays Induce an Exponential Memory Gap for Rendezvous in Trees

Participants: Pierre Fraigniaud, Pelc Andrzej.

The aim of rendezvous in a graph is meeting of two mobile agents at some node of an unknown anonymous connected graph. In this paper [8], we focus on rendezvous in trees, and, analogously to the efforts that have been made for solving the exploration problem with compact automata, we study the size of memory of mobile agents that permits to solve the rendezvous problem deterministically. We assume that the agents are identical, and move in synchronous rounds. We first show that if the delay between the starting times of the agents is *arbitrary*, then the lower bound on memory required for rendezvous is $\Omega(\log n)$ bits, even for the line of length n. This lower bound meets a previously known upper bound of $O(\log n)$ bits for rendezvous in arbitrary graphs of size at most n. Our main result is a proof that the amount of memory needed for rendezvous *with simultaneous start* depends essentially on the number ℓ of leaves of the tree, and is exponentially less impacted by the number n of nodes. Indeed, we present two identical agents with $O(\log \ell + \log\log n)$ bits of memory that solve the rendezvous problem in all trees with at most n nodes and at most ℓ leaves. Hence, for the class of trees with polylogarithmically many leaves, there is an exponential gap in minimum memory size needed for rendezvous between the scenario with arbitrary delay and the scenario with delay zero. Moreover, we show that our upper bound is optimal by proving that $\Omega(\log \ell + \log\log n)$ bits of memory are required for rendezvous, even in the class of trees with degrees bounded by 3.

5.3.5. On the Manipulability of Voting Systems: Application to Multi-Operator Networks

Participants: François Durand, Fabien Mathieu, Ludovic Noirie.

Internet is a large-scale and highly competitive economic ecosystem. In order to make fair decisions, while preventing the economic actors from manipulating the natural outcome of the decision process, game theory is a natural framework, and voting systems represent an interesting alternative that, to our knowledge, has not yet being considered. They allow competing entities to decide among different options. In this paper [20], we investigate their use for end-to-end path selection in multi-operator networks, analyzing their manipulability by tactical voting and their economic efficiency.We show that Instant Runoff Voting is much more efficient and resistant to tactical voting than the natural system which tries to get the economic optimum.

5.4. Communication and Fault Tolerance in Distributed Networks

5.4.1. Linear Space Bootstrap Communication Schemes

Participants: Carole Delporte-Gallet, Hugues Fauconnier, Eli Gafni, Sergio Rajsbaum.

We consider in [18], a system of n processes with ids not a priori known, that are drawn from a large space, potentially unbounded. How can these n processes communicate to solve a task? We show that n a priori allocated Multi-Writer Multi-Reader (MWMR) registers are both needed and sufficient to solve any read-write wait free solvable task. This contrasts with the existing possible solution borrowed from adaptive algorithms that require $\Theta(n^2)$ MWMR registers. To obtain these results, the paper shows how the processes can non blocking emulate a system of n Single-Writer Multi-Reader (SWMR) registers on top of n MWMR registers. It is impossible to do such an emulation with $n-1$ MWMR registers. Furthermore, we want to solve a sequence of tasks (potentially infinite) that are sequentially dependent (processes need the previous task's outputs in order to proceed to the next task). A non blocking emulation might starve a process forever. By doubling the space complexity, using $2n-1$ rather than just n registers, the computation is wait free rather than non blocking.

5.4.2. Black Art: Obstruction-Free k-set Agreement with |MWMR registers| < |proccesses|

Participants: Carole Delporte-Gallet, Hugues Fauconnier, Eli Gafni, Sergio Rajsbaum.

When n processes communicate by writing to and reading from $k < n$ MWMR registers the "communication bandwidth" precludes emulation of SWMR system, even non-blocking.

Nevertheless, recently a positive result was shown that such a system either wait-free or obstruction-free can solve an interesting one-shot task. This paper demonstrates another such result. It shows that $(n-1)-$set agreement can be solved obstruction-free with merely 2 MWMR registers. Achieving $k-$set agreement with $n-k+1$ registers is a challenge. In [17], we make the first step toward it by showing $k-$set agreement with $2(n-k)$ registers.

5.4.3. Adaptive Register Allocation with a Linear Number of Registers

Participants: Carole Delporte-Gallet, Hugues Fauconnier, Eli Gafni, Leslie Lamport.

In [16], we give an adaptive algorithm in which processes use multi-writer multi- reader registers to acquire exclusive write access to their own single-writer, multi-reader registers. It is the first such algorithm that uses a number of registers linear in the number of participating processes. Previous adaptive algorithms require at least $\Theta(n^{3/2})$ registers

5.4.4. Uniform Consensus with Homonyms and Omission Failures

Participants: Carole Delporte-Gallet, Hugues Fauconnier, Hung Tran-The.

In synchronous message passing models in which some processes may be homonyms, i.e. may share the same id, we consider the consensus problem. Many results have already been proved concerning Byzantine failures in models with homonyms, we complete in [19], the picture with crash and omission failures.

Let n be the number of processes, t the number of processes that may be faulty ($t < n$) and $l(1 \leq l \leq n)$ the number of identifiers. We prove that for crash failures and send-omission failures, uniform consensus is solvable even if $l = 1$, that is with fully anonymous processes for any number of faulty processes.

Concerning omission failures, when the processes are numerate, i.e. are able to count the number of copies of identical messages they received in each round, uniform consensus is solvable even for fully anonymous processes for $n > 2t$. If processes are not numerate, uniform consensus is solvable if and only if $l > 2t$.

All the proposed protocols are optimal both in the number of communication steps needed, and in the number of processes that can be faulty.

All these results show, (1) that identifiers are not useful for crash and send-omission failures or when processes are numerate, (2) for general omission or for Byzantine failures the number of different ids becomes significant.

5.4.5. Byzantine agreement with homonyms

Participants: Carole Delporte-Gallet, Hugues Fauconnier, Rachid Guerraoui, Anne-Marie Kermarrec, Hung Tran-The.

So far, the distributed computing community has either as- sumed that all the processes of a distributed system have distinct identifiers or, more rarely, that the processes are anonymous and have no identifiers. These are two extremes of the same general model: namely, n processes use l dif- ferent authenticated identifiers, where $1 \leq l \leq n$. In this paper [3], we ask how many identifiers are actually needed to reach agreement in a distributed system with t Byzantine processes. We show that having $3t+1$ identifiers is necessary and suf- ficient for agreement in the synchronous case but, more sur- prisingly, the number of identifiers must be greater than $(n+3t)/2$ in the partially synchronous case. This demonstrates two differences from the classical model (which has l = n): there are situations where relaxing synchrony to partial syn- chrony renders agreement impossible; and, in the partially synchronous case, increasing the number of correct processes can actually make it harder to reach agreement. The im- possibility proofs use the fact that a Byzantine process can send multiple messages to the same recipient in a round. We show that removing this ability makes agreement easier: then, $t+1$ identifiers are sufficient for agreement, even in the partially synchronous model.

5.4.6. Byzantine agreement with homonyms in synchronous systems

Participants: Carole Delporte-Gallet, Hugues Fauconnier, Hung Tran-The.

We consider in [4], the Byzantine agreement problem 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 $l \frac{(n-r)t}{n-t-\min(t,r)}$ 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.

5.4.7. Convergence of the D-iteration algorithm: convergence rate and asynchronous distributed scheme

Participants: Dohy Hong, Fabien Mathieu, Gérard Burnside.

In this paper [25], we define the general framework to describe the diffusion operators associated to a positive matrix. We define the equations associated to diffusion operators and present some general properties of their state vectors. We show how this can be applied to prove and improve the convergence of a fixed point problem associated to the matrix iteration scheme, including for distributed computation framework. The approach can be understood as a decomposition of the matrix-vector product operation in elementary operations at the vector entry level.

5.5. Discrete Optimization Algorithms

5.5.1. Shrinking Maxima, Decreasing Costs: New Online Packing and Covering Problems

Participants: Pierre Fraigniaud, Magnús M. Halldórsson, Boaz Patt-Shamir, Dror Rawitz, Adi Rosén.

We consider in [23], two new variants of online integer programs that are duals. In the packing problem we are given a set of items and a collection of knapsack constraints over these items that are revealed over time in an online fashion. Upon arrival of a constraint we may need to remove several items (irrevocably) so as to maintain feasibility of the solution. Hence, the set of packed items becomes smaller over time. The goal is to maximize the number, or value, of packed items. The problem originates from a buffer-overflow model in communication networks, where items represent information units broken into multiple packets. The other problem considered is online covering: There is a universe to be covered. Sets arrive online, and we must decide for each set whether we add it to the cover or give it up. The cost of a solution is the total cost of sets taken, plus a penalty for each uncovered element. The number of sets in the solution grows over time, but its cost goes down. This problem is motivated by team formation, where the universe consists of skills, and sets represent candidates we may hire. The packing problem was introduced for the special case where the matrix is binary; in this paper we extend the solution to general matrices with non-negative integer entries. The covering problem is introduced in this paper; we present matching upper and lower bounds on its competitive ratio.

5.5.2. Generalized Subdifferentials of the Sign Change Counting Function

Participants: Dominique Fortin, Ider Tseveendorj.

A natural generalization of piecewise linear approximation of non convex problems relies on piecewise convex approximation; along the way to solve the piecewise convex maximization problem [30] both effectively and efficiently, optimality conditions have to be addressed in two ways: either the violation of necessary conditions should lead to a direction of improvement from a local solution, or a sufficient condition for global optimality has to be fulfilled. The way to either goal is paved with subdifferentials and their generalizations on a per problem basis.

In the article [29], the counting function on binary values is extended to the signed case in order to count the number of transitions between contiguous locations. A generalized subdifferential for the sign change counting function is given where classical subdifferentials remain intractable. An attempt to prove global optimality at some point, for the 4-dimensional first non trivial example, is made by using a sufficient condition specially tailored among all the cases for this subdifferential.

GECO Project-Team

6. New Results

6.1. New results: geometric control

We start by presenting some results on motion planning and tracking algorithms.

- In [22] we study the complexity of the motion planning problem for control-affine systems. Such complexities are already defined and rather well-understood in the particular case of nonholonomic (or sub-Riemannian) systems. Our aim is to generalize these notions and results to systems with a drift. Accordingly, we present various definitions of complexity, as functions of the curve that is approximated, and of the precision of the approximation. Due to the lack of time-rescaling invariance of these systems, we consider geometric and parametrized curves separately. Then, we give some asymptotic estimates for these quantities.
- In [23] we study the problem of controlling an unmanned aerial vehicle (UAV) to provide a target supervision and to provide convoy protection to ground vehicles. We first present a control strategy based upon a Lyapunov–LaSalle stabilization method to provide supervision of a stationary target. The UAV is expected to join a pre-designed admissible circular trajectory around the target which is itself a fixed point in the space. Our strategy is presented for both HALE (High Altitude Long Endurance) and MALE (Medium Altitude Long Endurance) types UAVs. A UAV flying at a constant altitude (HALE type) is modeled as a Dubins vehicle (i.e. a planar vehicle with constrained turning radius and constant forward velocity). For a UAV that might change its altitude (MALE type), we use the general kinematic model of a rigid body evolving in $\mathbb{R}^3$. Both control strategies presented are smooth and unlike what is usually proposed in the literature these strategies asymptotically track a circular trajectory of exact minimum turning radius. We then consider the problem of adding to the tracking task an optimality criterion. In particular, we present the time-optimal control synthesis for tracking a circle by a Dubins vehicle. This optimal strategy, although much simpler than the point-to-point time-optimal strategy obtained by P. Souéres and J.-P. Laumond in the 1990s, is very rich. Finally, we propose control strategies to provide supervision of a moving target, that are based upon the previous ones.
- In [26] we prove the continuity and the Hölder equivalence w.r.t. an Euclidean distance of the value function associated with the L^1 cost of the control-affine system $\dot{q} = f_0(q) + \sum_{j=1}^m u_j f_j(q)$, satisfying the strong Hörmander condition. This is done by proving a result in the same spirit as the Ball-Box theorem for driftless (or sub-Riemannian) systems. The techniques used are based on a reduction of the control-affine system to a linear but time-dependent one, for which we are able to define a generalization of the nilpotent approximation and through which we derive estimates for the shape of the reachable sets. Finally, we also prove the continuity of the value function associated with the L^1 cost of time-dependent systems of the form $\dot{q} = \sum_{j=1}^m u_j f_j^t(q)$.

Let us list some new results in sub-Riemannian geometry and hypoelliptic diffusion.

- In [1] we provide normal forms for 2D almost-Riemannian structures, which are generalized Riemannian structures on surfaces for which a local orthonormal frame is given by a Lie bracket generating pair of vector fields that can become collinear. Generically, there are three types of points: Riemannian points where the two vector fields are linearly independent, Grushin points where the two vector fields are collinear but their Lie bracket is not, and tangency points where the two vector fields and their Lie bracket are collinear and the missing direction is obtained with one more bracket. We consider the problem of finding normal forms and functional invariants at each type of point. We also require that functional invariants are "complete" in the sense that they permit to recognize locally isometric structures. The problem happens to be equivalent to the one of finding a smooth canonical parameterized curve passing through the point and being transversal to the distribution.

For Riemannian points such that the gradient of the Gaussian curvature K is different from zero, we use the level set of K as support of the parameterized curve. For Riemannian points such that the gradient of the curvature vanishes (and under additional generic conditions), we use a curve which is found by looking for crests and valleys of the curvature. For Grushin points we use the set where the vector fields are parallel. Tangency points are the most complicated to deal with. The cut locus from the tangency point is not a good candidate as canonical parameterized curve since it is known to be non-smooth. Thus, we analyze the cut locus from the singular set and we prove that it is not smooth either. A good candidate appears to be a curve which is found by looking for crests and valleys of the Gaussian curvature. We prove that the support of such a curve is uniquely determined and has a canonical parametrization.

- The curvature discussed in [14] is a rather far going generalization of the Riemann sectional curvature. We define it for a wide class of optimal control problems: a unified framework including geometric structures such as Riemannian, sub-Riemannian, Finsler and sub-Finsler structures; a special attention is paid to the sub-Riemannian (or Carnot–Caratheodory) metric spaces. Our construction of the curvature is direct and naive, and it is similar to the original approach of Riemann. Surprisingly, it works in a very general setting and, in particular, for all sub-Riemannian spaces.
- In [15] we provide the small-time heat kernel asymptotics at the cut locus in three relevant cases: generic low-dimensional Riemannian manifolds, generic 3D contact sub-Riemannian manifolds (close to the starting point) and generic 4D quasi-contact sub-Riemannian manifolds (close to a generic starting point). As a byproduct, we show that, for generic low-dimensional Riemannian manifolds, the only singularities of the exponential map, as a Lagragian map, that can arise along a minimizing geodesic are A_3 and A_5 (in the classification of Arnol'd's school). We show that in the non-generic case, a cornucopia of asymptotics can occur, even for Riemannian surfaces.
- In [19] we study the evolution of the heat and of a free quantum particle (described by the Schroedinger equation) on two-dimensional manifolds endowed with the degenerate Riemannian metric $ds^2 = dx^2 + |x|^{-2\alpha} d\theta^2$, where $x \in \mathbb{R}$, $\theta \in \mathbb{T}$ and the parameter $\alpha \in \mathbb{R}$. For $\alpha \leq -1$ this metric describes cone-like manifolds (for $\alpha = -1$ it is a flat cone). For $\alpha = 0$ it is a cylinder. For $\alpha \geq 1$ it is a Grushin-like metric. We show that the Laplace–Beltrami operator Δ is essentially self-adjoint if and only if $\alpha \notin (-3, 1)$. In this case the only self-adjoint extension is the Friedrichs extension Δ_F, that does not allow communication through the singular set $\{x = 0\}$ both for the heat and for a quantum particle. For $\alpha \in (-3, -1]$ we show that for the Schroedinger equation only the average on θ of the wave function can cross the singular set, while the solutions of the only Markovian extension of the heat equation (which indeed is Δ_F) cannot. For $\alpha \in (-1, 1)$ we prove that there exists a canonical self-adjoint extension Δ_B, called bridging extension, which is Markovian and allows the complete communication through the singularity (both of the heat and of a quantum particle). Also, we study the stochastic completeness (i.e., conservation of the L^1 norm for the heat equation) of the Markovian extensions Δ_F and Δ_B, proving that Δ_F is stochastically complete at the singularity if and only if $\alpha \leq -1$, while Δ_B is always stochastically complete at the singularity.

6.2. New results: quantum control

New results have been obtained for the control of the bilinear Schrödinger equation.

- In [4] we show the approximate rotational controllability of a polar linear molecule by means of three nonresonant linear polarized laser fields. The result is based on a general approximate controllability result for the bilinear Schroedinger equation, with wavefunction varying in the unit sphere of an infinite-dimensional Hilbert space and with several control potentials, under the assumption that the internal Hamiltonian has discrete spectrum. A further general results, extending the above approach, are obtained in [16].
- In [5] we provide a short introduction to modern issues in the control of infinite dimensional closed quantum systems, driven by the bilinear Schroedinger equation. The first part is a quick presentation

of some of the numerous recent developments in the fields. This short summary is intended to demonstrate the variety of tools and approaches used by various teams in the last decade. In a second part, we present four examples of bilinear closed quantum systems. These examples were extensively studied and may be used as a convenient and efficient test bench for new conjectures. Finally, we list some open questions, both of theoretical and practical interest.

- In [6] we study the so-called spin-boson system, namely a spin-1/2 particle in interaction with a distinguished mode of a quantized bosonic field. We control the system via an external field acting on the bosonic part. Applying geometric control techniques to the Galerkin approximation and using perturbation theory to guarantee non-resonance of the spectrum of the drift operator, we prove approximate controllability of the system, for almost every value of the interaction parameter.
- In [9] and [25] we investigate the controllability of a quantum electron trapped in a two-dimensional device. The problem is modeled by the Schroedinger equation in a bounded domain coupled to the Poisson equation for the electrical potential. The controller acts on the system through the boundary condition on the potential, on a part of the boundary modeling the gate. We prove that, generically with respect to the shape and boundary conditions on the gate, the device is controllable. In [25] We also consider control properties of a more realistic nonlinear version of the device, taking into account the self-consistent electrostatic Poisson potential.
- In [18] we study the controllability of a closed control-affine quantum system driven by two or more external fields. We provide a sufficient condition for controllability in terms of existence of conical intersections between eigenvalues of the Hamiltonian in dependence of the controls seen as parameters. Such spectral condition is structurally stable in the case of three controls or in the case of two controls when the Hamiltonian is real. The spectral condition appears naturally in the adiabatic control framework and yields approximate controllability in the infinite-dimensional case. In the finite-dimensional case it implies that the system is Lie-bracket generating when lifted to the group of unitary transformations, and in particular that it is exactly controllable. Hence, Lie algebraic conditions are deduced from purely spectral properties. Another contribution of [18] is the proof that approximate and exact controllability are equivalent properties for general finite-dimensional quantum systems.

6.3. New results: neurophysiology

- In recent papers models of the human locomotion by means of an optimal control problem have been proposed. In this paradigm, the trajectories are assumed to be solutions of an optimal control problem whose cost has to be determined. The purpose of [3] is to analyze the class of optimal control problems defined in this way. We prove strong convergence result for their solutions on the one hand for perturbations of the initial and final points (stability), and on the other hand for perturbations of the cost (robustness).
- [8] analyses a class of optimal control problems on geometric paths of the euclidean space, that is, curves parametrized by arc length. In the first part we deal with existence and robustness issues for such problems and we define the associated inverse optimal control problem. In the second part we discuss the inverse optimal control problem in the special case of planar trajectories and under additional assumptions. More precisely we define a criterion to restrict the study to a convenient class of costs based on the analysis of experimentally recorded trajectories. This method applies in particular to the case of human locomotion trajectories.
- The article [17] presents an algorithm implementing the theory of neurogeometry of vision, described by Jean Petitot in his book. We propose a new ingredient, namely working on the group of translations and discrete rotations $\mathrm{SE}(2,N)$. We focus on the theoretical and numerical aspects of integration of an hypoelliptic diffusion equation on this group. Our main tool is the generalized Fourier transform. We provide a complete numerical algorithm, fully parallellizable.

6.4. New results: switched systems

- In [2] we study the control system $\dot{x} = Ax + \alpha(t)bu$ where the pair (A, b) is controllable, $x \in \mathbb{R}^2$, $u \in \mathbb{R}$ is a scalar control and the unknown signal $\alpha : \mathbb{R}_+ \to [0, 1]$ is (T, μ)-persistently exciting (PE), i.e., there exists $T \geq \mu > 0$ such that, for all $t \in \mathbb{R}_+$, $\int_t^{t+T} \alpha(s)ds \geq \mu$. We are interested in the stabilization problem of this system by a linear state feedback $u = -Kx$. In [2], we positively answer a question asked in [52] and prove the following: Assume that the class of (T, μ)-PE signals is restricted to those which are M-Lipschitzian, where $M > 0$ is a positive constant. Then, given any $C > 0$, there exists a linear state feedback $u = -Kx$ where K only depends on (A, b) and T, μ, M so that, for every M-Lipschitzian (T, μ)-PE signal, the rate of exponential decay of the time-varying system $\dot{x} = (A - \alpha(t)bK)x$ is greater than C.

  In [20] we consider a family of linear control systems $\dot{x} = Ax + \alpha Bu$ where α belongs to a given class of persistently exciting signals. We seek maximal α-uniform stabilisation and destabilisation by means of linear feedbacks $u = Kx$. We extend previous results obtained for bidimensional single-input linear control systems to the general case as follows: if the pair (A, B) verifies a certain Lie bracket generating condition, then the maximal rate of convergence of (A, B) is equal to the maximal rate of divergence of $(-A, -B)$. We also provide more precise results in the general single-input case, where the above result is obtained under the sole assumption of controllability of the pair (A, B).

  The paper [24] considers the stabilization to the origin of a persistently excited linear system by means of a linear state feedback, where we suppose that the feedback law is not applied instantaneously, but after a certain positive delay (not necessarily constant). The main result is that, under certain spectral hypotheses on the linear system, stabilization by means of a linear delayed feedback is indeed possible, generalizing a previous result already known for non-delayed feedback laws.

  Several problems and results related with persistent excitation and stabilization are discussed in the survey [11]. These problems and results deal with both finite- and infinite-dimensional systems.

- In [7] we consider several time-discretization algorithms for singularly perturbed switched systems. The algorithms correspond to different sampling times and the discretization procedure respects the splitting of each mode in fast and slow dynamics. We study whether such algorithms preserve the asymptotic or quadratic stability of the original continuous-time singularly perturbed switched system.

- In [10] we consider affine switched systems as perturbations of linear ones, the equilibria playing the role of perturbation parameters. We study the stability properties of an affine switched system under arbitrary switching, assuming that the corresponding linear system is uniformly exponentially stable. It turns out that the affine system admits a minimal invariant set Ω, whose properties we investigate. In the two-dimensional bi-switched case when both subsystems have non-real eigenvalues we are able to characterize Ω completely and to prove that all trajectories of the system converge to Ω. We also explore the behavior of minimal-time trajectories in Ω by constructing optimal syntheses.

- In [21] we give a collection of converse Lyapunov–Krasovskii theorems for uncertain retarded differential equations. We show that the existence of a weakly degenerate Lyapunov–Krasovskii functional is a necessary and sufficient condition for the global exponential stability of the linear retarded functional differential equations. This is carried out using the switched system transformation approach.

GENSCALE Project-Team

6. New Results

6.1. NGS methodology

Participants: Dominique Lavenier, Claire Lemaitre, Pierre Peterlongo, Guillaume Rizk, Anaïs Gouin, Fabrice Legeai.

- **Efficient Kmer counting**: Counting all the substrings of length k (k-mers) in DNA/RNA sequencing reads is the preliminary step of many bioinformatics applications. However, state of the art k-mer counting methods require that a large data structure resides in memory. Such structure typically grows with the number of distinct k-mers to count. We have developed a new streaming algorithm for that purpose which only requires a fixed user-defined amount of memory and disk space. This approach realizes a memory, time and disk trade-off. DSK is the first approach that is able to count all the 27-mers of a human genome dataset using only 4.0 GB of memory and moderate disk space (160 GB), in 17.9 h. DSK can replace a popular k-mer counting software (Jellyfish) on small-memory servers. [24]
- **Questionning the classical re-sequencing analyses approach**: Classical re-sequencing analyses are based on a first step of read mapping, then only mapped reads are taken into account in following analyses such as variant calling. We investigated the sources of unmapped reads in aphid re-sequencing data of 33 individuals, and we demonstrated that these reads contain valuable information that should not be discarded as usually done in such analyses. We proposed also an approach to extract this information, based on assembly and re-mapping. [34]
- **Repeat detection** A new algorithm was developed for detecting long similar fragments occurring at least twice in a set of biological sequences. The problem becomes computationally challenging when the frequency of a repeat is allowed to increase and when a non-negligible number of insertions, deletions and substitutions are allowed. The proposed algorithm, called Rime (for Repeat Identification: long, Multiple, and with Edits) performs this task, and manages instances whose size and combination of parameters cannot be handled by other currently existing methods. To the best of our knowledge, Rime is the first algorithm that can accurately deal with very long repeats (up to a few thousands), occurring possibly several times, and with a rate of differences (substitutions and indels) allowed among copies of a same repeat of 10-15% or even more. [17]

6.2. NGS applications

Participants: Dominique Lavenier, Claire Lemaitre, Pierre Peterlongo, Guillaume Rizk, Fabrice Legeai.

- **Participation to an international competition of assembly**: The process of generating raw genome sequence data continues to become cheaper, faster, and more accurate. However, assembly of such data into high-quality finished genome sequences remains challenging. Many genome assembly tools are available, but they differ greatly in terms of their performance and in their final output. More importantly, it remains largely unclear how to best assess the quality of assembled genome sequences. In this context, we have participated to the Assemblathon-2 competitions, which purpose was to assess current state-of-the-art methods in genome assembly. Globally, the cumulative z-scores of different assembly criteria set our assembly strategy in the 4th position compared to other competitors (21 groups). [12]
- **Assembly on Raspberri Pi**: Current Assembly tools require computers with large memory configuration. In order to demonstrate the efficiency of our low memory footprint assembly tools, we assemble the genome of C. Elegans (100 Mbp) on the raspberry PI computer, a small system equipped with only 512 MB RAM and 32 GB flash drive. [42]

- **SNP detection on the tick** We took part of a population genetic study on the tick species *Ixodes ricinus*, the main vector species of human and animal vector-borne diseases in Europe. In this framework, we proposed the first identification of a set of SNPs isolated from the genome of *I. ricinus*, by applying, among others a new tool developed in the GenScale team: discoSnp. The main advantage of this tool is to be able to detect SNPs without the use of a reference genome, which is crucially lacking for the tick species. Among the detected SNPs, 384 were selected, according to their minimal and maximal coverage and context sequences for experimental validation. Among them, 368 (95.8%) were biologically validated, demonstrating the precision of discoSNP.[23]
- **NGS analyses on insect models** We achieved the transcriptome assembly and analyzed the differential expression of an important noctuid pest. [22], [18]. Using gene expression data (RNA-Seq) in males, sexual females and asexual females of the pea aphid, we confirm theoretical models suggesting that the evolution of sex-biased gene expression may restrict the product of a sexually antagonistic allele to the sex it benefits.[19]
- **Genome sequencing and annotation**: We participated in the sequencing and annotation of several bacterial species of the Mollicute group. These bacteria are important pathogens of ruminants. The sequencing and annotation of their genomes confirmed their pathogenic features and phylogenetic location in the tree of Mollicutes. This is the first step before comparative genome analyses to unravel the genetic basis of mycoplasma pathogenicity and host specificity. [15], [16], [21]

6.3. HPC and parallelism

Participants: Dominique Lavenier, Rumen Andonov, Guillaume Chapuis, François Moreews, Charles Deltel.

- **Improving time performances of Mapping quantitative trait loci (QTL)** : we have developed a fast implementation of QTLMap, which takes advantage of the data parallel nature of the problem by offsetting heavy computations to a graphics processing unit (GPU). This new implementation performs up to 75 times faster than the previous multicore implementation, while maintaining the same results and level of precision . This speedup allows one to perform more complex analyses, such as linkage disequilibrium linkage analyses (LDLA) and multiQTL analyses, in a reasonable time frame. [13]
- **Integration of parallelism in bioinformatics workflows**: We propose a Model-Driven Architecture approach for capturing the complete design process of bioinformatics workflows. This approach is applied to graphical workflow editors and allows to quickly convert a workflow prototype in a parallel implementation. This work can have an impact on the way bioinformaticians implement their analysis and increase their productivity.[30]
- **Parallel assembly on FPGAs**: This research work proposes a method to reduce the overall time for assembly by using pre-processing of the short read data on FPGAs and processing its output using Velvet. We demonstrate significant speed-ups with slight or no compromise on the quality of the assembled output.[32]
- **All-Pairs Shortest Paths with multi-GPU** We propose a new algorithm for the All-Pairs Shortest Paths problem for graphs with good partitioning properties and its multi-GPU implementation. Our implementation targets large graphs (up to 10^6 vertices) and allows graphs with negative edges to be computed. [35]

6.4. Protein structures

Participants: Rumen Andonov, Guillaume Chapuis, Dominique Lavenier, Mathilde Le Boudic-Jamin, Antonio Mucherino, Douglas Goncalves.

- **A book on distance geometry problems (DGP).** This is a collection of invited papers on the topic "distance geometry" [38]. Among the other contributions, it contains a survey on "distance geometry" and "structural biology", which tries to function as a bridge between two scientific communities: computer science and biology. It presents some recent developments in the field by using a language common to the two communities [37]. In another contribution, the complexity of the DGP is discussed: even if this problem is NP-hard in general, we noticed a polynomial complexity on instances of DGP related to protein conformations (in the case all the available distances are exact)[36].

- **DGP with interval data.** In our preliminary works on the discretization of the Distance Geometry Problem (DGP), we considered instances where all distances were supposed to be exactly known. When biological molecules are concerned, however, this is not generally the case. We worked therefore for considering the full-atom representation of the protein backbone, where some of the distances are subject to uncertainty within a given nonnegative interval. We showed that the discretization is still possible in this case, and proposed the iBP algorithm to solve the discretized DGP. [20]
- **New pruning device for DGP.** After the discretization, DGPs can be solved by a branch-and-prune (BP) algorithm, which is potentially able to enumerate the entire solution set. This solution set, however, can be very large for some instances, while only the most energetically stable conformations are of interest. We worked therefore for integrating the BP algorithm with two new energy-based pruning devices. Our computational experiments showed that the newly added pruning devices were actually able to improve the performance of the algorithm, as well as the quality (in terms of energy) of the conformations in the solution set. [28]
- **Discretization orders for the DGP.** The main assumption that allows for the discretization of DGPs is strongly based on the order in which the atoms of the molecule are considered. The "natural" order of the atoms in the amino acid chain does not always allow for the discretization. We tried to find discretization orders in several ways, based on different approaches. In [31], we extended a previously proposed greedy algorithm that is able to deal with interval data (inexact distances). In [27], we handcrafted some discretization orders for the side chains of the amino acids involved in the protein synthesis. In [29], we proposed a heuristic, which outperforms, on large instances, the greedy algorithm previously proposed.
- **DGP with Clifford Algebra** The BP algorithm for the DGP is based on a search on the tree, where nodes of the tree belonging to a common layer provide the possible positions for the same atom of the molecule. When interval data are given, a curve in 3d (containing the possible positions for the atom) can be associated to one of such nodes. Since it is generally not necessary to have protein conformations with a precision higher than 1A, sample points on these curves can be chosen. The way to choose these sample points is not, however, a simple task. This is the reason why we are trying to make this selection process adaptive, by exploiting Clifford Algebra to this purpose. Preliminary studies in this direction were presented in [25]
- **Parallel seed-based approach to protein structure similarity detection** We have developed a new parallel heuristic-based approach to structural similarity detection between proteins that discovers multiple pairs of similar regions. We prove that returned alignments have RMSDc and RMSDd lower than a given threshold. Computational complexity is addressed by taking advantage of both fine- and coarse-grain parallelism. [26]
- **Datamining.** The selection of features that describe samples in sets of data is a typical problem in data mining. A crucial issue is to select a maximal set of pertinent features, because the scarce knowledge of the problem under study often leads to consider features which do not provide a good description of the corresponding samples. The concept of consistent biclustering of a set of data has been introduced to identify such a maximal set. The problem can be modeled as a 0–1 linear fractional program, which is NP-hard. We reformulated this optimization problem as a bilevel program, and we proposed a heuristic for its solution [39].

GEOMETRICA Project-Team

6. New Results

6.1. Mesh Generation and Geometry Processing

6.1.1. Splat-based Surface Reconstruction from Defect-Laden Point Sets.

Participant: Mariette Yvinec.

In collaboration with Pierre Alliez (EPI Titane), Ricard Campos (University of Girona), Raphael Garcia (University of Girona)

We introduce a method for surface reconstruction from point sets that is able to cope with noise and outliers. First, a splat-based representation is computed from the point set. A robust local 3D RANSAC-based procedure is used to filter the point set for outliers, then a local jet surface – a low-degree surface approximation – is fitted to the inliers. Second, we extract the reconstructed surface in the form of a surface triangle mesh through Delaunay refinement. The Delaunay refinement meshing approach requires computing intersections between line segment queries and the surface to be meshed. In the present case, intersection queries are solved from the set of splats through a 1D RANSAC procedure. [14].

6.1.2. Constructing Intrinsic Delaunay Triangulations of Submanifolds

Participants: Jean-Daniel Boissonnat, Ramsay Dyer.

In collaboration with Arijit Ghosh (Indian Statistical Institute)

We describe an algorithm to construct an intrinsic Delaunay triangulation of a smooth closed submanifold of Euclidean space [42]. Using results established in a companion paper on the stability of Delaunay triangulations on δ-generic point sets, we establish sampling criteria which ensure that the intrinsic Delaunay complex coincides with the restricted Delaunay complex and also with the recently introduced tangential Delaunay complex. The algorithm generates a point set that meets the required criteria while the tangential complex is being constructed. In this way the computation of geodesic distances is avoided, the runtime is only linearly dependent on the ambient dimension, and the Delaunay complexes are guaranteed to be triangulations of the manifold.

6.1.3. Delaunay Triangulation of Manifolds

Participants: Jean-Daniel Boissonnat, Ramsay Dyer.

In collaboration with Arijit Ghosh (Indian Statistical Institute)

We present an algorithmic framework for producing Delaunay triangulations of manifolds [44]. The input to the algorithm is a set of sample points together with coordinate patches indexed by those points. The transition functions between nearby coordinate patches are required to be bi-Lipschitz with a constant close to 1. The primary novelty of the framework is that it can accommodate abstract manifolds that are not presented as submanifolds of Euclidean space. The output is a manifold simplicial complex that is the Delaunay complex of a perturbed set of points on the manifold. The guarantee of a manifold output complex demands no smoothness requirement on the transition functions, beyond the bi-Lipschitz constraint. In the smooth setting, when the transition functions are defined by common coordinate charts, such as the exponential map on a Riemannian manifold, the output manifold is homeomorphic to the original manifold, when the sampling is sufficiently dense.

6.1.4. Anisotropic Delaunay Meshes of Surfaces

Participants: Jean-Daniel Boissonnat, Mariette Yvinec.

In collaboration with Jane Tournois (GeometryFactory) and Kan-Le Shi (Tsing Hua University)

Anisotropic simplicial meshes are triangulations with elements elongated along prescribed directions. Anisotropic meshes have been shown to be well suited for interpolation of functions or solving PDEs. They can also significantly enhance the accuracy of a surface representation. Given a surface S endowed with a metric tensor field, we propose a new approach to generate an anisotropic mesh that approximates S with elements shaped according to the metric field [13], [47]. The algorithm relies on the well-established concepts of restricted Delaunay triangulation and Delaunay refinement and comes with theoretical guarantees. The star of each vertex in the output mesh is Delaunay for the metric attached to this vertex. Each facet has a good aspect ratio with respect to the metric specified at any of its vertices. The algorithm is easy to implement. It can mesh various types of surfaces like implicit surfaces, polyhedra or isosurfaces in 3D images. It can handle complicated geometries and topologies, and very anisotropic metric fields.

6.2. Topological and Geometric Inference

6.2.1. An Efficient Data Structure for Computing Persistent Cohomology

Participants: Jean-Daniel Boissonnat, Clément Maria.

In collaboration with Tamal Dey (Ohio State University)

Persistent homology with coefficients in a field F coincides with the same for cohomology because of duality. We propose an implementation of a recently introduced algorithm for persistent cohomology that attaches annotation vectors with the simplices. We separate the representation of the simplicial complex from the representation of the cohomology groups, and introduce a new data structure for maintaining the annotation matrix, which is more compact and reduces substancially the amount of matrix operations. In addition, we propose a heuristic to further simplify the representation of the cohomology groups and improve both time and space complexities. The paper provides a theoretical analysis, as well as a detailed experimental study of our implementation and comparison with state-of-the-art software for persistent homology and cohomology [41], [29].

6.2.2. Multi-Field Persistent Homology

Participants: Jean-Daniel Boissonnat, Clément Maria.

In [46], we introduce the *multi-field persistence diagram* for the persistence homology of a filtered complex. It encodes compactly the *superimposition* of the persistence diagrams of the complex with several field coefficients, and provides a substantially more precise description of the topology of the filtered complex. Specifically, the multi-field persistence diagram encodes the Betti numbers of integral homology and the prime divisors of the torsion coefficients of the underlying shape. Moreover, it enjoys similar stability properties as the ones of standard persistence diagrams, with the appropriate notion of distance. These properties make the multi-field persistence diagram a useful tool in computational topology. The multi-field algorithms are, in practice, as fast as algorithms that compute persistent homology in a single field.

6.2.3. Zigzag Zoology: Rips Zigzags for Homology Inference

Participants: Steve Oudot, Donald Sheehy.

For points sampled near a compact set X, the persistence barcode of the Rips filtration built from the sample contains information about the homology of X as long as X satisfies some geometric assumptions. The Rips filtration is prohibitively large, however zigzag persistence can be used to keep the size linear. We present several species of Rips-like zigzags and compare them with respect to the signal-to-noise ratio, a measure of how well the underlying homology is represented in the persistence barcode relative to the noise in the barcode at the relevant scales. Some of these Rips-like zigzags have been available as part of the Dionysus library for several years while others are new. Interestingly, we show that some species of Rips zigzags will exhibit less noise than the (non-zigzag) Rips filtration itself. Thus, Rips zigzags can offer improvements in both size complexity and signal-to-noise ratio. Along the way, we develop new techniques for manipulating and comparing persistence barcodes from zigzag modules. In particular, we give methods for reversing arrows and removing spaces from a zigzag while controlling the changes occurring in its barcode. We also discuss factoring zigzags and a kind of interleaving of two zigzags that allows their barcodes to be compared. These techniques were developed to provide our theoretical analysis of the signal-to-noise ratio of Rips-like zigzags, but they are of independent interest as they apply to zigzag modules generally [33].

6.2.4. Efficient and Robust Topological Data Analysis on Metric Spaces

Participants: Mickaël Buchet, Frédéric Chazal, Steve Oudot, Donald Sheehy.

We extend the notion of the distance to a measure from Euclidean space to probability measures on general metric spaces as a way to perform topological data analysis in a way that is robust to noise and outliers. We then give an efficient way to approximate the sub-level sets of this function by a union of metric balls and extend previous results on sparse Rips filtrations to this setting. This robust and efficient approach to topological data analysis is illustrated with several examples from an implementation [54].

6.2.5. Noise-Adaptive Shape Reconstruction from Raw Point Sets

Participant: David Cohen-Steiner.

In collaboration with Pierre Alliez (EPI Titane), Simon Giraudot (EPI Titane)

We propose a noise-adaptive shape reconstruction method specialized to smooth, closed hypersurfaces. Our algorithm takes as input a defect-laden point set with variable noise and outliers, and comprises three main steps. First, we compute a novel type of robust distance function to the data. As a robust distance function, its sublevel-sets have the correct homotopy type when the data is a sufficiently good sample of a regular shape. The new feature is a built-in scale selection mechanism that adapts to the local noise level, under the assumption that the inferred shape is a smooth submanifold of known dimension. Second, we estimate the sign and confidence of the function at a set of seed points, based on estimated crossing parities along the edges of a uniform random graph. That component is inspired by the classical MAXCUT relaxation, except that we only require a linear solve as opposed to an eigenvector computation. Third, we compute a signed implicit function through a random walker approach with soft constraints chosen as the most confident seed points computed in previous step. The resulting pipeline is scalable and offers excellent behavior for data exhibiting variable noise levels [19].

6.2.6. Optimal Rates of Convergence for Persistence Diagrams in Topological Data Analysis

Participants: Frédéric Chazal, Marc Glisse, Bertrand Michel.

In collaboration with Catherine Labruère (Université de Bourgogne).

Computational topology has recently known an important development toward data analysis, giving birth to the field of topological data analysis. Topological persistence, or persistent homology, appears as a fundamental tool in this field. In this paper [57] (to appear in proc. ICML 2014), we study topological persistence in general metric spaces, with a statistical approach. We show that the use of persistent homology can be naturally considered in general statistical frameworks and persistence diagrams can be used as statistics with interesting convergence properties. Some numerical experiments are performed in various contexts to illustrate our results.

6.2.7. Bootstrap and Stochastic Convergence for Persistence Diagrams and Landscapes

Participant: Frédéric Chazal.

In collaboration with B. Fasy (Tulane University), F. Lecci, A. Rinaldo, A. Singh, L. Wasserman (Carnegie Mellon University).

Persistent homology probes topological properties from point clouds and functions. By looking at multiple scales simultaneously, one can record the births and deaths of topological features as the scale varies. We can summarize the persistent homology with the persistence landscape, introduced by Bubenik, which converts a diagram into a well-behaved real-valued function. We investigate the statistical properties of landscapes, such as weak convergence of the average landscapes and convergence of the bootstrap. In addition, we introduce an alternate functional summary of persistent homology, which we call the silhouette, and derive an analogous statistical theory [55].

6.2.8. Gromov-Hausdorff Approximation of Metric Spaces with Linear Structure

Participant: Frédéric Chazal.

In collaboration with S. Jian (Tsinghua University).

In many real-world applications data come as discrete metric spaces sampled around 1-dimensional filamentary structures that can be seen as metric graphs. In this paper [58] we address the metric reconstruction problem of such filamentary structures from data sampled around them. We prove that they can be approximated, with respect to the Gromov-Hausdorff distance by well-chosen Reeb graphs (and some of their variants) and we provide an efficient and easy to implement algorithm to compute such approximations in almost linear time. We illustrate the performances of our algorithm on a few synthetic and real data sets.

6.2.9. Analysis and Visualization of Maps Between Shapes

Participants: Frédéric Chazal, Maks Ovsjanikov.

In collaboration with L. Guibas (Stanford University), M. Ben Chen (Technion).

In this work we propose a method for analyzing and visualizing individual maps between shapes, or collections of such maps [23]. Our method is based on isolating and highlighting areas where the maps induce significant distortion of a given measure in a multi-scale way. Unlike the majority of prior work which focuses on discovering maps in the context of shape matching, our main focus is on evaluating, analyzing and visualizing a given map, and the distortion(s) it introduces, in an efficient and intuitive way. We are motivated primarily by the fact that most existing metrics for map evaluation are quadratic and expensive to compute in practice, and that current map visualization techniques are suitable primarily for global map understanding, and typically do not highlight areas where the map fails to meet certain quality criteria in a multi-scale way. We propose to address these challenges in a unified way by considering the functional representation of a map, and performing spectral analysis on this representation. In particular, we propose a simple multi-scale method for map evaluation and visualization, which provides detailed multi-scale information about the distortion induced by a map, which can be used alongside existing global visualization techniques.

6.2.10. Map-Based Exploration of Intrinsic Shape Differences and Variability

Participants: Frédéric Chazal, Maks Ovsjanikov.

In collaboration with L. Guibas and Raif Rustamov (Stanford University), M. Ben Chen and O. Azencot (Technion).

We develop a novel formulation for the notion of shape differences, aimed at providing detailed information about the location and nature of the differences or distortions between the two shapes being compared [27]. Our difference operator, derived from a shape map, is much more informative than just a scalar global shape similarity score, rendering it useful in a variety of applications where more refined shape comparisons are necessary. The approach is intrinsic and is based on a linear algebraic framework, allowing the use of many common linear algebra tools (e.g, SVD, PCA) for studying a matrix representation of the operator. Remarkably, the formulation allows us not only to localize shape differences on the shapes involved, but also to compare shape differences across pairs of shapes, and to analyze the variability in entire shape collections based on the differences between the shapes. Moreover, while we use a map or correspondence to define each shape difference, consistent correspondences between the shapes are not necessary for comparing shape differences, although they can be exploited if available. We give a number of applications of shape differences, including parameterizing the intrinsic variability in a shape collection, exploring shape collections using local variability at different scales, performing shape analogies, and aligning shape collections.

6.2.11. An operator Approach to Tangent Vector Field Processing

Participants: Frédéric Chazal, Maks Ovsjanikov.

In collaboration with M. Ben Chen and O. Azencot (Technion).

In this work [34], we introduce a novel coordinate-free method for manipulating and analyzing vector fields on discrete surfaces. Unlike the commonly used representations of a vector field as an assignment of vectors to the faces of the mesh, or as real values on edges, we argue that vector fields can also be naturally viewed as operators whose domain and range are functions defined on the mesh. Although this point of view is common in differential geometry it has so far not been adopted in geometry processing applications. We recall the theoretical properties of vector fields represented as operators, and show that composition of vector fields with other functional operators is natural in this setup. This leads to the characterization of vector field properties through commutativity with other operators such as the Laplace-Beltrami and symmetry operators, as well as to a straight-forward definition of differential properties such as the Lie derivative. Finally, we demonstrate a range of applications, such as Killing vector field design, symmetric vector field estimation and joint design on multiple surfaces.

6.3. Data Structures and Robust Geometric Computation

6.3.1. The Stability of Delaunay triangulations

Participants: Jean-Daniel Boissonnat, Ramsay Dyer.

In collaboration with Arijit Ghosh (Indian Statistical Institute)

We introduce a parametrized notion of genericity for Delaunay triangulations which, in particular, implies that the Delaunay simplices of δ-generic point sets are thick [45]. Equipped with this notion, we study the stability of Delaunay triangulations under perturbations of the metric and of the vertex positions. We quantify the magnitude of the perturbations under which the Delaunay triangulation remains unchanged.

6.3.2. Delaunay Stability via Perturbations

Participants: Jean-Daniel Boissonnat, Ramsay Dyer.

In collaboration with Arijit Ghosh (Indian Statistical Institute)

We present an algorithm that takes as input a finite point set in Euclidean space, and performs a perturbation that guarantees that the Delaunay triangulation of the resulting perturbed point set has quantifiable stability with respect to the metric and the point positions [43]. There is also a guarantee on the quality of the simplices: they cannot be too flat. The algorithm provides an alternative tool to the weighting or refinement methods to remove poorly shaped simplices in Delaunay triangulations of arbitrary dimension, but in addition it provides a guarantee of stability for the resulting triangulation.

6.3.3. Deletions in 3D Delaunay Triangulation

Participant: Olivier Devillers.

In collaboration with Kevin Buchin (Technical University Eindhoven, The Netherlands), Wolfgang Mulzer (Freie Universität Berlin, Germany), Okke Schrijvers, (Stanford University, USA) and Jonathan Shewchuk (University of California at Berkeley, USA)

Deleting a vertex in a Delaunay triangulation is much more difficult than inserting a new vertex because the information present in the triangulation before the deletion is difficult to exploit to speed up the computation of the new triangulation.

The removal of the tetrahedra incident to the deleted vertex creates a hole in the triangulation that need to be retriangulated. First we propose a technically sound framework to compute incrementally a triangulation of the hole vertices: *the conflict Delaunay triangulation*. The conflict Delaunay triangulation matches the hole boundary and avoid to compute extra tetrahedra outside the hole. Second, we propose a method that uses *guided randomized reinsertion* to speed up the point location during the computation of the conflict triangulation. The hole boundary is a polyhedron, this polyhedron is simplified by deleting its vertices one by one in a random order maintaining a polyhedron called *link Delaunay triangulation*, then the points are inserted in reverse order into the conflict Delaunay triangulation using the information from the link Delaunay triangulation to avoid point location [30].

6.3.4. A Convex Body with a Chaotic Random Polytope

Participants: Olivier Devillers, Marc Glisse, Rémy Thomasse.

Consider a sequence of points in a convex body in dimension d whose convex hull is dynamically maintained when the points are inserted one by one, the convex hull size may increase, decrease, or being constant when a new point is added. Studying the expected size of the convex hull when the points are evenly distributed in the convex is a classical problem of probabilistic geometry that yields to some surprising facts. For example, although it seems quite natural to think that the expected size of the convex hull is increasing with n the number of points, this fact is only formally proven for n big enough [16]. The asymptotic behavior of the expected size is known to be logarithmic for a polyhedral body and polynomial for a smooth one. If for a polyhedral or a smooth body, the asymptotic behavior is *somehow* “nice" it is possible to construct strange convex objects that have no such nice behaviors and we exhibit a convex body, such that the behavior of the expected size of a random polytope oscillates between the polyhedral and smooth behaviors when n increases [51].

6.3.5. Delaunay Triangulations and Cycles on Closed Hyperbolic surfaces

Participants: Mikhail Bogdanov, Monique Teillaud.

This work [40] is motivated by applications of *periodic* Delaunay triangulations in the Poincaré disk conformal model of the hyperbolic plane $\mathbb{H}^2$. A periodic triangulation is defined by an infinite point set that is the image of a finite point set by a (non commutative) discrete group G generated by hyperbolic translations, such that the hyperbolic area of a Dirichlet region is finite (i.e., a cocompact Fuchsian group acting on $\mathbb{H}^2$ without fixed points).

We consider the projection of such a Delaunay triangulation onto the closed orientable hyperbolic surface $M = \mathbb{H}^2/G$. The graph of its edges may have cycles of length one or two. We prove that there always exists a finite-sheeted covering space of M in which there is no cycle of length ≤ 2. We then focus on the group defining the Bolza surface (homeomorphic to a torus having two handles), and we explicitly construct a sequence of subgroups of finite index allowing us to exhibit a covering space of the Bolza surface in which, for any input point set, there is no cycle of length one, and another covering space in which there is no cycle of length two. We also exhibit a small point set such that the projection of the Delaunay triangulation on the Bolza surface for any superset has no cycle of length ≤ 2.

The work uses mathematical proofs, algorithmic constructions, and implementation.

6.3.6. Universal Point Sets for Planar Graph Drawings with Circular Arcs

Participant: Monique Teillaud.

In collaboration with Patrizio Angelini (Roma Tre University), David Eppstein (University of California, Irvine), Fabrizio Frati (The University of Sydney), Michael Kaufmann (MPI, Tübingen), Sylvain Lazard (EPI VEGAS), Tamara Mchedlidze (Karlsruhe Institute of Technology), and Alexander Wolff (Universität Würzburg).

We prove that there exists a set S of n points in the plane such that every n-vertex planar graph G admits a plane drawing in which every vertex of G is placed on a distinct point of S and every edge of G is drawn as a circular arc. [25]

6.3.7. A Generic Implementation of dD Combinatorial Maps in CGAL

Participant: Monique Teillaud.

In collaboration with Guillaume Damiand (Université de Lyon, LIRIS, UMR 5205 CNRS)

We present a generic implementation of dD combinatorial maps and linear cell complexes in CGAL. A combinatorial map describes an object subdivided into cells; a linear cell complex describes the linear geometry embedding of such a subdivision. In this paper [49], we show how generic programming and new techniques recently introduced in the C++11 standard allow a fully generic and customizable implementation of these two data structures, while maintaining optimal memory footprint and direct access to all information. To the best of our knowledge, the CGAL software packages presented here [59], [60] offer the only available generic implementation of combinatorial maps in any dimension.

6.3.8. Silhouette of a Random Polytope

Participant: Marc Glisse.

In collaboration with Sylvain Lazard and Marc Pouget (EPI VEGAS*) and Julien Michel (LMA-Poitiers).*

We consider random polytopes defined as the convex hull of a Poisson point process on a sphere in $\mathbb{R}^3$ such that its average number of points is n. We show [52] that the expectation over all such random polytopes of the maximum size of their silhouettes viewed from infinity is $\Theta(\sqrt{n})$.

6.3.9. A New Approach to Output-Sensitive Voronoi Diagrams and Delaunay Triangulations

Participant: Donald Sheehy.

In collaboration with Gary Miller (Carnegie Mellon University)

We describe [35] a new algorithm for computing the Voronoi diagram of a set of n points in constant-dimensional Euclidean space. The running time of our algorithm is $O(f \log n \log \Delta)$ where f is the output complexity of the Voronoi diagram and Δ is the spread of the input, the ratio of largest to smallest pairwise distances. Despite the simplicity of the algorithm and its analysis, it improves on the state of the art for all inputs with polynomial spread and near-linear output size. The key idea is to first build the Voronoi diagram of a superset of the input points using ideas from Voronoi refinement mesh generation. Then, the extra points are removed in a straightforward way that allows the total work to be bounded in terms of the output complexity, yielding the output sensitive bound. The removal only involves local flips and is inspired by kinetic data structures.

6.3.10. A Fast Algorithm for Well-Spaced Points and Approximate Delaunay Graphs

Participant: Donald Sheehy.

In collaboration with Gary Miller and Ameya Velingker (Carnegie Mellon University)

We present [32] a new algorithm that produces a well-spaced superset of points conforming to a given input set in any dimension with guaranteed optimal output size. We also provide an approximate Delaunay graph on the output points. Our algorithm runs in expected time $O(2^{O(d)}(n \log n + m))$, where n is the input size, m is the output point set size, and d is the ambient dimension. The constants only depend on the desired element quality bounds.

To gain this new efficiency, the algorithm approximately maintains the Voronoi diagram of the current set of points by storing a superset of the Delaunay neighbors of each point. By retaining quality of the Voronoi diagram and avoiding the storage of the full Voronoi diagram, a simple exponential dependence on d is obtained in the running time. Thus, if one only wants the approximate neighbors structure of a refined Delaunay mesh conforming to a set of input points, the algorithm will return a size $2^{O(d)}m$ graph in $2^{O(d)}(n \log n + m)$ expected time. If m is superlinear in n, then we can produce a hierarchically well-spaced superset of size $2^{O(d)}n$ in $2^{O(d)}n \log n$ expected time.

6.3.11. Geometric Separators and the Parabolic Lift

Participant: Donald Sheehy.

A geometric separator for a set U of n geometric objects (usually balls) is a small (sublinear in n) subset whose removal disconnects the intersection graph of U into roughly equal sized parts. These separators provide a natural way to do divide and conquer in geometric settings. A particularly nice geometric separator algorithm originally introduced by Miller and Thurston has three steps: compute a centerpoint in a space of one dimension higher than the input, compute a conformal transformation that "centers" the centerpoint, and finally, use the computed transformation to sample a sphere in the original space. The output separator is the subset of S intersecting this sphere. It is both simple and elegant. We show [36] that a change of perspective (literally) can make this algorithm even simpler by eliminating the entire middle step. By computing the centerpoint of the points lifted onto a paraboloid rather than using the stereographic map as in the original method, one can sample the desired sphere directly, without computing the conformal transformation.

GEOSTAT Project-Team

6. New Results

6.1. Nonlinear dynamics and Mild Therapeutic Hypothermia (MTH)

Participants: Binbin Xu [correspondant], Oriol Pont, Hussein Yahia, Ihu Liryc.

The neurological damage after cardiac arrest constitutes a big challenge of hospital discharge. The mild therapeutic hypothermia (MTH) (34°C - 32°C) has shown its benefit to reduce this type of damage. However, it can have many adverse effects, among which the cardiac-arrhythmia-generation-a-posteriori (CAGP) can represent up to 34%. So it's important to understand the mechanism of CAGP in order to improve this therapy. Our study with a cardiac culture in vitro showed that at 35°C the CAGP can be induced. Spiral waves, commonly considered as a sign of cardiac arrhythmia, are observed. The process of MTH can be represented by a Pitchfork bifurcation, which could explain the different ratio of arrhythmia among the adverse effects after this therapy. This nonlinear dynamics suggests that a variable speed of cooling / rewarming, especially when passing 35°C, would help to decrease the ratio of post-hypothermia arrhythmia and then improve the hospital output. See figures 5 , 6 .

../../../../projets/geostat/IMG/fig-phasespace.png

Figure 5. Phase space reconstruction for extracellular potential signals at 37°C, 35°C and 33°C.

Publications: [33], [35], [39].

6.2. Characterizing cardiac arrhythmias and their mechanisms by means of nonlinear and robust methods

Participants: Oriol Pont [correspondant], Binbin Xu, Hussein Yahia, Ihu Liryc.

../../../../projets/geostat/IMG/fig-illus-bifurcation.png

Figure 6. Illustration of Bifurcation or Trifurcation (type Pitchfork) of hypothermia effect.

Nonlinear analysis provides appropriate tools to characterize cardiac dynamics. Singularity analysis and phase-space reconstruction are physically meaningful complexity measures with minimal assumptions on the underlying models. These methods are based on effective descriptions derived from first principles, and as a consequence, parameters are robustly estimated. We have validated this approach on ECG, endocavitary catheter measures and electrocardiographic maps.

Key parameters vary infrequently and exhibit sharp transitions, which show where information concentrates and correspond to actual dynamical regime changes. Singularity exponents sift a simple fast dynamics from its slow modulation. In space domain, extreme values highlight arrhythmogenic areas. We observe a correspondence of time lag fluctuations of phase-space reconstructions with atrial fibrillation episodes in the same way as with the dynamical changes coming from singularity exponents. This characterization of information transitions could be used in the regularization of inverse-problem mapping of electrocardiographic epicardial maps. Furthermore, this opens the way for improved model-independent complexity descriptors to be used in non-invasive, automatic diagnosis support and ablation guide for electrical insulation therapy, in cases of arrhythmias such as atrial flutter and fibrillation.
Publications: [34], [29], [32], [37], [21].

6.3. Multifractal Deep Convolutional Pooling for Robust Texture Discrimination

Participants: Hicham Badri [correspondant], Hussein Yahia, Khalid Daoudi.

A robust and fast affine invariant texture classification system is presented. The new approach consists first in filtering the input image with multiple wavelet filters of different scales and orientations followed by a dual-pooling operation to increase the local invariance. The process is repeated for different wavelet sets and multiple image resolutions. This can be seen as a deep convolutional network where the outputs correspond to the pooling responses. The next step consists in extracting a robust affine invariant descriptor based on the scale invariance prior observed in natural images ; a multifractal log exponent histogram is calculated for each output node of the network. These log- histograms are combined to form the main descriptor. The final step consists in features post-processing based on the sparse wavelet coefficients prior to reduce the influence of small perturbations. For the training, we propose a combination of the generative PCA classifier with multiclass SVMs which improves classification rates. We also propose to use multi-illumination and multi-scale training ; two simple strategies to significantly boost classification results when dealing with small and homogeneous training sets. Experiments demonstrate that the proposed solution outperforms existing methods on three challenging public benchmark datasets.

Work submitted to CVPR 2014.

6.4. Nonlinear reconstruction of optical phase perturbated by atmospheric turbulence in Adaptive Optics

Participants: Suman Maji [correspondant], Hussein Yahia, Thierry Fusco.

A new approach to wavefront phase reconstruction in Adaptive Optics (AO) from the low-resolution gradient measurements provided by a wavefront sensor, using a nonlinear approach derived from the Microcanonical Multiscale Formalism (MMF). MMF comes from established concepts in statistical physics, it is naturally suited to the study of multiscale properties of complex natural signals, mainly due to the precise numerical estimate of geometrically localized critical exponents, called the singularity exponents. These exponents quantify the degree of predictability, locally, at each point of the signal domain, and they provide information on the dynamics of the associated system. We show that multiresolution analysis carried out on the singularity exponents of a high-resolution turbulent phase (obtained by model or from data) allows a propagation along the scales of the gradients in low-resolution (obtained from the wavefront sensor), to a higher resolution. We compare our results with those obtained by linear ap- proaches, which allows us to offer an innovative approach to wavefront phase reconstruction in Adaptive Optics.
Supporting grant: Conseil Régional Aquitaine project and funding *OPTAD*.
PhD thesis defended: Suman Kumar Maji, *Multiscale Methods in Signal Processing for Adaptive Optics* , University Bordeaux-1, PhD defended on November 14th, 2013, supervisor: H. Yahia [14].
Publications: [19], [20], [28], [14].

6.5. Nonlinear Speech Analysis

Participants: Vahid Khanagha [correspondant], Khalid Daoudi, Safa Mrad, Nicolas Vinuesa, Blaise Bertrac.

1. *MMF for speech analysis* : we continued our research on the adaptation and application of the MMF to speech analysis and started a research theme on pathological voice analysis. We proposed a novel a compact representation of speech which consists in reconstructing a speech signal form its most singular manifold. This leads us to build a speech waveform coder which outperforms the G.726 standard. We then used our recently developed algorithm for Glottal Closure Instants (GCI) detection to improve the performance of our sparse linear prediction method. We also used this algorithm to develop new acoustic perturbation measures for normal/pathological voice classification.
2. *Matching pursuit for speech analysis* : we first showed that the Gabor dictionary is actually more efficient than the Gammatone dictionary for speech coding using the matching pursuit (MP) algorithm. This results mitigates some famous findings on the neural coding at the human auditory nerve. Second, we shoed that one single parameter, derived from MP decomposition of speech, allow discrimination between normal and dysphonic voices with an accuracy which is significantly higher than all existing methods.

Supporting grant: Inria CORDIS.
PhD thesis defended: Vahid Khanagha *Novel Multiscale Methods for Nonlinear Speech Analysis* , University Bordeaux-1, PhD defended on January 16th, 2013, supervisors: K. Daoudi and H. Yahia [13].
Publications: [17], [27], [41], [40], [44].

6.6. Discriminative learning for Automatic speaker recognition

Participants: Khalid Daoudi [correspondant], Reda Jourani, Régine André Obrecht, Driss Aboutajdine.

We proposed a speaker identification which combines SVM and Large Margin Gaussian Mixture Models (LM-GMM) which outperforms the performance of our LM-GMM system.
Publication: [26].

6.7. Learning Multifractal Structures in Images

Participants: Hicham Badri [correspondant], Hussein Yahia, Driss Aboutajdine.

Learning dictionaries has become a powerful tool in many image processing applications. However, standard learning methods such as K-SVD and Online learning do not take into account the the structure of the patches : each patch is expressed as a linear combination of atoms of one global dictionary. We present a new dictionary learning method which takes into account the nature of each patch by performing a multifractal decomposition of the image. As a result, each fractal set will have a specific dictionary and each dictionary contains atoms of a certain singularity degree. Each patch can therefore be expressed much more efficiently compared to global dictionary learning methods. Current experiments in image denoising show that the proposed method outperforms the global dictionary learning methods.
Work in progress.

6.8. Super resolution maps of partial pressure pCO_2 between the ocean and the atmosphere

Participants: Hussein Yahia [correspondant], Véronique Garçon (laboratoire d'Etudes En Géophysique Et Océanographie Spatiales (legos)), Joël Sudre (laboratoire d'Etudes En Géophysique Et Océanographie Spatiales (legos)), Christoph Garbe (university Of Heidelberg), Christophe Maes (laboratoire d'Etudes En Géophysique Et Océanographie Spatiales (legos)), André Butz (karlsruhe Institute Of Technology (kit)), Boris Dewitte (laboratoire d'Etudes En Géophysique Et Océanographie Spatiales (legos)), Isabelle Dadou (université Paul Sabatier).

Multiresolution analysis computed on singularity exponents estimated from physical variables is used to produce submesoscale (pixel size: 4 kms) of partial pressures pCO_2 maps between the ocean and the atmosphere. Low resolution pCO_2 information coming from models and data is propagated across the scale of the specific multriresolution analysis to infer super resolution pCO_2 maps. Validation with model outputs and boat campaigns.
Supporting grant: OceanFlux project.
Publications: [30], [38].

6.9. Turbulent ocean dynamics at super resolution: validation

Participants: Hussein Yahia [correspondant], Véronique Garçon, Joël Sudre.

Synoptic determination of ocean circulation using data acquired from space, with a coherent depiction of its turbulent characteristics, from large scale ocean circulation down to super resolution of remote sensing optical sensors, remains a fundamental challenge in Oceanography. This determination has the potential of revealing all aspects of the ocean's dynamic variability on a wide range of spatio-temporal scales and will enhance our understanding of ocean-atmosphere exchanges at super resolution, as required in the present context of climate change. We show a 4-year time series of spatial super resolution (4 kms) turbulent ocean dynamics generated from satellite data using emerging nonlinear physics, low resolution dynamics and super resolution oceanic sea surface temperature data. The method at its core consists in propagating across the scales the low resolution dynamics in a multi resolution analysis computed on adimensional critical transition information. The resulting vector field is validated with Lagrangian buoy data at super resolution obtained from NASA Global Drifter Program.
A movie showing the evolution of turbulent ocean dynamics around South Africa in the Algunas current has been made with the help and support of Inria DIRCOM project (C. Blonz, P.-O. Gaumin).
Supporting grant: ICARODE project.
Publications: [30], [38].

6.10. Upwelling

Participants: Ayoub Tamim [correspondant], Khalid Daoudi, Hussein Yahia, Joël Sudre.

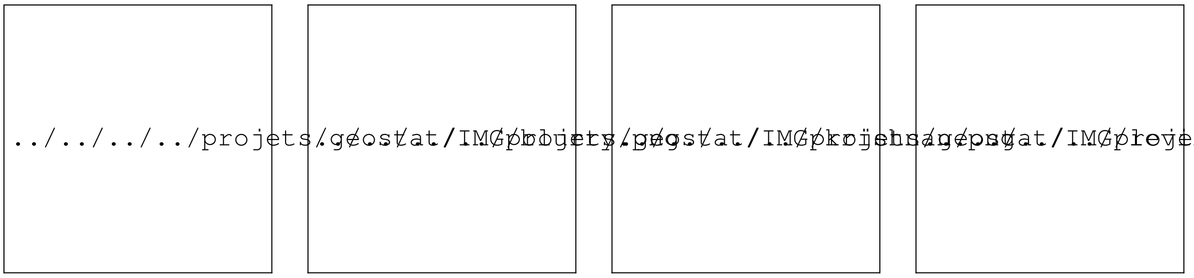
Based on fuzzy clustering, we developed a new algorithm for the segmentation of upwelling regions in the southern atlantic Morrocan coast using Sea Surface temperature images. This method has the advantage of being more efficient and more accurate than a state-of-the-art method. It is followed by a work under way of determinig oriented contours using the MMF.

Publication: [31].

6.11. Combining Local and Non-Local Priors For Image Deconvolution

Participants: Hicham Badri [correspondant], Hussein Yahia.

Non-blind deconvolution consists in recovering a sharp latent image from a blurred image with a known kernel. Deconvolved images usually contain unpleasant artifacts due to the ill-posedness of the problem even when the kernel is known. Making use of natural sparse priors has shown to reduce ringing artifacts but handling noise remains limited. On the other hand, non-local priors have shown to give the best results in image denoising. We propose in this project to combine both local and non-local priors in one framework. By studying the distribution of the singularity exponents as well as the distribution of the eigenvalues of similar patches, we show that the blur increases the self-similarity within an image and thus makes the non-local prior a good choice for denoising blurred images. The blurred image is denoised using only the self-similarties within the image, without any prior specific to the blur, via low rank estimation. However, denoising introduces outliers which are not Gaussian and should be well modeled. Experiments show that our method produces a much better image reconstruction both visually and empirically compared to some popular methods. See figure 7 . Work in progress.



(a) Blurred and Noisy, $\sigma = 2\%$ (b) Krishnan (NIPS 2009) (c) Levin (SIGGRAPH 2007) (d) Proposed

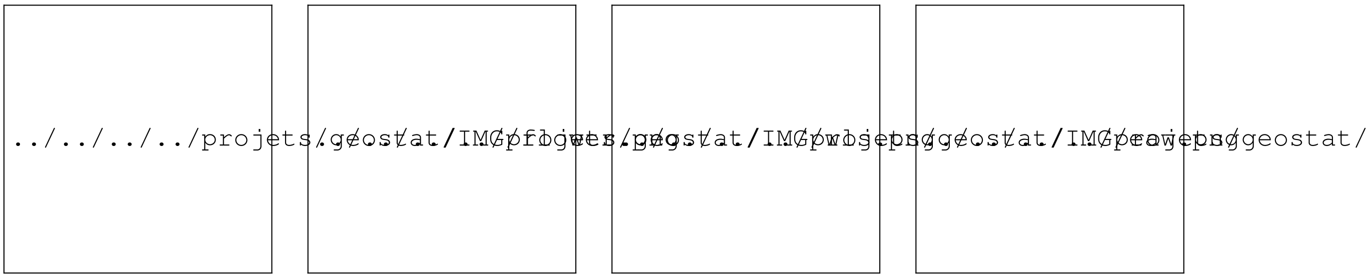
Figure 7. Various deconvolution results. The proposed method produces a much better reconstruction ; note the background noise in the methods (b) and (c). The PSNR is higher with our method 29.56 dB compared to the methods (b) 27.22 dB and (c) 28.11 dB.

6.12. Fast Multi-Scale Detail Decomposition via Accelerated Iterative Shrinkage

Participants: Hicham Badri [correspondant], Hussein Yahia, Driss Aboutajdine.

Edge-aware smoothing is one of the most important operations in computer graphics and vision. It is the building-block for a wide range of applications including : smoothing, detail manipulation, HDR tone-mapping, to cite a few. However, good quality edge-aware smoothing operators are relatively slow. We present a fast solution for performing high-quality edge-aware smoothing, particularly efficient for edge manipulation applications. Our strategy to perform smoothing consists in using a half-quadratic solver with a non-convex sparsity-inducing norm, accelerated using a first order approximation. First, we show how to solve optimization problems with complex non-convex norms using a first order proximal estimation. This step is of

paramount importance not just for smoothing, but for many applications requiring the use non-convex norms. Secondly, we design two norms inspired by natural image statistics. We incorporate these norms with a first order proximal estimation to design the main smoothing operator. Finally, we propose a warm-start solution to accelerate the solver. Experiments show that our method produces high quality results, sometimes better than some state-of-the-art methods, with reduced processing time. We demonstrate the performance of the proposed approach on various applications such as smoothing, multi-scale detail manipulation of low and high dynamic range images as well as high definition video manipulation. See figure 8 .
Work presented at SIGGRAPH Asia 2013 (technical brief), Hong Kong [24].



(a) (b) (c) (d)

Figure 8.

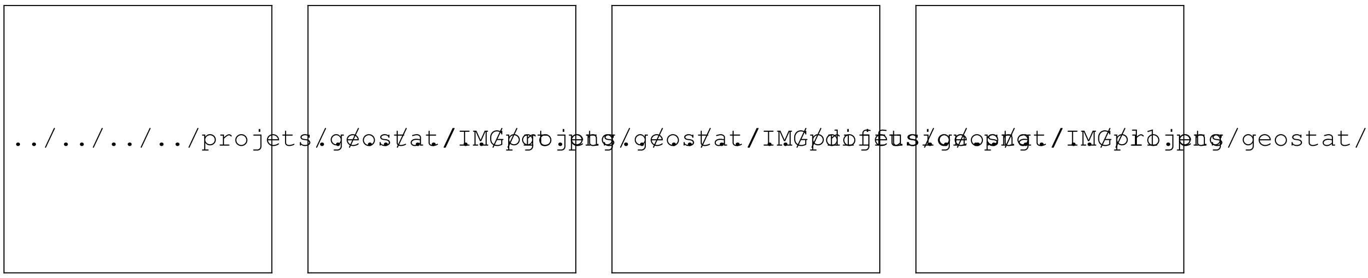
Detail manipulation example. The proposed method produces a high-quality result with reduced processing time. From left to right:(a) Input, (b) WLS, (c) EAW, (d) Proposed method.

6.13. Robust Surface Reconstruction via Triple Sparsity

Participants: Hicham Badri [correspondant], Hussein Yahia, Driss Aboutajdine.

Reconstructing a surface/image from corrupted gradient fields is a crucial step in many imaging applications where a gradient field is subject to both noise and unlocalized outliers, resulting typically in a non-integrable field. The methods presented so far can only handle a small amount of outliers and noise due to the limited performance of their models. We present in this project a powerfull method for robust surface reconstruction. The proposed formulation is based on a triple sparsity prior : a sparse prior on the residual gradient field and a double sparse prior on the surface itself. A double prior corrects the outliers in the field, while the third sparsity prior smooths the surface to reduce the noise. We develop an efficient alternate minimization strategy to solve the proposed optimization problem. The method is able to recover a good quality surface from severely corrupted gradients thanks to its ability to handle both noise and outliers. We demonstrate the performance of the proposed method on synthetic and real data. Experiments show that the proposed solution outperforms some existing methods in the three possible cases : noise only, outliers only and mixed noise/outliers. See figure 9 .

Work submitted to CVPR 2014.



(a) Ground-truth (b) Diffusion (ECCV 2006) (c) l_1-norm (CVPR09) (d) Proposed

Figure 9. Photometric stereo reconstruction from noisy images. The proposed method produces amuch better reconstruction compared to two state-of-the-art methods.

GRACE Project-Team

6. New Results

6.1. Diffusion layers for block ciphers

MDS matrices allow the construction of optimal linear diffusion layers in block ciphers. However, MDS matrices usually have a large description (for example, they can never be sparse), and this results in costly software/hardware implementations. We can solve this problem using *recursive MDS matrices*, which can be computed as a power of a simple companion matrix—and thus have a compact description suitable for constrained environments. Until now, finding recursive MDS matrices required an exhaustive search on families of companion matrices; this clearly limited the size of MDS matrices that one could look for. We have found a new direct construction, based on shortened BCH codes, which allows us to efficiently construct these matrices for arbitrary parameter sizes.

6.2. Rank metric codes over the rationals

Rank metric and Gabidulin codes over the rationals promise interesting applications to space-time coding. We have constructed optimal codes similar to Gabidulin codes but with complex coefficients, using number fields and Galois automorphisms. Using these codes we can completely bypass intermediate constructions using finite fields, which were the stumbling-block in classic constructions.

6.3. Cryptanalysis of McEliece cryptosystems based on Generalised Reed–Solomon codes

The McEliece encryption scheme based on binary Goppa codes was one of the first public-key encryption schemes [24]. Niederreiter [25] dramatically reduced the (huge) key size—a major problem with McEliece's original proposal—using Generalised Reed–Solomon (GRS) codes, but his modified scheme was broken by Sidelnikov and Shestakov [26] in 1992. There have been several attempts at repairing these smaller-key McEliece schemes. In collaboration with P. Gaborit, V. Gautier, A. Otmani and J.-P. Tillich, Alain Couvreur found polynomial time attacks on these schemes using the distinguishability of GRS codes from random codes.

6.4. New Identities relating Goppa codes

Goppa codes are strongly related to AG codes based on curves of genus 0. Among other applications, these codes are very famous for their cryptographic potential: they are one of the very few families of algebraic codes proposed for the McEliece encryption scheme which have not been broken up to now. At least for this reason, getting further knowledge on the structure of such codes is of interest. In [19], Alain Couvreur, A. Otmani and J.-P. Tillich proved a new identity yielding many improvements in the designed parameters of Goppa codes.

6.5. Root finding algorithms over local rings

Guillaume Quintin, in collaboration with J. Berthomieu and G. Lecerf, has developed new algorithms computing the roots of polynomials over complete local unramified rings [7]; this is important in the second stage of Guruswami–Sudan list decoding algorithms for codes over finite rings. Quintin has implemented these algorithms in MATHEMAGIX, using his FINITEFIELDZ and QUINTIX librairies.

6.6. Codes over rings

M. Barbier, C. Chabot and Guillaume Quintin proposed a new description for quasi–cyclic codes using the ring of matrices with polynomial entries, thus defining the new class of *quasi-BCH* codes. Guillaume Quintin proved that these codes can be regarded as interleaved subcodes of Reed–Solomon codes; this allowed them to define a polynomial-time decoding algorithm for quasi-BCH codes. Guillaume Quintin also generalized list decoding algorithms to codes over non commutative rings [8].

6.7. Quantum LDPC codes

For some time it was feared that quantum computers could not be built because of distortions of quantum states due to interaction with the environment. This issue could be addressed by the use of quantum codes. *Quantum LDPC codes* are very interesting candidates here, because their very fast decoding algorithm allows high error correction rates. But the design of good quantum LDPC codes is far more complicated than for their classical counterparts, and cannot be done by random generation. The best-known constructions come from algebraic topology and simplicial homology, but their limits were unknown. Nicolas Delfosse used Riemannian geometry theorems of Gromov to prove that an $[[n,k,d]]$-quantum code constructed from the homology of a simplicial surface satisfies $kd^2 \leq C(\log k)^2 n$ for some constant C [21].

Color codes are quantum LDPC codes constructed from 3–regular surface tilings whose set of faces is 3–colorable. Delfosse used morphisms of chain complexes to prove that the decoding of a color code can be reduced to the decoding of three associated surface codes; hence, every decoding algorithm for surface codes yields a decoding algorithm for color codes. From this result, Delfosse obtained theoretical lower bounds on the error threshold of a family of color codes [20].

6.8. New families of fast elliptic curves

Benjamin Smith has pioneered the use of mod-p reductions of Q-curves to produce elliptic curves with efficient scalar multiplication algorithms—which translates into faster encryption, decryption, signing, and signature verification operations on these curves. A theoretical article was presented at ASIACRYPT 2013 [9], and the Journal of Cryptology has invited the submission of a longer version. The theory was put into practice in collaboration with Craig Costello (Microsoft Research) and Huseyin Hisil (Yasar University). Their resulting publicly available implementation, which represents the state of the art in constant-time (side-channel conscious) elliptic curve scalar multiplication on 64-bit Intel platforms at the 128-bit security level, can carry out a constant-time scalar multiplication in 145k cycles on Ivy Bridge architectures. This work will appear in EUROCRYPT 2014 [17].

6.9. Tensor rank of multiplication over finite fields

Determining the tensor rank of multiplication over finite fields is a problem of great interest in algebraic complexity theory, but it also has practical importance: it allows us to obtain multiplication algorithms with a low bilinear complexity, which are of crucial significance in cryptography. In collaboration with S. Ballet and J. Chaumine [12], Julia Pieltant obtained new asymptotic bounds for the symmetric tensor rank of multiplication in finite extensions of finite fields $\mathbb{F}_q$. In the more general (not-necessarily-symmetric) case, Pieltant and H. Randriam obtained new uniform upper bounds for multiplication in extensions of $\mathbb{F}_q$. They also gave purely asymptotic bounds substantially improving those coming from uniform bounds, by using a family of Shimura curves defined over $\mathbb{F}_q$. This work will appear in Mathematics of Computation [22].

GRAND-LARGE Project-Team

5. New Results

5.1. Automated Code Generation for Lattice Quantum Chromodynamics

Participants: Denis Barthou, Konstantin Petrov, Olivier Brand-Foissac, Olivier Pène, Gilbert Grosdidier, Michael Kruse, Romain Dolbeau, Christine Eisenbeis, Claude Tadonki.

This ongoing work is about a Domain Specific Language which aims to simplify Monte-Carlo simulations and measurements in the domain of Lattice Quantum Chromodynamics. The tool-chain, called Qiral, is used to produce high-performance OpenMP C code from LaTeX sources. We discuss conceptual issues and details of implementation and optimization. The comparison of the performance of the generated code to the well-established simulation software is also made.[33][20][37]

5.2. A Fine-grained Approach for Power Consumption Analysis and Prediction

Participants: Alessandro Ferreira Leite, Claude Tadonki, Christine Eisenbeis, Alba Cristina de Melo.

Power consumption has became a critical concern in modern computing systems for various reasons including financial savings and environmental protection. With battery powered devices, we need to care about the available amount of energy since it is limited. For the case of supercomputers, as they imply a large aggregation of heavy CPU activities, we are exposed to a risk of overheating. As the design of current and future hardware is becoming more and more complex, energy prediction or estimation is as elusive as that of time performance. However, having a good prediction of power consumption is still an important request to the computer science community. Indeed, power consumption might become a common performance and cost metric in the near future. A good methodology for energy prediction could have a great impact on power-aware programming, compilation, or runtime monitoring. In this paper, we try to understand from measurements where and how power is consumed at the level of a computing node. We focus on a set of basic programming instructions, more precisely those related to CPU and memory. We propose an analytical prediction model based on the hypothesis that each basic instruction has an average energy cost that can be estimated on a given architecture through a series of micro-benchmarks. The considered energy cost per operation includes all of the overhead due to context of the loop where it is executed. Using these precalculated values, we derive an linear extrapolation model to predict the energy of a given algorithm expressed by means of atomic instructions. We then use three selected applications to check the accuracy of our prediction method by comparing our estimations with the corresponding measurements obtained using a multimeter. We show a 9.48% energy prediction on sorting.[35]

5.3. Switcheable scheduling

Participants: Lénaïc Bagnères, Cédric Bastoul, Taj Khan.

Parallel applications used to be executed alone until their termination on partitions of supercomputers. The recent shift to multicore architectures for desktop and embedded systems is raising the problem of the coexistence of several parallel programs. Operating systems already take into account the *affinity* mechanism to ensure a thread will run only onto a subset of available processors (e.g., to reuse data remaining in the cache since its previous execution). But this is not enough, as demonstrated by the large performance gaps between executions of a given parallel program on desktop computers running several processes. To support many parallel applications, advances must be made on the system side (scheduling policies, runtimes, memory management...). However, automatic optimization and parallelization can play a significant role by generating programs with dynamic-auto-tuning capabilities to adapt themselves to the complete execution context, including the system load.

Our approach is to design at compile-time programs that can adapt at run-time to the execution context. The originality of our solution is to rely on *switcheable scheduling*, a selected set of program restructuring which allows to swap between program versions at some meeting points without backtracking. A first step selects pertinent versions according to their performance behavior on some execution contexts. The second step builds the auto-adaptive program with the various versions. Then at runtime the program selects the best version by a low overhead sampling and profiling of the versions, ensuring every computation is useful.

This work has been started at Paris-Sud University by Cédric Bastoul before he joined the Inria CAMUS project team during this year. The first results have been presented in 2013 at the HiPEAC System Week and at the Rencontres Françaises de Compilation.

5.4. Solving Navier-Stokes equations on heterogeneous parallel architectures

Participants: Marc Baboulin, Jack Dongarra, Joël Falcou, Yann Fraigneau, Olivier Lemaître, Yushan Wang.

The Navier-Stokes equations describe a large class of fluid flows but are difficult to solve analytically because of their nonlinearity. We implemented a parallel solver for the 3-D Navier-Stokes equations of incompressible unsteady flows with constant coefficients, discretized by the finite difference method. We applied the prediction-projection method which transforms the Navier-Stokes equations into three Helmholtz equations and one Poisson equation. For each Helmholtz system, we applied the Alternating Direction Implicit (ADI) method resulting in three tridiagonal systems. The Poisson equation is solved using partial diagonalization which transforms the Laplacian operator into a tridiagonal one. Our implementation is based on MPI where the computations are performed on each subdomain and information is exchanged on the interfaces, and where the tridiagonal system solutions are accelerated using vectorization techniques. We provided performance results on a current multicore system.[31]

5.5. Optimizing NUMA effects in dense linear algebra software

Participants: Marc Baboulin, Adrien Rémy, Brigitte Rozoy, Masha Sosonkina.

We studied the impact of non-uniform memory accesses (NUMA) on the solution of dense general linear systems using an LU factorization algorithm. In particular we illustrated how an appropriate placement of the threads and memory on a NUMA architecture can improve the performance of the panel factorization and consequently accelerate the global LU factorization. We applied these placement strategies and presented performance results for a hybrid multicore/GPU LU algorithm as it is implemented in the public domain library MAGMA.

GRAPHIK Project-Team

6. New Results

6.1. Ontology-Based Query Answering with Existential Rules

Participants: Jean-François Baget, Fabien Garreau, Mélanie König, Michel Leclère, Marie-Laure Mugnier, Swan Rocher, Michaël Thomazo.

Note that for this section, as well as all sections in New Results, participants are given in alphabetical order.

This year we continued to work on the existential rule framework in the context of Ontology-Based Query Answering (a.k.a. Ontology-Based Data Access, OBDA). See the 2011-2012 activity reports for details on this framework (a.k.a. Tuple-Generating Dependencies or Datalog+/-). The ontology-based query answering issue consists in querying data while taking into account inferences enabled by an ontology. This ontology is here described by existential rules, a very expressive formalism which generalizes the lightweight description logics used for OBDA (e.g. the tractable fragments of the Semantic Web language OWL 2).

From 2009 to 2011, we mainly investigated decidability and complexity issues. In 2012, we tackled the next step, which consists in developing algorithms with good theoretical properties (they should at least run in the "right" worst-case complexity class) and with good performance in practice. There are two main ways of processing rules, namely forward chaining and backward chaining, which are also known as "materialization" and "query rewriting". In forward chaining, rules are applied to enrich the initial data and query answering can then be solved solved by evaluating the query against the "saturated" database (as in a classical database system, i.e., forgetting the rules). Backward chaining process can be divided into two steps: first, the initial query is rewritten using the rules into a first-order query (typically a union of conjunctive queries, UCQ); then the rewritten query is evaluated against the initial database (again, as in a classical database system).

In 2013, on the one hand we focussed on the improvement of query rewriting algorithms, on the other hand we began to investigate extensions of our framework.

6.1.1. Improvement of Query Rewriting Algorithms

The advantage of the query rewriting approach is that the data are not modified (hence no write access permission is required and the data do not grow; moreover, there is no materialization that would need to be updated when data change). However, the practicability of this approach is questionable due to (1) the weak expressivity of classes for which efficient rewriters have been implemented, and (2) the large size of rewritings using UCQ.

With respect to the first point, we improved the algorithm designed in 2012. This algorithm accepts as input any set of existential rules and stops if this set of rules fulfills so-called finite unification set (fus) property, meaning that the set of rules allows to rewrite any query as a first-order query, e.g. a UCQ (this property is not true in general, where no finite rewriting may exist). We also studied properties of rewriting operators that ensure the correctness and the termination of a generic breadth-first rewriting algorithm and analyzed some operators with respect to these properties.

- *Work published in IJCAI 2013 [37] and RR 2013 (Rules and Web Reasoning) [36]*

With respect to the second point, we defined semi-conjunctive queries (SCQs), which are a syntactical extension of conjunctive queries. We designed and implemented an algorithm called Compact, which computes sound and complete rewritings of a conjunctive query in the form of a union of SCQs (USCQs). As in the above work, any kind of existntial rules can be considered, however the algorithm is ensured to stop only for fus rules. First experiments show that USCQs are both very efficiently computable and more efficiently evaluable than their equivalent UCQs.

- *Work published in IJCAI 2013 [41]*

6.1.2. Ongoing Work: Extensions of the Framework

Inconsistent-tolerant query answering. It may be the case that the data are inconsistent with the ontology, specially when there are several data sources. The classic logical framework becomes inappropriate since an inconsistent logical theory entails everything. Therefore, inconsistency-tolerant semantics have been defined to get meaningful answers. These semantics are based on the notion of repairs, which are maximal subsets of the data consistent with the ontology. In the most natural semantics, a tuple is an answer to the query if it is an answer in each repair. This issue is relevant to Pagoda and Qualinca, two ANR projects respectively started in 2013 and 2012 (see Section 8.1). Swan Rocher's master thesis was devoted to a query answering algorithm in this framework, where the ontology is described by existential rules and negative constraints.

Existential Rules with non-monotonic negation. Non-monotonic negation is very useful for modeling purposes. We added non-monotonic negation to existential rules, under stable model semantics. This brought us close to logic programs considered in the area called Answer Set Programming. First results were obtained on the semantics and decidability of query answering with these rules. This work is part of ASPIQ project started in 2013 (see Section 8.1).

- *Paper currently submitted to an international conference.*

6.1.3. Others

Michael Thomazo defended his PhD thesis entitled "Conjunctive Query Answering Under Existential Rules —Decidability, Complexity, and Algorithms" (Oct. 2013). The main contributions of this thesis are the following: first, a unified view of the currently known existential rule classes ensuring decidability of query answering, together with a complexity analysis and a worst-case optimal algorithm for a new generic class, which generalizes a family of very expressive decidable classes (see the gbts class in 2012 activity report); second, a generic algorithm for query rewriting, which overcomes some causes of combinatorial explosion that make classical approaches inapplicable.

- *See the PhD thesis [15] and the extended abstract published in IJCAI 2013 [42].*

The journal version extending the papers at IJCAI 2011 and KR 2012, in collaboration with Sebastian Rudolph (TU Dresden), is still in preparation but almost finished (postponement due to the addition of complementary results).

6.2. Reasoning with Imperfect Information and Priorities

Participants: Madalina Croitoru, Jérôme Fortin, Souhila Kaci, Tjitze Rienstra, Rallou Thomopoulos.

6.2.1. Monotonic and Non-monotonic Inference for Abstract Argumentation

An argumentation framework (or framework, for short) consists of a set of arguments, whose content may be left unspecified, together with an attack relation encoding conflict between arguments. Given a framework, a semantics specifies which sets of arguments (called extensions) are rationally acceptable. This formalism captures many different types of reasoning considered in the area of AI. In many applications, a framework somehow represents (part of) an agent's belief state. Beliefs are then formed on the basis of acceptable sets of arguments. For example, a 'grounded reasoner' forms beliefs on the basis of the framework's grounded extension, a 'preferred reasoner' on the basis of the preferred extensions, and so on. There is a problem with this account, however. Two different argumentation frameworks may be considered equivalent as soon as they lead to the same extensions. A more appropriate notion of equivalence is strong equivalence. Given a semantics, two frameworks are said to be strongly equivalent if their extensions are the same given every possible addition of new arguments and attacks. But still, it leaves open the question of how to form beliefs on the basis of a framework, so that different frameworks can be meaningfully distinguished, even if their extensions are the same. We addressed this problem and presented a new approach to reasoning about the outcome of an argumentation framework, where an agent's reasoning with a framework and semantics is represented by an inference relation defined over a logical labeling language. We first studied a monotonic type of inference which is, in a sense, more general than an acceptance function, but equally expressive. In order to overcome the limitations of this expressiveness, we studied a non-monotonic type of inference which allows counterfactual inferences. We precisely characterized the classes of frameworks distinguishable by the non-monotonic inference relation for the admissible semantics.

- *Joint work with R. Booth and L. van der Torre (Univ. of Luxembourg), published in FLAIRS 2013 [27]*

6.2.2. Dynamics in Abstract Argumentation

Recent years have seen a considerable work on dynamics in argumentation framework (AF). We addressed dynamics in abstract argumentation using a logical theory where an agent's belief state consists of an argumentation framework and a constraint that encodes the outcome the agent believes the argumentation framework should have. Dynamics enters in two ways: (1) the constraint is strengthened upon learning that the AF should have a certain outcome and (2) the argumentation framework is expanded upon learning about new arguments/attacks. A problem faced in this setting is that a constraint may be inconsistent with the AF's outcome. We discussed two ways to address this problem: First, it is still possible to form consistent fallback beliefs, i.e., beliefs that are most plausible given the agent's argumentation framework and constraint. Second, we showed that it is always possible to find argumentation framework expansions to restore consistency. Our work combines various individual approaches in the literature on argumentation dynamics in a general setting.

- *Joint work with R. Booth and L. van der Torre (Univ. of Luxembourg), published in SUM 2013. [26]*

Preferences have been intensively studied in argumentation framework. Preference-based argumentation frameworks are instantiation of Dung's framework in which the defeat relation (in the sense of Dung) is computed from an attack relation and a preference relation over the set of arguments. We distinguish between different ways to derive preferences over arguments, e.g., from their relative specificity, relative strength or from values promoted by the arguments. However an underexposed aspect in these models is change of preferences. We proposed a dynamic model of preferences in argumentation, centering on what we call property-based AFs. It is based on Dietrich and List's model of property-based preference and it provides an account of how and why preferences in argumentation may change. The idea is that preferences over arguments are derived from preferences over properties of arguments, and change as the result of moving to different motivational states. We also provided a dialogical proof theory that establishes whether there exists some motivational state in which an argument is accepted.

- *Joint work with R. Booth (Univ. of Luxembourg), published in ADT 2013.*

6.2.3. Representing Synergy Among Arguments with Choquet Integral

Preference-based argumentation frameworks are instantiation of Dung's framework in which the defeat relation (in the sense of Dung) is computed from an attack relation and a preference relation over the set of arguments. Value-based argumentation framework is a preference-based argumentation framework where the preference relation over arguments is derived from a preference relation over values they promote. We extended value-based argumentation framework with collective defeats and arguments promoting values with various strengths. In the extended framework, we defined a function which computes the strength of a collective defeat. We also defined desired properties for the proposed function. Surprisingly, we showed that this function obeying the corresponding properties is Choquet integral, a well-known aggregation function at work in multiple criteria decision.

- *Joint work with C. Labreuche (Thales), published in EC-SQARU 2013 [35]*

6.2.4. Compiling Preference Queries in Qualitative Constraint Problems

Comparative preference statements are the basic ingredients of conditional logics for representing users' preferences in a compact way. These statements may be strict or not and obey different semantics. Algorithms have been developed in the literature to compute a preference relation over outcomes given a set of comparative preference statements and one or several semantics. These algorithms are based on insights from non-monotonic reasoning (more specifically, minimal and maximal specificity principles) enforcing the preference relations to be a complete preorder. The main limitation of these logics however relies in preference queries when comparing two outcomes. Indeed given two outcomes having the same preference w.r.t. the preference relation, there is no indication whether this equality results from an equality between two preference statements or the outcomes are in fact incomparable and equality has been enforced by specificity principles. On the other hand, comparative preference statements and their associated semantics can be translated into qualitative constraint satisfaction problems in which one can have a precise ordering over two outcomes. We investigated this bridge and provided a compilation of conditional logics-based preference queries in qualitative constraint problems.

- *Joint work with J.-F. Condotta (CRIL), published in FLAIRS 2013 [31]*

6.2.5. Argumentation for Reasoning with Inconsistencies

We investigate the use of argumentation when reasoning over an inconsistent knowledge base. We use argumentation in this context given the explanation power that it may bring (and that is currently under investigation).

We have investigated logical based argumentation following two methods. First, we have defined our own argument and attack notion (given the logical language at hand) and showed that such instantiation respects desirable properties of consistency and maximality (called rationality postulates in the field). This work has showed that the ICR, AR, IAR semantics investigated by inconsistent query answering (see Pagoda, Section 8.1) are the same as skeptically preferred or stable semantics, grounded and universally stable or preferred. Such result is encouraging as it bridges the two communities (argumentation and inconsistent query answering) allowing to use results from one field in order to enrich the other. We have also investigated the practical applicability of such argument definition and approach in the selection of flour for bread.

- *Joint work with Srdjan Vesic (Univ. of Luxembourg), published in RIA 2013 [23] and SUM 2013 [32]*

.

On the other hand we have also looked at using a generic logical argumentation framework (ASPIC) in order to instantiate it with a simple logic in the EcoBioCap project (see Section 8.2). We have extended previous results to enrich bipolar queries. A software tool is under construction.

- *Work published in RIA 2013 [21]*

6.3. Semantic Data Integration

Participants: Michel Chein, Madalina Croitoru, Léa Guizol, Michel Leclère.

It often happens that different references (i.e. data descriptions), possibly coming from heterogeneous data sources, concern the same real world entity. In such cases, it is necessary: (i) to detect whether different data descriptions really refer to the same real world entity and (ii) to fuse them into a unique representation. Since the seminal paper [59], this issue has been been studied under various names: “record linking”, “entity resolution”, “reference resolution”, “de-duplication”, “object identification”, “data reconciliation”, etc., mostly in databases (cf. the bibliography by William E. Winckler [60]). It has become one of the major challenges in the Web of Data, where the objective is to link data published on the web and to process them as a single distributed database.

We investigate this problem in the specific context of bibliographic databases. Indeed, people working in bibliographical information systems have a lasting tradition of using norms and have integrated, along collections of documents notices (e.g. bibliographic records), collections of *authority notices* that categorize the different named entities used to describe documents (people, organizations, places, ...). In current databases, documents notices do not use directly the names of named entities to fill a particular field (author, editor, ...), but the unique identifier of the authority notice representing that named entity.

Past years, we began a collaboration with ABES (National Bibliographic Agency for Universities) to develop a method and a prototype to perform entity resolution between on one hand the authors of a new bibliographic record, and, on the other the authority references of an authority catalog (and namely the Sudoc catalogue from the ABES agency). The prototype providing this service has been implemented on top of Cogui and experiments have been led in the context of the SudocAd project (jointly conducted by ABES and GraphIK).

Our proposed method can be stated as follows: first, enrich authority records with knowledge extracted from bibliographic records in which the authority is mentioned ; then, use logical rules which conclude on different levels of reconciliation, to compare the authors of a new bibliographic record with the enriched authority records ; finally, for each author of the new bibliographic record, order the authority identifiers by level of reconciliation.

- *Work published in [30].*

A problem with this approach is that it relies upon pre-established links between bibliographic records and authority notices. However, our experimentation and evaluation have shown that many existing links were erroneous, and thus led to the propagation of new linkage errors. We have thus began to work on methods and tools to repair linkage errors in bibliographical databases. This year, this work has been pursued along three different axis:

1. We have built a formal framework allowing to evaluate the quality of links in a documents database. We propose two different "quality" notions, based upon an identification predicate'*id* and a differentiation predicate *di* between pairs of authority notices identifiers (these predicates can be either given by an expert or computed using rules). We have first introduced the notion of a *well-founded* database, when *id* is an equivalence relation and *di* its complement. This property can be checked using logical inferences and combinatorial techniques. In the general case where a database is not necessarily well-founded, we have proposed different distances to a well-founded one. We have also introduced a more complex quality criterion that corresponds to *stability by substitution* (a fundamental property of logical equality that is not necessarily satisfied by *id*).
   - *A research report should lead to a publication in 2014.*
2. We developed a methodology for detecting linkage errors and fixing them, based upon a clustering method of authors in bibliographic records. Last year, the general schema of the methodology was defined. It is based upon a set of criteria which allows us to cluster "similar" authors together. Each criterion represents a point of view on the author: name, publication time span, publication domain, etc... This year, two aggregation semantics for such criteria have been developed, implemented and evaluated.
   - *Work published in AI-SGAI 2013 [34].*
3. We have studied methods allowing to automatically extract similarity criteria between named entities. This problem is very similar to the automatic discovery of composite key constraints in RDF data sources that conform to a given ontology. We have studied the different existing methods allowing to discover such keys, and have proposed logical semantics for these different keys. These semantics allow to understand and compare the results produced by these different methods. These methods have been evaluated against the documentary databases provided by our partners ABES and INA.
   - *Work described in a research report [48], at the moment, two papers are submitted.*

HIEPACS Project-Team

6. New Results

6.1. High-performance computing on next generation architectures

6.1.1. Composing multiple StarPU applications over heterogeneous machines: a supervised approach

Enabling HPC applications to perform efficiently when invoking multiple parallel libraries simultaneously is a great challenge. Even if a uniform runtime system is used underneath, scheduling tasks or threads coming from different libraries over the same set of hardware resources introduces many issues, such as resource oversubscription, undesirable cache flushes or memory bus contention.

This paper presents an extension of StarPU, a runtime system specifically designed for heterogeneous architectures, that allows multiple parallel codes to run concurrently with minimal interference. Such parallel codes run within *scheduling contexts* that provide confined execution environments which can be used to partition computing resources. Scheduling contexts can be dynamically resized to optimize the allocation of computing resources among concurrently running libraries. We introduce a *hypervisor* that automatically expands or shrinks contexts using feedback from the runtime system (e.g. resource utilization). We demonstrate the relevance of our approach using benchmarks invoking multiple high performance linear algebra kernels simultaneously on top of heterogeneous multicore machines. We show that our mechanism can dramatically improve the overall application run time (-34%), most notably by reducing the average cache miss ratio (-50%).
This work is developed in the framework of Andra Hugo's PhD. These contributions have been presented at the international workshop on Accelerators and Hybrid Exascale Systems [19] in Boston.

6.1.2. A task-based H-matrix solver for acoustic and electromagnetic problems on multicore architectures

$\mathcal{H}$-Matrix is a hierarchical, data-sparse approximate representation of matrices that allows the fast approximate computation of matrix products, LU and LDL^T decompositions, inversion and more. This representation is suitable for the direct solution of large dense linear systems arising from the Boundary Element Method in $O(N \log_2^\alpha(N))$ operations. This kind of formulation is widely used in the industry for the numerical simulation of acoustics and electromagnetism scattering by large objetcs. Applications of this approach include aircraft noise reduction and antenna sitting at Airbus Group. The recursive and irregular nature of these $\mathcal{H}$-Matrix algorithms makes an efficient parallel implementation very challenging, especially when relying on a "Bulk Synchronous Parallel" paradigm. We have considered an alternative parallelization for multicore architectures using a task-based approach on top of a runtime system, namely StarPU. We have showed that our method leads to a highly efficient, fully pipelined computation on large real-world industrial test cases provided by Airbus Group.
This research activity has been conduced in the framework of the EADS-ASTRIUM, Inria, Conseil Régional initiative in collaboration with the RUNTIME Inria project, and is part of Benoit Lize's PhD.

6.1.3. A task-based 3D geophysics application

Reverse Time Migration (RTM) technique produces underground images using wave propagation. A discretization based on the Discontinuous Galerkin (DG) method unleashes a massively parallel elastodynamics simulation, an interesting feature for current and future architectures. We have designed a task-based version of this scheme in order to enable the use of manycore architectures. At this stage, we have demonstrated the efficiency of the approach on homogeneous and cache coherent Non Uniform Memory Access (ccNUMA) multicore platforms (up to 160 cores) and designed a prototype version of a distributed memory version that can exploit multiple instances of such architectures. This work has been conducted in the context of the DIP Inria-Total strategic action in collaboration with the MAGIQUE3D Project-Team and thanks to the long-term visit of George Bosilca funded by TOTAL. Geroge's expertise ensured an optimum usage of the PaRSEC runtime system onto which our task-based scheme has been ported.

This work was presented during a PRACE workshop [28] as well as during a TOTAL scientific event [29].

6.1.4. Resiliency in numerical simulations

For the solution of systems of linear equations, various recovery-restart strategies have been investigated in the framework of Krylov subspace methods to address the situations of core failures. The basic underlying idea is to recover fault entries of the iterate via interpolation from existing values available on neighbor cores. The resulting results are reported in the research report [41] currently submitted to an international journal. In that resilience framework, we have extended the recovey-restart ideas to the solution of linear eigenvalue problems. Contrary to the linear system case, not only the current iterate can be interpolated but also part of the subspace where candidate eigenpairs are searched.
This work is developed in the framework of Mawussi Zounon's PhD funded by the ANR RESCUE. These contributions have been presented at the international workshop Sparse Days [27] in Toulouse. More details and results can be found in report RR-8324 [41]. Notice that theses activities are also part of our contribution to the G8 ESC (Enabling Climate Simulation at extreme scale).

6.2. High performance solvers for large linear algebra problems

6.2.1. Parallel sparse direct solver on runtime systems

The ongoing hardware evolution exhibits an escalation in the number, as well as in the heterogeneity, of the computing resources. The pressure to maintain reasonable levels of performance and portability, forces the application developers to leave the traditional programming paradigms and explore alternative solutions. Algorithms, especially those in critical domains such as linear algebra, need to undergo invasive structural changes and be adapted to new programming paradigms to be in agreement with the latest hardware advances. PaStiX is a parallel sparse direct solver, based on a dynamic scheduler for modern hierarchical architectures. In this paper, we study the replacement of the highly specialized internal scheduler in PaStiX by two generic runtime frameworks: PaRSEC and StarPU. The tasks graph of the factorization step is made available to the two runtimes, providing them with the opportunity to optimize it in order to maximize the algorithm efficiency for a predefined execution environment. A comparative study of the performance of the PaStiX solver with the three schedulers on different execution contexts is performed. The analysis highlights the similarities from a performance point of view between the different execution supports. These results demonstrate that these generic DAG-based runtimes provide a uniform and portable programming interface across heterogeneous environments, and are, therefore, a sustainable solution for hybrid environments.
This work is developed in the framework of Xavier Lacoste's PhD funder by the ANR ANEMOS. These contributions have been presented at the international workshop Sparse Days [37] in Toulouse. More details and results can be found in report RR-8446 [46].

6.2.2. Hybrid parallel implementation of hybrid solvers

In the framework of the hybrid direct/iterative MaPHyS solver, we have designed and implemented an hybrid MPI-thread variant. More precisely, the implementation rely on the multi-threaded MKL library for all the dense linear algebra calculations and the multi-threaded version of PaStiX. Among the technical difficulties, one was to make sure that the two multi-threaded libraries do not interfere with each other. The resulting software prototype is currently experimented to study its new capability to get flexibility and trade-off between the parallel and numerical efficiency. Parallel experiments have been conducted on the Plafrim plateform as well as on a large scale machine located at the USA DOE NERSC, which has a large number of CPU cores per socket.
This work is developed in the framework of the PhD thesis of Stojce Nakov funded by TOTAL. These contributions have been presented at the NVIDIA GPU Technology Conference [25] in San Jose.

6.2.3. Designing LU-QR hybrid solvers for performance and stability

New hybrid LU-QR algorithms for solving dense linear systems of the form $Ax = b$ have been introduced. Throughout a matrix factorization, these algorithms dynamically alternate LU with local pivoting and QR elimination steps, based upon some robustness criterion. LU elimination steps can be very efficiently parallelized, and are twice as cheap in terms of flops, as QR steps. However, LU steps are not necessarily stable, while QR steps are always stable. The hybrid algorithms execute a QR step when a robustness criterion detects some risk for instability, and they execute an LU step otherwise. Ideally, the choice between LU and QR steps must have a small computational overhead and must provide a satisfactory level of stability with as few QR steps as possible. In this paper, we introduce several robustness criteria and we establish upper bounds on the growth factor of the norm of the updated matrix incurred by each of these criteria. In addition, we describe the implementation of the hybrid algorithms through an extension of the PaRSEC software to allow for dynamic choices during execution. Finally, we analyze both stability and performance results compared to state-of-the-art linear solvers on parallel distributed multicore platforms. These contributions have been presented at the international conference IPDPS [18] in Phoenix.

6.3. High performance Fast Multipole Method for N-body problems

Last year we have worked primarily on developing an efficient fast multipole method for heterogeneous architecture. Some of the accomplishments for this year include:

1. Implementation of the FMM of multicore machines using StarPU. A new parallel scheduler was developed for this purpose. We implemented a state-of-the-art OpenMP version of the code for benchmarking purposes. It was found that StarPU significantly outperforms OpenMP. Figures show the traces of an execution of the FMM algorithm with our priority scheduler for the cube (volume) and ellipsoid (surface) with 20 million particles on a 4 deca-core Intel Xeon E7-4870 machine.
2. Implementation of the FMM of heterogeneous machines (CPU+GPU) using StarPU. The FMM was also used to demonstrate the flexibility of StarPU for handling different types of processors. In particular we demonstrated in that application that StarPU can automatically select the appropriate version of a computational kernel (CPU or GPU version) and run it on the appropriate processor in order to minimize the overall runtime. Significant speed-up were obtained on heterogeneous platforms compared to multicore only processors.

These contributions have been presented in minnisymposia at the SIAM conference on Comutational Sciences and Engineering [23], [33] in Boston and at NVIDIA GPU Technology Conference [24]. More details and results can be found in report RR-8277 [40], our paper is accepted for publication in the SIAM Journal on Scientific Computing [11].

Concerning dynamics dislocations (DD) kernels, an efficient formulation of the isotropic elastic far-field interactions between dislocations has been developed. This formulation is suitable for any polynomial interpolation based Fast Multipole Method (FMM) and is currently being implemented in OptiDis.

Meanwhile a much lighter and faster interpolation scheme based on a uniform grid (i.e. Lagrange interpolation) and the Fast Fourier Transform (FFT) was implemented into ScalFMM. This last feature was introduced in order to overcome the expensive cost of the Chebyshev FMM in the range of low interpolation orders (up to approx. 10). This should significantly improve the performances of the far-field computation in DD simulations where tensorial kernels are involved but only relatively low interpolation orders are required. This work is developed in the framework of Pierre Blanchard's PhD funded by ENS.

6.4. Efficient algorithmic for load balancing and code coupling in complex simulations

6.4.1. Dynamic load balancing for massively parallel coupled codes

As a preliminary step related to the dynamic load balancing of coupled codes, we focus on the problem of dynamic load balancing of a single parallel code, with variable number of processors. Indeed, if the workload

varies drastically during the simulation, the load must be redistributed regularly among the processors. Dynamic load balancing is a well studied subject but most studies are limited to an initially fixed number of processors. Adjusting the number of processors at runtime allows to preserve the parallel code efficiency or to keep running the simulation when the current memory resources are exceeded. We call this problem, *MxN graph repartitioning*. We propose some methods based on graph repartitioning in order to rebalance the load while changing the number of processors. These methods are split in two main steps. Firstly, we study the migration phase and we build a "good" migration matrix minimizing several metrics like the migration volume or the number of exchanged messages. Secondly, we use graph partitioning heuristics to compute a new distribution optimizing the migration according to the previous step results. Besides, we propose a direct k-way partitioning algorithm that allows us to improve our biased partitioning. Finally, an experimental study validates our algorithms against state-of-the-art partitioning tools. Our algorithms are implemented in the `LBC2` library and have been integrated in the partitioning tools `Scotch` as a prototype.
This work is developed in the framework of Clément Vuchener's PhD, that will be defended on February 2014. These contributions have been presented at the international conference ParCo [22] in Munchen.

Regarding the problem of dynamic balancing of parallel coupled codes, we start to reuse results on *MxN graph repartitioning*. Given two coupled codes A and B, the key idea is to develop an algorithm of *two-graph co-partitioning*, that partitions two *coupled* graphs G_A and G_B in respectively N_A and N_B with classic objectives (*i.e.*, balancing computational load and minimizing communication cost for each code) and that minimizes the number of messages exchanged between codes in the coupling phase.
This work is developed in the framework of Maria Predari's PhD, that just started in october 2013.

6.4.2. Graph partitioning for hybrid solvers

Nested Dissection has been introduced by A. George and is a very popular heuristic for sparse matrix ordering before numerical factorization. It allows to maximize the number of parallel tasks, while reducing the fill-in and the operation count. The basic standard idea is to build a "small separator" S of the graph associated with the matrix in order to split the remaining vertices in two parts P_0 and P_1 of "almost equal size". The vertices of the separator S are ordered with the largest indices, and then the same method is applied recursively on the two sub-graphs induced by P_0 and P_1. At the end, if k levels of recursion are done, we get 2^k sets of independents vertices separated from each other by $2^k - 1$ separators.
However, if we examine precisely the complexity analysis for the estimation of asymptotic bounds for fill-in or operation count when using Nested Dissection ordering, we can notice that the size of the halo of the separated sub-graphs (set of external vertices belonging to an old separator and previously ordered) plays a crucial role in the asymptotic behavior achieved. In the perfect case, we need halo vertices to be balanced among parts.
Considering now hybrid methods mixing both direct and iterative solvers such as `HIPS`, `MaPHyS`, obtaining a domain decomposition leading to a good balancing of both the size of domain interiors and the Scalable numerical schemes for scientific applications size of interfaces is a key point for load balancing and efficiency in a parallel context. This leads to the same issue: balancing the halo vertices to get balanced interfaces.
For this purpose, we revisit the algorithm introduced by Lipton, Rose and Tarjan which performed the recursion of nested dissection in a different manner: at each level, we apply recursively the method to the sub-graphs But, for each sub-graph, we keep track of halo vertices. We have implemented that in the Scotch framework, and have studied its main algorithm to build a separator, called greedy graph growing.
This work is developed in the framework of Astrid Casadei's PhD. These contributions have been presented at the international workshop on Nested Dissection [32] in Waterloo.

6.5. Application Domains

6.5.1. Dislocation dynamics simulations in material physics

This year we have focused on the hybrid parallelization of the OptiDis code. As dislocations move in their grain, they expand, shrink, collide and annihilate, which means that we are facing a extremely dynamic n-body problem. Also, we have introduced an adaptive cache conscious data structure to manage the dislocation mesh. Moreover, two main kernels, plugged in our `ScalFMM` library, was built to handle the pairwise force interactions and the collisions between dislocations. Finally the code is written using hybrid parallelism based on OpenMP tasks inside on node and MPI to exchange data between nodes. The code can run on both shared and distributed memories. Future works will mainly focus on tuning the code and manage dynamically this tuning to adapt to different kind of simulations and architectures. On the physical side, we have introduced more *split node* cases to simulate irradiated materials. Now we are able to run simulations with tens of thousand of defaults in materials. Typically, our simulation box can hold lot of tiny dislocation loops such as those induced by radiation on materials, so we can observe how Frank-Read sources interact while they cross the field of loop defects.
This work is developed in the framework of Arnaud Etcheverry's PhD funded by the ANR OPTIDIS.

6.5.2. Co-design for scalable numerical algorithms in scientific applications

The study of the **thermo-acoustic stability of large combustion chambers** requires the solution of a nonlinear eigenvalue problem. The nonlinear problem is linearized using a fixed point iteration procedure. This leads to a sequence of linear eigenproblems which must be solved iteratively in order to obtain one nonlinear eigenpair. Therefore, efficient and robust parallel eigensolvers for the solution of linear problems have been investigated, and strategies to accelerate the solution of the sequence of linear eigenproblems have also been proposed. Among the numerical techniques that have been considered (Krylov-Schur, Implicitly Restarted Arnoldi, Subspace iteration with Chebyshev acceleration) the Jacobi-Davidson method was the best suited to be combined with techniques to recycle spectral information between the nonlinear iterations. The robustness of the parallel numerical techniques were illustrated on large problems with a few millions unknowns solved on a few tens of cores.
These results are part of the outcome of Pablo Salas PhD thesis that has been defended on November 15th.

The **Time-domain Boundary Element Method** (TD-BEM) has not been widely study but represent an interesting alternative to its frequency counterpart. Usually based on inefficient Sparse Matrix Vector-product (SpMV), we investigate other approaches in order to increase the sequential flop-rate. We have implement extremely efficient operator using intrinsic SIMD or even ASM64 instructions. We are using this novel approaches to parallelize both in shared and distributed memory and target execution on hundreds of clusters. All the implementations should be in high quality in the Software Engineering sense since the resulting library is going to be used by industrial applications.
This work is developed in the framework of Bérenger Bramas's PhD and contributes to the EADS-ASTRIUM, Inria, Conseil Régional initiative.

In a preliminary work, a **3D Cartesian SN solver**`DOMINO` has been designed and implemented using two nested levels of parallelism (multicore+SIMD) on shared memory computation nodes. `DOMINO` is written in C++, a multi-paradigm programming language that enables the use of powerful and generic parallel programming tools such as Intel TBB and Eigen. These two libraries allow us to combine multi-thread parallelism with vector operations in an efficient and yet portable way. As a result, `DOMINO` can exploit the full power of modern multi-core processors and is able to tackle very large simulations, that usually require large HPC clusters, using a single computing node. The very high Flops/Watt ratio of `DOMINO` makes it a very interesting building block for a future many-nodes nuclear simulation tool.
This work is developed in the framework of Salli Moustafa's PhD in collaboration with EDF. These contributions have been presented at the international conference on Supercomputing on Nuclear Applications [21] in Paris.

Concerning the numerical simulation of **the turbulence of plasma particules inside a tokamak**, two software tools, providing a post-mortem analysis, have been designed to manage the memory optimization of GYSELA [20]. The first one is a visualization tool. It plots the memory consumption of the code along an execution. This tool helps the developer to localize where happens the memory peak and to wonder how he can modify the code to decrease it. On the same graphic, the names of the allocated structures are labelled, which gives a significant hint on the modifications to apply. The second tool concerns the prediction of the peak memory. Given an input set of parameters, we can replay the allocations of the code in an offline mode. With this tool, we can deduce accurately the value of the memory peak and where it happens. Thank to this prediction we know which size of mesh is possible under a given architecture.
This work is carried on in the framework of Fabien Rozar's PhD in collaboration with CEA Cadarache.

In the first part of our research work concerning the parallel **aerodynamic code**FLUSEPA, an intermediate version based on the previous one has been developped. By using an hybrid OpenMP/MPI parallelism based on a domain decomposition, we achieved a faster version of the code and the temporal adaptive method used without bodies in relative motion has been tested successfully for real complex 3D-cases using up to 400 cores. Moreover, an asynchronous strategy for computing bodies in relative motion and mesh intersections has been developed and the test of this feature is currently in progress. The next step will be to design a new fully asynchronous code based on a task graph description to be executed on a modern runtine system like StarPU.
This work is carried on in the framework of Jean-Marie Couteyen's PhD in collaboration with Astrium Les Mureaux.

HIPERCOM2 Team

6. New Results

6.1. Wireless Sensor Networks

6.1.1. Node activity scheduling and routing in Wireless Sensor Networks

Participants: Cédric Adjih, Ichrak Amdouni, Pascale Minet.

The need to maximize network lifetime in wireless ad hoc networks and especially in wireless sensor networks requires the use of energy efficient algorithms and protocols. Motivated by the fact that a node consumes the least energy when its radio is in sleep state, we achieve energy efficiency by scheduling nodes activity. Nodes are assigned time slots during which they can transmit and they can turn off their radio when they are neither transmitting nor receiving. Compared to classical TDMA-based medium access scheme, spatial bandwidth use is optimized: non interfering nodes are able to share the same time slots, collisions are avoided and overhearing and interferences are reduced. In our work about time slots assignment, two cases are studied. First, when nodes require equal channel access, we use node coloring. Second, when nodes have heterogeneous traffic demands, we designed the traffic aware time slot assignment algorithm TRASA. Unlike the majority of previous works, we generalize the definition of node coloring and slot allocation problems. Indeed, we set the maximum distance between two interfering nodes as a parameter of these problems. We prove that they are NP-complete, making heuristic approaches inevitable in practice. A central directive of this thesis is to design self-adaptive solutions. This adaptivity concerns many aspects such as the mission given by the application, the heterogeneity of node traffic demands, the network density, the regularity of network topology, and the failure of wireless links.

In the GETRF project, we target the energy efficiency in wireless sensor networks. We proposed node activity scheduling approaches that determine active and inactive slots for sensor nodes as to enable them to turn off their radio and save energy in the inactive slots.

1. First, we proposed a scheduling algorithm based on node coloring of grid sensor networks called VCM. This proposal was strengthen with mathematical analysis of the optimal number of colors needed to color an infinite grid. VCM produced an optimal number of colors when the transmission range tends to infinity. Also, this algorithm does not require message exchange between sensors to determine colors.

2. Second, this work was extended to adapt it to general graphs: the graph is divided into cells and the color of the cell is the color of the node on the left bottom of the cell. Nodes inside the cell are scheduled successively.

In addition to the energy efficiency, we targeted the delay optimization for data collection applications in grid wireless sensor networks. We profit from the previous work VCM and integrate it with a new hierarchical routing method to minimize data collection delays.

6.1.2. Time slot and channel assignment in multichannel Wireless Sensor Networks

Participants: Pascale Minet, Ridha Soua, Erwan Livolant.

Applying WSNs in industrial environment requires fast and reliable data gathering (or data convergecast). If packets are forwarded individually to the sink, it is called raw data convergecast. We resort to the multichannel paradigm to enhance the data gathering delay, the robustness against interferences and the throughput. Since some appli- cations require deterministic and bounded convergecast delays, we target conflict free joint time slot and channel assignment solutions that minimize the schedule length. Such solutions allow nodes to save energy by sleeping in any slot where they are not involved in transmissions. We extend existing multichannel results to take into account a sink equipped with multiple radio interfaces and heterogeneous traffic demands. Indeed, we compute the theoretical bounds, that is the minimum number of time slots needed to complete convergecast, in various topologies with different traffic demands. These bounds are provided for different acknowledgment policies. For each of them, we provide a graph-based interference model. We also give optimal schedules that achieve these optimal bounds. We formalize the problem of multichannel slot assignment using integer linear programming and solve with GLPK tool for small configurations.

We propose MODESA, a centralized joint time slot and channel assignment algorithm. We prove the optimality of MODESA in lines, multilines and balanced trees topologies. By simulations, we show that MODESA outperforms TMCP, a well known subtree-based scheduling. We improve MODESA with different channel allocation strategies depending on the channel selection criteria (channels load balancing or preference of channels with the best qualities). Moreover, we show that resorting to multipath routing minimizes the convergecast delay. This work is extended in MUSIKA to take into account multi-sinks WSNs and traffic differentiation: the problem is formalized using integer linear programming and solved with GLPK. Simulations results show that the schedule length is minimized and the buffer size is reduced. We then address the adaptivity challenge. The slot assignment should be more flexible and able to adapt to application and environment variability (e.g., alarms, temporary additional demands). Theoretical bounds on the number of additional slots intro- duced to cope with traffic changes, are given. AMSA, an incremental solution, is proposed. Its performances are evaluated in two cases: retransmissions or temporary changes in appli- cation needs.

6.1.3. WSN Redeployment

Participants: Pascale Minet, Saoucene Ridene, Ines Khoufi.

This is a joint work with Telecom SudParis: Anis Laouiti.
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. DVFA, is our 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 DVFA gives very good coverage rate (between 98% and 100%) and ensures the connectivity between sensors.

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.

6.1.4. Opportunistic routing cross-layer schemes for low duty-cycle wireless sensor networks

Participants: Mohamed Zayani, Paul Muhlethaler.

This is a joint work with Nadjib Aitsaadi from University of Paris 12.
The opportunistic aspect of routing is suitable with such networks where the topology is dynamic and protocols based on topological information become inefficient. Previous work initiated by Paul Muhlethaler and Nadjib Aitsaadi consisted in a geographical receiver-oriented scheme based on RI-MAC protocol (Receiver-Initiated MAC). This scheme is revised and a new contribution proposes to address the same problem with a sender-oriented approach. After scrutinising different protocols belonging to this classification, the B-MAC protocol is chosen to build a new opportunistic cross-layer scheme. Our choice is motivated by the ability of this protocol to provide to a sender the closest neighbor to the destination (typically a sink). In other words, such a scheme enables us to obtain shorter paths in terms of hops which would increase the efficiency of information delivery. In counterparts, as it relies on long preambles (property of B-MAC) to solicit all the neighborhood, it needs larger delays and energy consumption (1% of active time). Nevertheless, this proposal remains interesting as the studied networks are dedicated to infrequent event detection and are not real time-oriented.

Starting from a simulator coded by Nadjib Aitsaadi for the receiver-oriented scheme, the new scheme has been coded under many variants. On top of ideal techniques, a realistic variant has been considered and modelled. Its particularity can be summarized in the election process of the next hop. Indeed, it is based on sending bursts by the potential candidates to receive a packet from a sender. These bursts express the closeness of each candidate to the destination and correspond to the binary complement of the distance to this destination.

The opportunistic cross-layer scheme, when designed with RI-MAC, has shown solid performances in carrying the information about a rare event detection to a sink. This is verified for an event detected by several nodes. Nevertheless, the efficiency of such a design becomes less obvious when the detection is performed by a very small number of nodes. the opportunistic routing using RI-MAC relies on a minor set of potential candidates to forward a packet. In other words, a sender can only select an awake neighbor (typically closer to the sink) as the next hop. To overcome this limitation, we initially proposed to limit the number of hops to reach the sink.The principle of B-MAC perfectly matches with this idea. It is also important to highlight the ability of an opportunistic cross-layer built over B-MAC to avoid collisions. B-MAC- and RI-MAC-based proposals are suitable to convey emergency packets in dense and large WSNs when the event is reported by a significant set of nodes. When this set is limited, the sake of efficiency rather suggests a scheme based on B-MAC. It should be remembered that the proposed schemes extremely limit the energy consumption compared to classical networks.

6.1.5. Data dissemination in Urban Environment

Participants: Belhaoua Asma, Nadjib Achir, Paul Muhlethaler.

Over the last decade, wireless sensor networks have brought valid solutions to real-world monitoring problems. Sensors are now incorporated in all our modern life facilities, such as mobile phones, vehicles, buses, bus stations, bikes, etc. For example, mobile phones, with their increasing capabilities are used as voice communication device but also as a sensing device able to collect data such as image, audio, GPS position, speed, etc. All these sensors could play an important role in the provisioning of a multitude of dynamic information about their environmental trends. Considering that, WSN could be considered as a valid solution to urban monitoring problems by bringing new services for the city or for the citizens. According to the last requirement, the main question that we need to answer is how the data could be collected and/or transmitted? Several algorithms were developed recently for sensor data gathering in WSN. However, the majority of existing works on WSN has focused only on specific areas applications, such as environmental monitoring, military target tracking, weather forecast, home automation, intrusion detection, etc. In this training we studied the existing strategies of dissemination in Delay/ Disruption Tolerant Networks (DTN). The main objective is to identify those that can be applied to urban environments. We implemented and tested several strategies in the WSNet network simulator on a dense network.

6.2. Cognitive Radio Networks

6.2.1. Multichannel time slot assignment in Cognitive Radio Sensor Networks

Participants: Ons Mabrouk, Pascale Minet, Ridha Soua, Ichrak Amdouni.

This is a joint work with Hanen Idoudi and Leila Saidane from ENSI, Tunisia.
Current Wireless Sensor Networks (WSNs) are deployed over unlicensed frequency bands that face an increased level of interference from various wireless systems. Cognitive Radio Sensor Networks (CRSNs) overcome this problem by allowing sensor nodes to access new spectrum bands to minimize interferences. In this paper, we focus on the MultiChannel Time Slot Assignment problem (MC-TSA) in CRSN. Each sensor node is assigned the number of time slots it needs to transfer its own data as well as the data received from its children in the rooting tree rooted at the sink without interfering with other secondary users. Besides, sensor nodes cannot transmit on a channel occupied by a primary user. Our objective is to increase the network throughput offered to sensor nodes. We start by formulating the MC-TSA problem as an Integer Linear Program where the goal is to minimize the number of slots in the schedule.We then propose an Opportunistic centralized TIme slot assignment in COgnitive Radio sensor networks (OTICOR). We evaluate its performance in terms of number of slots and throughput.

6.2.2. Leader election in Cognitive Radio Networks

Participants: Paul Muhlethaler, Dimitrios Milioris.

This is a joint work with Philippe Jacquet from Alcatel-Lucent Bell Labs.
In this sudy we have introduced a new algorithm (green election) to achieve a dis- tributed leader election in a broadcast channel that is more efficient than the classic Part-and-Try algorithm. The algorithm has the adavantage of having a reduced overhead $log(log(N))$ rather than $log(N)$. More importantly the algorithm has the a greatly reduced energy consumption since it requires $O(N^{1/k})$ burst transmissions instead of $O(N/k)$, per election, k being a parameter depending on the physical properties of the medium of communication.

One of the applications of green election is for wireless col- lision algorithms in particular in cognitive wireless networks where the secondary network is WiFi IEEE 802.11. Since the green election is low energy consuming, it can be used as a systematic and repetitive medium access control that will naturally prevail over the WiFi CSMA scheme.

6.3. Development, implementation and distribution of the Ey-Wifi module for the NS3 simulation tool

Participants: Hana Baccouch, Cédric Adjih, Paul Muhlethaler.

Ey-Wifi module is an ns-3 module developed within the Mobsim project. Ey-Wifi stands for Elimination-Yield for WiFi networks. The main goal of Ey-Wifi is to integrate the features of the EY-NPMA channel access scheme in the ns-3 Wifi module. EY-NPMA (Elimination-Yield Non-Pre-emptive Priority Multiple Access) is a contention based protocol that has been used as the medium access scheme in HIPERLAN type 1. The main advantages of EY-NPMA are: low collision rate, more determinism and priority support. EY-NPMA is based on active signaling (black burst): a node requests access to the medium by transmitting a burst signal. More precisely, the channel access cycle comprises three phases : priority phase, elimination phase and yield phase. Compared to Wifi, EY-NPMA adds the transmission of a burst in the elimination phase: it reduces the number of nodes, that will compete in next "yield" phase (equivalent to the contention window based access of WiFi).

Furthermore, the performances of Ey-Wifi have been evaluated and compared with those of Wifi with ns-3. Distribution of Ey-Wifi module: The module and a tutorial explaining how to use it, are available at: http://hipercom.inria.fr/Ey-Wifi

6.4. Mobile ad hoc and mesh networks

6.4.1. Geographic routing and location services

Participants: Selma Boumerdassi, Pascale Minet, Paul Muhlethaler.

Thanks to its scalable nature, geographic routing is an interesting alternative to topological routing for ad-hoc networks. In fact, in order to set up such a network, each node needs to know the location of the others and location services are in charge to provide such an information.

Two kinds of location services have been provided using either a flooding or a rendez-vous, a node in the network being chosen as a server for the rendez-vous. In the scope of our research, we have proposed different mechanisms based on social groups and/or communities and studied their impact on the control traffic of various protocols. For example, based on the simulations of SLS and SFLS using NS-2, we have demonstrated that the social behaviour of nodes has a strong impact on location services and therefore that next-generation location services should take the relationships between the network users into account.

6.4.2. Optimized Broadcast Scheme for Mobile Ad hoc Networks

Participants: Ahmed Amari, Nadjib Achir, Paul Muhlethaler.

In this training we propose an optimized broadcasting mechanism, which uses very limited signaling overhead. The main objective is to select the most appropriate relay nodes according to a given cost function. Basically, after receiving a broadcast packet each potential relay node computes a binary code according to a given cost function. Then, each node starts a sequence of transmit/listen intervals following this code. In other words, each 0 corresponds to a listening interval and each 1 to a transmit interval. During this active acknowledgment signaling period, each receiver applies the following rule: if it detects a signal during any of its listening intervals, it quits the selection process, since a better relay has also captured the packet. Finally, we split the transmission range into several sectors and we propose that all the nodes within the same sector use the same CDMA orthogonal spreading codes to transmit their signals. The CDMA codes used in two different sectors are orthogonal, which guarantees that the packet is broadcast in all possible directions.

6.5. Learning for an efficient and dynamic management of network resources and services

Participants: Dana Marinca, Pascale Minet.

To guarantee an efficient and dynamic management of network resources and services we intend to use a powerful mathematical tool: prediction and learning from prediction. Prediction will be concerned with guessing the short-term, average-term and long-term evolution of network or network components state, based on knowledge about the past elements and/or other available information. Basically, the prediction problem could be formulated as follows: a forecaster observes the values of one or several metrics giving indications about the network state (generally speaking the network represents the environment). At each time t, before the environment reveals the new metric values, the forecaster predicts the new values based on previous observations. Contrary to classical methods where the environment evolution is characterized by stochastic process, we suppose that the environment evolution follows an unspecified mechanism, which could be deterministic, stochastic, or even adaptive to a given behavior. The prediction process should adapt to unpredictable network state changes due to its non-stationary nature. To properly address the adaptivity challenge, a special type of forecasters is used: the experts. These experts analyse the previous environment values, apply their own computation and make their own prediction. The experts predictions are given to the forecaster before the next environment values are revealed. The forecaster can then make its own prediction depending on the experts' "advice". The risk of a prediction may be defined as the value of a loss function measuring the discrepancy between the predicted value and the real environment value. The principal notion to optimize the behavior of the forecasters is the regret, seen as a difference between the forecaster's accumulated loss and that of each expert. To optimize the prediction process means to construct a forecasting strategy that guarantees a small loss with respect to defined experts. Adaptability of the forecaster is reflected in the manner in which it is able to follow the better expert according to the context. We intend to use and improve this prediction technique to design dynamically adaptive regret matching algorithms that will be applied to dynamically manage the resources in wireless networks, especially in sensor networks. These algorithms will allow the network to choose an optimal behavior, otherwise called a correlated equilibrium, from a defined behaviors' set. This behavior will be able to evolve in time to adapt to the network context evolution. We will focus mainly but not exclusively on applications like: the choice of communication channels depending on the predicted quality of transmission, energy efficiency, network nodes deployment, efficient routing, and intelligent switching between available technologies in a multi-technology context.

6.6. Vehicular Ad hoc NETworks (VANETs) for car merging

Participant: Paul Muhlethaler.

This is a joint work with Oyunchimeg Shagdar from the IMARA team.
Cooperative Adaptive Cruise Control (CACC) systems are intended to make driving safer and more efficient by utilizing information exchange between vehicles (V2V) and/or between vehicles and infrastructures (V2I). An important application of CACC is safe vehicle merging when vehicles join a main road, achieved by compiling information on the movement of individual main road vehicles. To support such road safety applications, the IEEE standardized the 802.11p amendment dedicated to V2V and V2I communications.

In this study, we have seek answers to the questions as to whether the IEEE 802.11p can support merging control and how the communications performance is translated into the CACC performance. We have built an analytical model of the IEEE 802.11p medium access control (MAC) for transmissions of the ETSI-standardized Cooperative Awareness Messages (CAM) and Decentralized Environmental Notification Messages (DENM) to support merging control. We have also developed a highway merging decision algorithm. Using computer simulations, packet delivery ratio (PDR), and packet inter-reception (PIR) time of IEEE 802.11p-based V2V and V2I communications and their impact on the CACC performance have been investigated. Our study has disclosed several useful insights including that PIR and throughput provide a good indication of the CACC performance, while improving PDR does not necessarily enhance the CACC performance. Moreover, thanks to its ability to reliably provide information at constant time intervals, the V2I structure offers a better support for CACC than V2V.

HYBRID Project-Team

6. New Results

6.1. 3D interactive techniques

6.1.1. Navigating in virtual environments with omnidirectional rendering

Participants: Jérôme Ardouin [contact], Anatole Lécuyer [contact], Maud Marchal.

The "FlyVIZ" enables humans to experience a real-time 360o vision of their surroundings for the first time. The visualization device combines a panoramic image acquisition system (positioned on top of the user's head) with a Head-Mounted Display (HMD). The omnidirectional images are transformed to fit the characteristics of HMD screens. As a result, the user can see his/her surroundings, in real-time, with 360o images mapped into the HMD field-of- view.

../../../../projets/hybrid/IMG/3DUI-FlyVIZ.png

Figure 2. The "FlyVIZ" enables humans to experience in real-time a 360-degree vision of their surroundings.

In order to safely simulate and evaluate our approach, we designed and evaluated [27] several visualization techniques, for navigating in virtual environments (VE). We have conducted an evaluation of different methods compared to a rendering method of reference, i.e. a perspective projection, in a basic navigation task. Our results confirm that using any omnidirectional rendering method could lead to more efficient navigation in terms of average task completion time. Among the different 360o projection methods, the subjective preference was significantly given to a cylindrical projection method (equirectangular). Taken together, our results suggest that omnidirectional rendering could be used in virtual reality applications in which fast navigation or full and rapid visual exploration are important. They pave the way to novel kinds of visual cues and visual rendering methods in virtual reality. This work was a collaboration with the Lagadic team (Inria Rennes).

6.1.2. Advances in locomotion interfaces for virtual environments

Participants: Anatole Lécuyer [contact], Maud Marchal [contact], Bruno Arnaldi.

Navigation, a fundamental task in Virtual Reality (VR), is greatly influenced by the locomotion interface being used, by the specificities of input and output devices, and by the way the virtual environment is represented. No matter how virtual walking is controlled, the generation of realistic virtual trajectories is absolutely required for some applications, especially those dedicated to the study of walking behaviors in VR, navigation through virtual places for architecture, rehabilitation and training.

First, we have studied the realism of unconstrained trajectories produced during virtual walking.We proposed a comprehensive evaluation framework consisting on a set of trajecto-graphical criteria and a locomotion model to generate reference trajectories [16]. We considered a simple locomotion task where users walk between two oriented points in space. The travel path was analyzed both geometrically and temporally in comparison to simulated reference trajectories. This work was a collaboration with the Mimetic team (Inria Rennes).

Secondly, we have introduced novel “Camera Motions” (CMs) to improve the sensations related to locomotion in virtual environments (VE) [26]. Traditional CMs are artificial oscillating motions applied to the subjective viewpoint when walking in the VE, and they are meant to evoke and reproduce the visual flow generated during a human walk. Our novel CMs are: (1) multistate, (2) personified, and (3) they can take into account the topography of the virtual terrain. In addition, they can then take into account avatar’s fatigue and recuperation, and the topography for updating visual CMs accordingly. Taken together, our results suggest that our new CMs could be introduced in Desktop VR applications involving first-person navigation, in order to enhance sensations of walking, running, and sprinting, with potentially different avatars and over uneven terrains, such as for training, virtual visits or video games.

6.1.3. 3D manipulation of virtual objects: 3-Point++

Participants: Thierry Duval [contact], Thi Thuong Huyen Nguyen.

Manipulation in immersive Virtual Environments (VEs) is often difficult and inaccurate because humans have difficulty in performing precise positioning tasks or in keeping the hand motionless in a particular position without any help of external devices or haptic feedback. To address this problem, we proposed a set of four manipulation points attached to objects (called a 3-Point++ tool, including three handle points and their barycenter), by which users can control and adjust the position of objects precisely [39]. By determining the relative position between the 3-Point++ tool and the objects, and by defining different states of each manipulation point (called locked/unlocked or inactive/active), these points can be freely configured to be adaptable and flexible to enable users to manipulate objects of varying sizes in many kinds of positioning scenarios.

6.1.4. A survey of 3D object selection techniques for virtual environments

Participant: Ferran Argelaguet Sanz [contact].

Computer graphics applications controlled through natural gestures are gaining increasing popularity these days due to recent developments in low-cost tracking systems and gesture recognition technologies. Although interaction techniques through naturalgestures have already demonstrated their benefits in manipulation, navigation and avatar-control tasks, effective selection with pointing gestures remains an open problem. We surveyed the state-of-the-art in 3D object selection techniques [13]. We reviewed important findings in human control models, analyze major factors influencing selection performance, and classify existing techniques according to a number of criteria. Unlike other components of the application's user interface, pointing techniques need a close coupling with the rendering pipeline, introducing new elements to be drawn, and potentially modifying the object layout and the way the scene is rendered. Conversely, selection performance is affected by rendering issues such as visual feedback, depth perception, and occlusion management. We thus reviewed existing literature paying special attention to those aspects in the boundary between computer graphics and human computer interaction.

6.1.5. Novel pseudo-haptic based interfaces

Participants: Pierre Gaucher, Ferran Argelaguet Sanz, Anatole Lécuyer [contact], Maud Marchal.

Pseudo-haptics is a technique meant to simulate haptic sensations using visual feedback and properties of human visuo-haptic perception. In this course of action, we have extended its usage for gestural interfaces [32] and exploring its usage for the simulation of the local elasticity of images [].

Interacting with virtual objects through free-hand gestures do not allow users to perceive the physical properties of virtual objects. To provide enhanced interaction, we explored how the usage of a pseudo-haptic approach could be introduced while interacting with a 3D Carrousel [32]. In our approach, which is envisioned for showcasting purposes, virtual products are presented using a 3D carousel augmented with physical behavior and a pseudo-haptic effect aiming to attract the user to specific items. The user, through simple gestures, controls the rotation of the carousel, and can select, examine and manipulate the objects presented. Several demos can be tested on-line at Hybrid website.

Secondly, we have introduced the Elastic Images, a novel pseudo-haptic feedback technique which enables the perception of the local elasticity of images without the need of any haptic device []. The proposed approach focuses on whether visual feedback is able to induce a sensation of stiffness when the user interacts with an image using a standard mouse. The user, when clicking on a Elastic Image, is able to deform it locally according to its elastic properties. A psychophysical experiment was conducted to quantify this novel pseudo-haptic perception and determine its perceptual threshold (or its Just Noticeable Difference). The results showed that users were able to recognize up to eight different stiffness values with our method and confirmed that it provides a perceivable and exploitable sensation of elasticity.

6.1.6. Experiencing the past in virtual reality

Participant: Valérie Gouranton [contact].

We designed a public experience and exhibition organized during the French National Days of Archaeology. This was the result of an interdisciplinary collaboration between archaeologists and computer scientists, centered on the immersive virtual reality platform Immersia, a node of the European Visionair project. This public exhibition had three main goals: (i) presenting our interdisciplinary collaboration, (ii) communicating on the scientific results of this collaboration, and (iii) offering an immersive experience in the past for visitors. In [33] we could present the scientific context of the event, its organization, and a discussion on feedbacks.

In the frame of the CNPAO project (section 8.1.3) we have also worked on the reconstitution of six archaeological sites located in the west of France ranging from prehistory to the Middle Ages: the Cairn of Carn Island, the covered pathway of Roh Coh Coet, the GohMin Ru megalithic site, the gallo-roman mansion of Vanesia, the keep of the Château de Sainte-Suzanne, the Porte des Champs of the Château d'Angers. Other proposals are currently under study [29].

6.1.7. Perception of affordances in virtual reality

Participants: Anatole Lécuyer [contact], Maud Marchal.

../../../../projets/hybrid/IMG/3DUI-TouchingThePast.png

Figure 3. "Touching the past" experience during the French National Days of Archaeology.

The perception of affordances could be a potential tool for sensorimotor assessment of physical presence, that is, the feeling of being physically located in a virtual place. We have evaluated the perception of affordances for standing on a virtual slanted surface [25]. Participants were asked to judge whether a virtual slanted surface supported up right stance. The objective was to evaluate if this perception was possible in virtual reality (VR) and comparable to previous works conducted in real environments. We found that the perception of affordances for standing on a slanted surface in virtual reality is possible and comparable (with an underestimation) to previous studies conducted in real environments. We also found that participants were able to extract and to use virtual information about friction in order to judge whether a slanted surface supported an upright stance. Finally, results revealed that the person's position on the slanted surface is involved in the perception of affordances for standing on virtual grounds. Taken together, our results show quantitatively that the perception of affordances can be effective in virtual environments, and influenced by both environmental and person properties. Such a perceptual evaluation of affordances in VR could guide VE designers to improve their designs and to better understand the effect of these designs on VE users.

6.2. Haptic Feedback and Physical Simulation

6.2.1. Haptic feedback to improve audiovisual experience

Participants: Fabien Danieau, Anatole Lécuyer [contact].

Haptics have been employed in a wide set of applications ranging from teleoperation and medical simulation to arts and design, including entertainment, aircraft simulation and virtual reality. As for today, there is also a growing attention from the research community on how haptic feedback can be integrated with profit to audiovisual systems. We have first reviewed [18] the techniques, formalisms and key results on the enhancement of audiovisual experience with haptic feedback. We first reviewed the three main stages in the pipeline which are (i) production of haptic effects, (ii) distribution of haptic effects and (iii) rendering of haptic effects. We then highlighted the strong necessity for evaluation techniques in this context and discuss the key challenges in the field. By building on technology and results from virtual reality, and tackling the specific challenges in the enhancement of audiovisual experience with haptics, we believe the field presents exciting research perspectives for which financial and societal stakes are significant.

We have then developped a novel approach called HapSeat for simulating motion sensations in a consumer environment.Multiple force-feedbacks are applied to the seated user's body to generate a 6DoF sensation of motion while experiencing passive navigation as illustrated Figure 4 . A set of force-feedback devices such as mobile armrests or headrests are arranged around a seat so that they can apply forces to the user. The forces are computed consistently with the visual content (visual acceleration) in order to generate motion sensations. This novel display device has been patented and was demonstrated this year at ACM SIGGRAPH 2013 Emerging-Technologies [55], and ACM CHI 2013 Interactivity [54].

../../../../projets/hybrid/IMG/hapseat.png

Figure 4. The HapSeat device: force-feedback is applied on the user's body with mobile armrests or headrests in order to generate motion sensations that are consistent with the visual content.

This work was a collaboration with the Mimetic team (Inria Rennes).

6.2.2. Vibrotactile rendering of splashing fluids

Participants: Anatole Lécuyer, Maud Marchal [contact].

Compelling virtual reality scenarios involving physically based virtual materials have been demonstrated using hand- based and foot-based interaction with visual and vibrotactile feedback. However, some materials, such as water and other fluids, have been largely ignored in this context. For VR simulations of real-world environments, the inability to include interaction with fluids is a significant limitation. Potential applications include improved training involving fluids, such as medical and phobia simulators, and enhanced user experience in entertainment, such as when interacting with water in immersive virtual worlds. We introduced the use of vibrotactile feedback as a rendering modality for solid-fluid interaction, based on the physical processes that generate sound during such interactions [15]. This rendering approach enables the perception of vibrotactile feedback from virtual scenarios that resemble the experience of stepping into a water puddle or plunging a hand into a volume of fluid.

6.2.3. Six-DoF haptic interaction with fluids, solids, and their transitions

Participants: Anatole Lécuyer, Maud Marchal [contact].

Haptic interaction with different types of materials in the same scene is a challenging task, mainly due to the specific coupling mechanisms that are usually required for either fluid, deformable or rigid media. Dynamically-changing materials, such as melting or freezing objects, present additional challenges by adding another layer of complexity in the interaction between the scene and the haptic proxy. We have addressed these issues through a common simulation framework, based on Smoothed-Particle Hydrodynamics, and enable haptic interaction simultaneously with fluid, elastic and rigid bodies, as well as their melting or freezing [30]. We introduced a mechanism to deal with state changes, allowing the perception of haptic feedback during the process, and a set of dynamic mechanisms to enrich the interaction through the proxy. We decouple the haptic and visual loops through a dual GPU implementation. An initial evaluation of the approach was performed through performance and feedback measurements, as well as a small user study assessing the capability of users to recognize the different states of matter they interact with.

6.2.4. Bimanual haptic manipulation

Participants: Anatole Lécuyer [contact], Maud Marchal [contact], Anthony Talvas.

Bimanual haptics is a specific kind of multi-finger interaction that focuses on the use of both hands simultaneously. Several haptic devices enable bimanual haptic interaction, but they are subject to a certain number of limitations for interacting with virtual environments (VEs), such as workspace size issues or manipulation difficulties, notably with single-point interfaces. Interaction techniques exist to overcome these limitations and allow users to perform specific two-handed tasks, such as the bimanual exploration of large VEs and grasping of virtual objects. We have proposed an overview of the current limitations in bimanual haptics and the interaction techniques developed to overcome them. Novel techniques based on the Bubble technique are more specifically presented, with a user evaluation that assesses their efficiency. These include bimanual workspace extension techniques as well as techniques to improve the grasping of virtual objects with dual single-point interfaces. This work was published as a chapter in a book on "Multi-finger Haptic Interaction" [51].

6.2.5. The god-finger method

Participants: Anatole Lécuyer, Maud Marchal [contact], Anthony Talvas.

In physically-based virtual environments, interaction with objects generally happens through contact points that barely represent the area of contact between the user's hand and the virtual object. This representation of contacts contrasts with real life situations where our finger pads have the ability to deform slightly to match the shape of a touched object. We have proposed a method called god-finger to simulate a contact area from a single contact point determined by collision detection, and usable in a rigid body physics engine [42]. The method uses the geometry of the object and the force applied to it to determine additional contact points that will emulate the presence of a contact area between the user's proxy and the virtual object. It could improve the manipulation of objects by constraining the rotation of touched objects in a similar manner to actual finger pads. An implementation in a physics engine shows that the method could make for more realistic behaviour when manipulating objects while keeping high simulation rates. This work was presented at IEEE 3DUI Symposium 2013 and has received the best technote award [42].

6.2.6. Collision detection for fracturing rigid bodies

Participant: Maud Marchal [contact].

../../../../projets/hybrid/IMG/FigureBimanual.png

Figure 5. 3D interaction techniques for bimanual haptic manipulation.

In complex scenes with many objects, collision detection plays a key role in the simulation performance. This is particularly true for fracture simulation, where multiple new objects are dynamically created. We have proposed novel algorithms and data structures for collision detection in real-time brittle fracture simulations [21]. We build on a combination of well-known efficient data structures, namely distance fields and sphere trees, making our algorithm easy to integrate on existing simulation engines. We proposed novel methods to construct these data structures, such that they can be efficiently updated upon fracture events and integrated in a simple yet effective self-adapting contact selection algorithm. Altogether, we drastically reduced the cost of both collision detection and collision response. We have evaluated our global solution for collision detection on challenging scenarios, achieving high frame rates suited for hard real-time applications such as video games or haptics. Our solution opens promising perspectives for complex brittle fracture simulations involving many dynamically created objects.

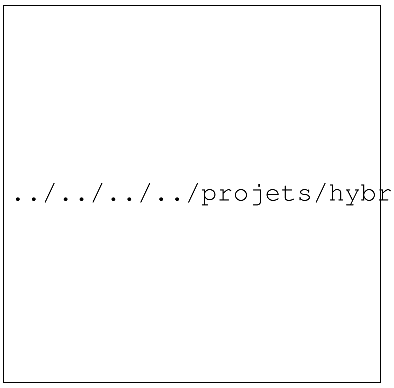


Figure 6. Example of brittle fracture with collision detection.

This work was a collaboration with the Mimetic team (Inria Rennes).

6.2.7. Collision detection with high performance computing on GPU

Participants: Bruno Arnaldi, Valérie Gouranton [contact], François Lehericey.

We have first proposed IRTCD, a novel Iterative Ray-Traced Collision Detection algorithm that exploits spatial and temporal coherency. Our approach uses any existing standard ray-tracing algorithm and we propose an iterative algorithm that updates the previous time step results at a lower cost with some approximations. Applied for rigid bodies, our iterative algorithm accelerate the collision detection by a speedup up to 33 times compared to non-iterative algorithms on GPU [34].

Then, we have presented two methods to efficiently control and reduce the interpenetration without noticeable computation overhead. The first method predicts the next potentially colliding vertices. These predictions are used to make our IRTCD algorithm more robust to the approximations, therefore reducing the errors up to 91%. We also present a ray re-projection algorithm that improves the physical response of ray-traced collision detection algorithm. This algorithm also reduces, up to 52%, the interpenetration between objects in a virtual environment. Our last contribution showed that our algorithm, when implemented on multi-GPUs architectures, is far faster [35].

Finally, we proposed a distributed and anticipative model for collision detection and propose a lead for distributed collision handling, two key components of physically-based simulations of virtual environments. This model is designed to improve the scalability of interactive deterministic simulations on distributed systems such as PC clusters. Our main contribution consists of loosening synchronism constraints in the collision detection and response pipeline to allow the simulation to run in a decentralized, distributed fashion.

../../../../projets/hybrid/IMG/collision.png

Figure 7. Real-time simulation of iterative ray-traced collision detection algorithm.

We could show the potential for distributed load balancing strategies based on the exchange of grid cells, and explain how anticipative computing may, in cases of short computational peaks, improve user experience by avoiding frame-rate drop-downs [31].

6.3. Brain-Computer Interfaces and Virtual Environments

6.3.1. Multi-user BCI video game

Participant: Anatole Lécuyer [contact].

How can we connect two brains to a video game by means of a BCI, and what will happen when we do so? How will the two users behave, and how will they perceive this novel common experience? We have created a multi-user videogame called “BrainArena” in which two users can play a simple football game by means of two BCIs [14], as illustrated Figure 8 . They can score goals on the left or right side of the screen by simply imagining left or right hand movements. To add another interesting element , the gamers can play in a collaborative manner (their two mental activities are combined to score in the same goal), or in a competitive manner (the gamers must push the ball in opposite directions). Two experiments were conducted to evaluate the performance and subjective experience of users in the different conditions. Taken together our results suggest that multi-user BCI applications can be operational, effective, and more engaging for participants.

This work was a collaboration with the Potioc team (Inria Bordeaux).

../../../../projets/hybrid/IMG/BrainArena.png

Figure 8. Multi-user football videogame in which two players can score goals to the left or right by imagining left or right hand movements. The users can play together using their brain activities either in a collaboration mode (same goal) or in a competitive mode (one versus the other).

6.3.2. Contextual SSVEP-based BCI control

Participants: Jozef Legény, Anatole Lécuyer [contact].

One main disadvantage of Brain-Computer Interfaces is that they are not completely reliable. In order to increase BCI performances, some adjustments can be made on low levels, such as signal processing and on high levels by modifying the controller paradigm. We have explored a novel, context-dependent, approach for SSVEP-based BCI controller [22]. This controller uses two kinds of behaviour alternation, commands can be added and removed if their use is irrelevant to the context or the actions resulting from their activation can be weighted depending on the likeliness of the actual intention of the user. This controller has been integrated within a BCI computer game and its influence in performance and mental workload has been addressed through a pilot experiment. Preliminary results have shown a workload reduction and performance improvement with the context-dependent controller while keeping the engagement levels untouched.

This work was a collaboration with the Universidad de Jaen (Spain).

6.3.3. Can we use a BCI and manipulate a mouse at the same time?

Participants: Jonathan Mercier-Ganady, Anatole Lécuyer [contact], Maud Marchal.

In most setups using a BCI, the user is explicitly asked to remain as motionless as possible, since muscular activity is commonly admitted to add noise and artifacts in brain electrical signals. Thus, as for today, people have been rarely let using other classical input devices such as mice or joysticks simultaneously to a BCI-based interaction. We have conducted an experimental study on the influence of manipulating an input device such as a standard computer mouse on the performance of a BCI system [37]. The study uses a simple virtual environment inspired by the well-known Pac-Man videogame and based on BCI and mouse controls. As expected the BCI performance was found to slightly decrease in presence of motor activity. However, we found that the BCI could still be successfully used in all conditions, even in presence of a highly-demanding mouse manipulation. These promising results pave the way to future experimental studies with more complex mental and motor activities, but also to novel 3D interaction paradigms that could mix BCI and other input devices for virtual reality and videogame applications.

6.3.4. Adaptive VR simulators combining visual, haptic, and BCIs

Participants: Anatole Lécuyer [contact], Maud Marchal.

What if the next generation of virtual reality simulators would take into account a novel user's input: his/her mental state, as measured with electrodes and Brain-Computer Interfaces ? This would lead to adaptive simulators that could match the “hidden” expectations of the user optimally? We have initiated and illustrated this promising path with a virtual reality setup in which the force-feedback of a guidance system is adapted in real-time to the “mental workload” of the user [23]. A first application of this approach is a medical simulator in which virtual assistances are automatically adapted to surgeon and trainee's mental activity as illustrated Figure 9 . Such results pave the way to future virtual reality systems which would automatically reconfigure and adapt to cerebral inputs and cognitive processes.

6.4. Collaborative Virtual Environments

6.4.1. Collaborative exploration in multi-scale shared virtual environments

Participants: Thierry Duval [contact], Thi Thuong Huyen Nguyen.

Exploration of large-scale 3D Virtual Environments (VEs) is often difficult because of lack of familiarity with complex virtual worlds, lack of spatial information that can be offered to users and lack of sensory details compared to the exploration of real environments. To address this problem, we presented a set of metaphors for assisting users in collaborative navigation to perform common exploration tasks in shared collaborative virtual environments [38], [56]. Our propositions consist in three guiding techniques in the form of navigation aids to enable one or several users to help one main user (exploring user) to explore the VE efficiently. These three techniques consist in drawing directional arrows, lighting up path to follow, and orienting a compass to show a direction to the exploring user. Our experimental results could show that although the directional arrows and compass surpassed the light source in a navigation task, these three techniques are completely appropriate for guiding a user in 3D complex VEs.

../../../../projets/hybrid/IMG/medicalBCI.png

Figure 9. Medical simulator adapted to a BCI output. The user manipulates a virtual needle and has to insert it into a virtual liver to reach a tumor. Visual and haptic assistances are activated when a high mental workload is detected which corresponds to a more difficult manipulation of the needle.

6.4.2. Improving the awareness of collaboration in 3D virtual environments

Participants: Thierry Duval [contact], Thi Thuong Huyen Nguyen, Valérie Gouranton.

When a user is fully immersed within a Virtual Environments (VE) through a large immersive display system, his feeling of presence can be altered because of disturbing interactions with his physical environment, such as collision with hardware parts of the system or loss of tracking. This alteration can be avoided by taking into account the physical features of the user and to embed them in the VE. In [19] we could present how we use the Immersive Interactive Virtual Cabin (IIVC) model to obtain such a virtual representation of the physical environment of the user and we illustrated how it can be used to guide efficiently a user for a navigation task in a VE. We also presented how we can add 3D representations of 2D interaction tools in order to cope with asymmetrical collaborative configurations, providing 3D cues for users in order to understand the actions of the other users even if they are not fully immersed in the shared virtual environment. Last, we explained how we could enhance 3D interaction and collaboration by embedding a symbolic 3D representation of the user that would give 3D information about his posture.

6.4.3. Sharing and bridging information: application to ergonomics

Participant: Thierry Duval [contact].

We introduced a collaborative virtual environment usable to conduct ergonomic design sessions, involving the worker, ergonomists and engineers [40]. We focused particularly on the representation of the ergonomic evaluation and the interaction between an ergonomist and the main user (worker). An ergonomic evaluation of the postures was presented. An interaction architecture between the main user and an ergonomist based on the combination of animation modes of two linked manikins was also proposed. Preliminary results and future developments of the CVE (e.g. additional ergonomic evaluation tools, graphical enhancement, interaction enhancement) were then presented.

6.4.4. User embodiment and collaboration in virtual environments for training

Participants: Bruno Arnaldi, Valérie Gouranton [contact], Thomas Lopez, Florian Nouviale, Rozenn Bouville Berthelot.

In Collaborative Virtual Environments for Training (CVET), a group can learn and practice the completion of a task as a team using all the assets provided by Virtual Reality. We presented a novel mechanism that allows real and virtual humans to dynamically exchange the control of their embodiment in virtual environments [41]. Such a mechanism raises two important issues: the possibility of dynamic embodiment exchanges between real humans and virtual humans and the continuity of actions of the team members after an exchange. To address these issues we introduce a new entity, the Perceptive Puppet that abstracts real and virtual humans into one common entity containing its own knowledge.

In addition, in CVET different roles need to be played by actors, i.e. virtual agents or users. In order to abstract an actor from its embodiment in the virtual world, we have introduced a new entity, the Shell [36]. Through the Shell, users and virtual agents are able to collaborate in the same manner during the training. In addition to the embodiment's control, the Shell gathers and carries knowledge and provides interaction inputs. This knowledge and those inputs can be accessed and used homogeneously by both users and virtual agents to help them to perform the procedure.

Hycomes Team

5. New Results

5.1. Hybrid Systems Modeling

Participants: Albert Benveniste, Benoît Caillaud.

5.1.1. Type-Based Analysis of Causality Loops In Hybrid Systems Modelers

Explicit hybrid systems modelers like Simulink / Stateflow allow for programming both discrete- and continuous-time behaviors with complex interactions between them. A key issue in their compilation is the static detection of algebraic or causality loops. Such loops can cause simulations to deadlock and prevent the generation of statically scheduled code. We have addressed this issue for a hybrid modeling language that combines synchronous Lustre-like data-flow equations with Ordinary Differential Equations (ODEs) [6], [9]. We introduce the operator $\mathrm{last}(x)$ for the left-limit of a signal x. This operator is used to break causality loops and permits a uniform treatment of discrete and continuous state variables. The semantics relies on non-standard analysis, defining an execution as a sequence of infinitesimally small steps. A signal is deemed causally correct when it can be computed sequentially and only progresses by infinitesimal steps outside of discrete events. The causality analysis takes the form of a simple type system. In well-typed programs, signals are proved continuous during integration and can be translated into sequential code for integration with off-the-shelf ODE solvers. The effectiveness of this system is illustrated with several examples written in Zélus [9], a Lustre-like synchronous language extended with hierarchical automata and ODEs.

5.1.2. Semantics of multi-mode DAE systems

Hybrid systems modelers exhibit a number of difficulties related to the mix of continuous and discrete dynamics and sensitivity to the discretization scheme. Modular modeling, where subsystems models can be simply assembled with no rework, calls for using Differential Algebraic Equations (DAE). In turn, DAE are strictly more difficult than ODE. They require sophisticated pre-processing using various notions of index before they can be submitted to a solver. We have studied some fundamental issues raised by the modeling and simulation of hybrid systems involving DAEs [10]. The objective of this work is to serve for the evolution and the design of future releases of the Modelica language for such systems. We focus on the following questions:

- What is the proper notion of index for a hybrid DAE system?
- What are the primitive statements needed for a DAE hybrid systems modeler?

The differentiation index for DAE explicitly relies on everything being differentiable. Therefore, generalizations to hybrid systems must be done with caution. We propose to rely on non-standard analysis for this. Non-standard analysis formalizes differential equations as discrete step transition systems with infinitesimal time basis. We can thus bring hybrid DAE systems to their non-standard form, where the notion of difference index can be firmly used. From this study, general hints for future releases of Modelica can be drawn.

5.2. Surgical Process Mining with Test and Flip Net Synthesis

Participant: Benoît Caillaud.

Surgical process modeling aims at providing an explicit representation of surgical procedural knowledge. *Surgical process models* are inferred from a set of surgical procedure recordings, and represent in a concise manner concurrency, causality and conflict relations between actions. In the context of the S3PM project (Section 6.1), we have investigated the use of *test and flip* nets, a mild extension of flip-flop nets, to represent surgical process models. A test and flip net synthesis algorithm, based on linear algebraic methods in the $Z/2Z$ ring is detailed. Experimental results regarding the use of this synthesis algorithm to automate the construction of simple surgical process models are also presented.

[9] http://zelus.di.ens.fr

I4S Project-Team

5. New Results

5.1. identification of linear systems

5.1.1. Evaluation of confidence intervals and computation of sensitivities for subspace methods

Participants: Michael Doehler, Laurent Mevel.

Stochastic Subspace Identification methods have been extensively used for the modal analysis of mechanical, civil or aeronautical structures for the last ten years. So-called stabilization diagrams are used, where modal parameters are estimated at successive model orders, leading to a graphical procedure where the physical modes of the system are extracted and separated from spurious modes. Recently an uncertainty computation scheme has been derived allowing the computation of uncertainty bounds for modal parameters at some given model order. In this paper, two problems are addressed. Firstly, a fast computation scheme is proposed reducing the computational burden of the uncertainty computation scheme by an order of magnitude in the model order compared to a direct implementation. Secondly, a new algorithm is proposed to derive the uncertainty bounds for the estimated modes at all model orders in the stabilization diagram. It is shown that this new algorithm is both computationally and memory efficient, reducing the computational burden by two orders of magnitude in the model order[14].

5.1.2. Subspace methods in frequency domain

Participants: Philippe Mellinger, Michael Doehler, Laurent Mevel.

In this paper a combined subspace algorithm and a way to quantify uncertainties of its resulting identified modal parameter has been presented. Even if the algorithm is data-driven, it was proven that uncertainties can still be quantified by using the square subspace matrix without any modification neither on the identified modal parameters or on the stabilization diagrams. A comparison between uncertainty quantification based on this data-driven combined subspace algorithm and the well-known covariance-driven stochastic subspace algorithm shows good results on this new method. Both values and confidence intervals are similar. However combined algorithm gives better results considering spurious modes. [27].

5.1.3. Subspace Identification for Linear Periodically Time-varying Systems

Participants: Laurent Mevel, Ahmed Jhinaoui.

Many systems such as turbo-generators, wind turbines and helicopters show intrinsic time-periodic behaviors. Usually, these structures are considered to be faithfully modeled as Linear Time-Invariant (LTI). In some cases where the rotor is anisotropic, this modeling does not hold and the equations of motion lead necessarily to a Linear Periodically Time-Varying (referred to as LPTV in the control and digital signal field or LTP in the mechanical and nonlinear dynamics world) model. Classical modal analysis methodologies based on the classical time-invariant eigenstructure (frequencies and damping ratios) of the system no more apply. This is the case in particular for subspace methods. For such time-periodic systems, the modal analysis can be described by characteristic exponents called Floquet multipliers. The aim of this paper is to suggest a new subspace-based algorithm that is able to extract these multipliers and the corresponding frequencies and damping ratios. The algorithm is then tested on a numerical model of a hinged-bladed helicopter on the ground. [22], [23], [18].

5.2. damage detection for mechanical structures

5.2.1. Damage detection and localisation

Participants: Michael Doehler, Luciano Gallegos, Laurent Mevel.

Mechanical systems under vibration excitation are prime candidate for being modeled by linear time invariant systems. Damage detection in such systems relates to the monitoring of the changes in the eigenstructure of the corresponding linear system, and thus reflects changes in modal parameters (frequencies, damping, mode shapes) and finally in the finite element model of the structure. Damage localization using both finite element information and modal parameters estimated from ambient vibration data collected from sensors is possible by the Stochastic Dynamic Damage Location Vector (SDDLV) approach. Damage is related to some residual derived from the kernel of the difference between transfer matrices in both reference and damage states and a model of the reference state. Deciding that this residual is zero is up to now done using an empirically defined threshold. In this paper, we show how the uncertainty in the estimates of the state space system can be used to derive uncertainty bounds on the damage localization residuals to decide about the damage location with a hypothesis test.[13], [21], [26].

5.2.2. Robust subspace damage detection

Participants: Michael Doehler, Laurent Mevel.

The detection of changes in the eigenstructure of a linear time invariant system by means of a subspace-based residual function has been proposed previously. While enjoying some success in its applicability in particular in the context of vibration monitoring, the robustness of this framework against changes in the noise properties has not been properly addressed yet. In this paper, a new robust residual is proposed and the robustness of its statistics against changes in the noise covariances is shown. The complete theory for hypothesis testing for fault detection is derived and a numerical illustration is provided[16].

5.2.2.1. Feasibility of reflectometry techniques for non destructive evaluation of external post-tensioned cables

Participant: Qinghua Zhang.

Nowadays a considerable number of bridges is reaching an age when renovating operations become necessary. For some bridges, external post-tension is realized with cables protected in ducts, with the residual internal space imperfectly filled with a fluid cement grout. Detecting the problems of injection in the ducts is visually impossible from the outside. Through a collaboration with the SISYPHE project-team, the feasibility of reflectometry techniques for cable health monitoring is investigated via numerical simulations and laboratory experiments. The main idea consists in adding electrically conductive tapes along a duct so that the duct and the added tapes can be treated as an electrical transmission line. It is then possible to apply advanced reflectometry methods developed by the SISYPHE project-team, initially for true electric cables.

IBIS Project-Team

5. New Results

5.1. Analysis of gene regulatory networks by means of piecewise-linear (PL) models

GENETIC NETWORK ANALYZER (GNA) is a tool for the qualitative modeling and simulation of the dynamics of gene regulatory networks by means of PLDE models, as described in Section 4.1 . GNA has been integrated with the other bioinformatics tools distributed by Genostar (http://www.genostar.com/). Version 8.5 of GNA was released by IBIS and Genostar this year. This version is an update of version 8.0, deposited at the Agence pour la Protection des Programmes (APP). Some bugs have been corrected in the new version and the program has been adapted to the latest versions of Java and the software platform of Genostar. Version 8.5 supports the SBML standard and is also capable of exporting its models to the newly-developed standard for qualitative models, SBML Qual. This standard has been elaborated by the community of developers of logical and related modeling tools (CoLoMoTo), in which the GNA developers participate, and has been described in detail in a paper published in *BMC Systems Biology* [6].

The predictions obtained with the help of GNA are purely qualitative, describing the dynamics of the network by means of a state transition graph. While a qualitative analysis is appropriate for certain problems, the absence of precise quantitative predictions may not be desirable in others, such as the analysis of a limit cycle or the design of a controller for a synthetic network. The quantitative study of PLDE models of gene regulatory networks is hindered by the fact that the step functions describing the logic of regulatory interactions lead to discontinuities in the right-hand side of the PLDE models (Section 3.1). This has motivated extensions of the PLDE models based on differential inclusions and Filippov solutions. As of now, no numerical simulation tool for the simulation of these Filippov extensions is available.

In collaboration with Vincent Acary and Bernard Brogliato of the BIPOP project-team, we have shown how tools developed for the simulation of nonsmooth mechanical, electrical and control systems can be adapted for this purpose, in a paper published in *Physica D* [2]. We have presented a method for the numerical analysis of one proposed extension, called Aizerman–Pyatnitskii (AP)-extension, by reformulating the PLDE models as mixed complementarity systems (MCSs). This allows the application of powerful methods developed for this class of nonsmooth dynamical systems, in particular those implemented in the SICONOS PLATFORM developed by BIPOP. We have also shown that under a set of reasonable biological assumptions, putting constraints on the right-hand side of the PLDE models, AP-extensions and classical Filippov extensions are equivalent. This means that the proposed numerical method is valid for a range of different solution concepts. We have illustrated the practical interest of our approach through the numerical analysis of three well-known networks developed in the field of synthetic biology.

5.2. Inference of bacterial regulatory networks from reporter gene data

The use of fluorescent and luminescent reporter genes allows real-time monitoring of gene expression, both at the level of individual cells and cell populations (Section 3.2). In order to fully exploit this technology, we need methods to rapidly construct reporter genes, both on plasmids and on the chromosome, mathematical models to infer biologically relevant quantities from the primary data, and computer tools to achieve this in an efficient and user-friendly manner. For instance, in a typical microplate experiment, 96 cultures are followed in parallel, over several hours, resulting in 10,000-100,000 measurements of absorbance and fluorescence and luminescence intensities. Over the past few years, we put into place an experimental platform and data analysis software, notably the WELLREADER program (Section 4.2), to allow biologists to make the most out of the information contained in reporter gene expression data. A web-based version of WELLREADER is currently in preparation. Valentin Zulkower has analyzed the measurement models underlying WELLREADER, work was presented at the *Journées Ouvertes Biologie, Informatique et Mathématiques (JOBIM'13)* [13] and submitted for publication.

The above tools have been used in a series of studies directed at the experimental mapping of gene regulatory networks in *E. coli*. A first example, which was carried out in the framework of the PhD thesis of former IBIS member Guillaume Baptist, concerns the development of a new screening methodology for identifying all genes that control the expression of a target gene through genetic or metabolic interactions. The screen combines mutant libraries with luciferase reporter constructs. Instead of a static picture of gene expression, this method allows dynamical monitoring in different environmental conditions. Mutants with interesting phenotypes can thus be selected based on multiple criteria, and the expression dynamics of the target gene can be extensively characterized. The method has been applied to the identification of the direct and indirect regulators of the gene *acs* in *Escherichia coli*. We confirmed known genetic regulators of the gene and identified new regulatory influences, many of which involve metabolic intermediates or metabolic sensing. An analysis of mutants involved in glycolysis and glucose transport demonstrates that the classical model of catabolite repression in *E. coli* needs to be amended. A paper describing the above work was published in *Nucleic Acids Research* this year [3].

A second example is a study, led by Stéphan Lacour in collaboration with Akira Ishihama and Hiroshi Ogasawara in Japan, on the lifestyle adaptation of *E. coli*. The study concerns the switch between swimming motility and biofilm formation in response to changes in environmental growth conditions. The stationary phase sigma factor RpoS is an important regulator of this switch since it stimulates adhesion and represses flagellar biosynthesis. By measuring the dynamics of gene expression, we show that RpoS inhibits the transcription of the flagellar sigma factor, FliA, in exponential growth phase. RpoS also partially controls the expression of CsgD and CpxR, two transcription factors important for bacterial adhesion. We have demonstrated that these two regulators repress the transcription of *fliA*, *flgM* and *tar*, and that this regulation is dependent on the growth medium. CsgD binds to the flgM and fliA promoters around their -10 promoter element, strongly suggesting direct repression. The results show that CsgD and CpxR also affect the expression of other known modulators of cell motility. An updated structure of the regulatory network controlling the choice between adhesion and motility was proposed in the paper based on this work, published in the *Journal of Bacteriology* [7].

A third study, published in *Research in Microbiology* [8], also focuses on the alternative sigma factor RpoS. The small protein Crl increases the interaction between RpoS and RNA polymerase and thereby activates certain RpoS-dependent promoters. However, the growth-phase dependence of the interaction of Crl with different forms of polymerase remains unknown. We have used 41 GFP transcriptional fusions to study the dynamics of gene regulation by RpoS and Crl during growth transition from exponential to stationary phase in *Escherichia coli*. This has confirmed that RpoS can regulate gene expression in exponential phase, both positively and negatively. Crl slightly stimulates transcription by RpoS in exponential phase and controls a subset of RpoS-dependent genes in stationary phase. Growth temperature strongly affects induction of specific promoters by RpoS, whereas its impact on gene regulation by Crl is much less significant. In addition, we have identified five new genes regulated by Crl (*ada*, *cbpA*, *glgS*, *sodC* and *flgM*), and demonstrated that Crl improves promoter binding and opening by RpoS-containing RNA polymerase at the *hdeA* promoter. The study also shows that Crl is a cognate enhancer of RpoS activity under different growth conditions, since its deletion has no effect on genes transcribed by other sigma factors.

In the framework of the PhD thesis of Diana Stefan, a network inference method developed by Eugenio Cinquemani and colleagues, first published in *Bioinformatics* in 2010, has been applied to reporter gene data from the network regulating motility of *E. coli*, described above. The results are currently being prepared for publication.

5.3. Models of carbon metabolism in bacteria

Kinetic models capture the dynamics of the large and complex networks of biochemical reactions that endow bacteria with the capacity to adapt their functioning to changes in the environment. In collaboration with Matteo Brilli and Daniel Kahn (Inra and Université Claude Bernard in Lyon), we previously developed an approximate model of central metabolism of *E. coli*, using linlog kinetics, and estimated the parameter values from metabolomics, transcriptome, proteomics data sets, as described in an article published in *Bioinformatics* in 2011. The results of this study revealed the fundamental role played by the identifiability of the model parameters, an issue often overlooked in systems biology. This prompted us for a thorough investigation of the concepts of structural identifiability (in presence of perfect, idealized data), practical identifiability (in presence of noisy and limited amounts of data), and the relations between the two. In addition, we looked into the implications of this analysis for the reduction of nonidentifable to identifiable models. While having a solid mathematical basis, the study was tailored to the actual experimental practice, and resulted in a practical model reduction method that improves upon our previous approach in case of large measurement noise. This study, and the results from its application to both *in-silico* case studies and state-of-the-art datasets, were reported in a paper that appeared in the *Journal of Mathematical Biology* this year [4]. Although the theoretical development has focused on linlog models and related classes of approximate kinetic models, it is important to note that the results also bear on more general classes of nonlinear models of metabolism.

A second line of work is based on the use of classical kinetic models that are, in comparison with the above-mentioned linlog models, much reduced in scope (the focus is on the metabolic and genetic regulation of the glycolysis pathway) and granularity (individual reactions are lumped together). The models, developed by Delphine Ropers, are being calibrated using experimental data from IBIS group and the group of Jean-Charles Portais at Inra/INSA in Toulouse, and will be used to understand some key mechanisms in the adaptation of *E. coli* to the exhaustion of glucose. The PhD thesis of Manon Morin, in the framework of a collaboration supported by a Contrat Jeune Scientifique Inra-Inria, will further develop these research directions. In the framework of their PhD theses, Stéphane Pinhal and Valentin Zulkower also study specific aspects of carbon metabolism, using both models and experimental data. In parallel, we collaborate with Myriam Ferro at CEA in Grenoble to investigate how state-of-the-art measurements of the absolute concentrations of enzymes in *E. coli* can be integrated with other high-throughput data sets and kinetic models. The first results of this collaboration were accepted for publication in *Molecular and Cellular Proteomics* early 2014.

5.4. Stochastic modeling and identification of gene regulatory networks in bacteria

At the single-cell level, the processes that govern gene expression are often better described by stochastic models. Modern techniques for the real-time monitoring of gene expression in single cells enable one to apply stochastic modelling to study the origins and consequences of random noise in response to various environmental stresses, and the emergence of phenotypic variability. The potential impact of single-cell stochastic analysis and modelling ranges from a better comprehension of the biochemical regulatory mechanisms underlying cellular phenotypes to the development of new strategies for the control of cell populations and even of single cells. General modeling paradigms, such as the Chemical Master Equation, exist for the description of stochastic dynamics at the single-cell level. However, due to the complexity of the interactions, current studies have often preferred to focus on specific cases of interest by *ad-hoc* modeling and analysis. In addition, theoretical and practical challenges inherent in the inference of stochastic models from biological experimental data have limited the development of general identification approaches.

Work in IBIS on the probabilistic modeling of gene expression and interaction dynamics at the level of individual cells is centered around two main challenges. On the one hand, we address identification from microscopy data and analysis of the arabinose uptake dynamics in *E. coli* upon glucose exhaustion. Starting from a reduced arabinose uptake model, Eugenio Cinquemani and Michel Page are working on methods for the estimation of unknown stochastic model parameters from statistical population snapshot data collected by fluorescence microscopy experiments. Analysis of the model focuses on the problem of model-based real-time single-cell state estimation, with feedback control applications in mind, in collaboration with Alfonso Carta (BIOCORE). Based on a stochastic model reflecting switch-like dynamics in the form of sigmoidal reaction rates, taking a Chemical Master Equation model with cell-dependent parameters as a gold standard desdcription of the system, a Chemical Langevin Equation approximation is proposed as a convenient approximation of the model for observer design purposes. On top of this model approximation, a so-called Square-Root Unscented Kalman filter (SRUKF) is designed. Based on simulations of a realistically tuned model, SRUKF is found to perform as good as much heavier particle filters based on the gold-standard model.

Results were presented at the *European Control Conference (ECC)* in 2013 [11], where we also showed that including extrinsic noise effects explicitly in the estimation process allows one to improve the knowledge of the hidden states.

On the other hand, we investigate the use mixed effects-modelling and identification techniques to characterize single-cell profiles in isogenic cell populations. Mixed-effects models are hierarchical models where parametric response profiles of individuals is subject to inter-individual parameter variability following a common population distribution. In collaboration with Gregory Batt (CONTRAINTES) and Giancarlo Ferrari-Trecate (University of Pavia, Italy), we are adapting and applying existing procedures from pharmacokinetics to the context of microfluidic data, with focus on the budding yeast response to osmolarity shocks. The first results of the work were presented at the *European Control Conference (ECC)* this year [12]. Rigorous model identification and validation steps are performed on data from real-time control experments performed in Pascal Hersen's lab at Université Paris Descartes, for both mixed-effects modelling and for the competing method of moment-based identification. Results show the tendency of mixed-effects modelling to avoid overfitting for this system, trading fitting performance for validation performance and hence predictive capabilities. The work is being further developed and the collaboration tightened by the ongoing visit of Andres Gonzalez, PhD candidate at the University of Pavia, to CONTRAINTES and IBIS. A first journal publication is in preparation, which will be followed by extensions and refinements of the method.

In parallel, work concerning the study of noise propagation in gene regulatory networks is carried out in collaboration with Irina Mihalcescu (Université Joseph Fourier). Finally, collaboration of Eugenio Cinquemani with Marianna Rapsomaniki, PhD student affiliated with Zoi Lygerou (University of Patras, Greece) and John Lygeros (ETH Zürich, Switzerland), has been devoted to the analysis of data from Fluorescence Recovery After Photobleaching (FRAP) experiments. It has given rise to a novel method for reconstructing physical diffusion and immobilization parameters at the level of single cells. The method has been applied to nuclear species of mammalian cells and results are part of a journal paper under revision.

5.5. Shared control of gene expression by global physiological effects and specific regulators

Gene expression is controlled by the joint effect of (1) the global physiological state of the cell, in particular the activity of the gene expression machinery, and (2) DNA-binding transcription factors and other specific regulators. While many studies have focused on networks of transcription factors, the analysis of the relative contributions of both transcription factors and global effects of the physiological state has received relatively little attention thus far.

In the framework of the PhD thesis of former IBIS member Sara Berthoumieux, we have developed a model-based approach to distinguish between these two effects using time-resolved measurements of promoter activities. We have demonstrated the strength of the approach by analyzing a circuit involved in the regulation of carbon metabolism in *E. coli*, consisting of two pleiotropic regulators of the cell (Crp and Fis), the gene *acs* encoding the enzyme acetyl-CoA synthetase (Acs), and the signaling metabolite cyclic AMP (cAMP) which activates Crp. *acs* is strongly expressed in the absence of glucose and is thus an excellent indicator of the transcriptional response of carbon metabolism to a growth-phase transition.

Our results show that the transcriptional response of the network is controlled by the physiological state of the cell and the signalling metabolite cAMP. The (surprising) absence of a strong regulatory effect of transcription factors suggests that they are not the main coordinators of gene expression changes during growth transitions, but rather that they complement the effect of global physiological control mechanisms. This change of perspective has important consequences for the interpretation of transcriptome data and the design of biological networks in biotechnology and synthetic biology. An article presenting the above results was published in *Molecular Systems Biology* this year [5] and selected as an Editor's choice in *Science* (http://ibis.inrialpes.fr/article1040.html).

In the above-mentioned work, the activity of the gene expression machinery was indirectly measured, by monitoring the activity of a constitutive gene, that is, a gene whose expression does not depend on any specific regulators but only on the activity of the gene expression machinery. There exists a huge literature on the molecular mechanisms coupling the activity of the gene expression machinery to changes in the nutritional quality of the environment, but a quantitative and dynamic picture of this very complicated regulatory system is still missing. Delphine Ropers and Edith Grac as well as Nils Giordano are developing models to achieve this, from bottom-up and top-down perspectives, respectively.

5.6. Control of regulatory networks in bacteria

A bacterial cell adapts its growth rate to the environment, notably to the availability of nutrients providing the molecular building blocks and the energy required for growth. Upon a change in the environment, the global physiology of the cell is adjusted in parallel with the adaptation of the growth rate. In the context of the PhD thesis of former IBIS member Jérôme Izard, we have studied the relation between the gene expression machinery, the global physiology of the cell, and the growth rate from a different perspective. Our aim was to change the mechanisms regulating the activity of the gene expression machinery in such a way so as to be able to externally control the growth rate of the cell.

More precisely, we have engineered an *E. coli* strain in which the transcription of an essential component of the global gene expression machinery, RNA polymerase, is under the tight control of an inducible promoter. By adjusting the inducer concentration in the medium we can adjust the RNA polymerase concentration and thereby reversibly tune the growth rate of the bacterium between zero and the maximal growth rate. The growth arrest is completely reversed when RNA polymerase is provided again. The analysis of the transcriptome at growth rates restricted by the concentration of RNA polymerase confirms that the concentration of RNA polymerase is the major determinant of changes in gene expression patterns. Our modified *E. coli* strain provides a novel way of setting growth rate in a tunable, reversible, modular, and medium-independent way. The strain, described in a paper submitted for publication, opens new perspectives for studying the mechanisms of growth control as well as for developing biotechnological applications, the subject of the post-doctoral fellowship of Cindy Gomez Balderas-Barillot. We have submitted a patent proposing such applications, which underlies the technology transfer activities undertaken in the recently-started Reset project (Section 7.2).

IMAGINE Project-Team

5. New Results

5.1. Introduction

We are developing user-centred, knowledge-based models in three main domains: shape, motion and narrative design, leading us to three research axes. The fourth one is the combination of these models with intuitive interaction tools, in order to set up interactive creative environments dedicated to specific categories of content. The following sections describe our activities in 2013 for each axis.

5.2. High level model for shapes

- **Scientist in charge**: Stefanie Hahmann
- **Other permanent researchers**: Marie-Paule Cani, Jean-Claude Léon.

5.2.1. Implicit modeling

Participants: Antoine Bégault, Adrien Bernhardt, Marie-Paule Cani, Mohamed-Galal Koraa, Cédric Zanni.

Implicit surfaces are an appealing representation for free-form, volumetric shapes. In addition to being able to represent shapes of arbitrary topological genius, they have the ability to be constructed by successively blending different components, which eases interactive modeling.

In collaboration with Loic Barthe in Toulouse, we contributed to a new binary blending operator, called Gradient Blending [7], which enables us to blend implicit shapes not only in function of their field values but also of their gradients. This solves a number long standing problems in implicit modeling: we can generate bulge-free blending, ensure that the topological genius of the blended shape remains the one of the union of the input one, and avoid the blur of small details.

Figure 4. Example of surface generated using our SCALIS approach.

Within Cédric Zanni's PhD [2] we introduced closed-form solutions for convolution surfaces along helical skeletons and extended Gabor-noise texturing to enable the creation of repetitive geometric details along implicit surfaces. We also developed a novel extension to convolution surfaces, so-called SCALe-invariant Integral Surfaces (SCALIS) [15], see Figure 4 . Thanks to their scale invariant blending properties, these surfaces have three main advantages: the radius of the surface around a skeleton can be explicitly controlled, shapes are self-similar regardless of the scale of the model, and thin components are not smoothed-out when blended into larger ones. This is done while preserving the main benefits of integral surfaces, namely n-ary blending with a simple plus, and shape preservation whatever the way the skeletons is spit into smaller primitives. We are currently extending this work to enable the topology of the implicit surface to always reflect the one of the skeleton

5.2.2. Analysis of CAD models

Participants: Francois Faure, Stefanie Hahmann, Jean-Claude Léon, Olivier Palombi, Flavien Boussuge, Ahmad Shahwan.

CAD models, as part of assemblies defining manufactured products, are often shaped in accordance with their physical counterpart. However, one can observe that the shape of some components, as modeled in CAD, may differ from that of their physical instance. In addition, assemblies representing products are most often reduced to a collection of CAD models representing each component and the designation of each component is neither a reliable information nor a faithful connection with one or more functions of a component. As a result, geometric interfaces between components are unknown and they cannot be reduced to contacts. Interferences may exist that are also relevant for several applications. Determining precisely, the geometric interfaces between components is a first requirement to enrich geometric models with functional information because a subset of functions derive from interfaces between components.

As an example, this is particularly useful for structural mechanics to be able to generate rapidly a Finite Element model of assemblies and it is especially critical when assemblies get very complex. [9] addresses the problem to generate automatically a class of geometric interfaces for very complex assemblies (see fig. 5). GPU-based algorithms have proved suitable to obtain reliable results on CAD models.

Using these geometric interfaces as well as the newly introduced concept of conventional interfaces, [6], [4] and [11], [12], [19] have proposed an approach using qualitative reasoning, ontology reasoning to connect CAD components, their geometric interfaces, to functions and functional designations of components: an intrinsic identifier of a component in an assembly that connects it to its function. As a result, it is shown how geometric models of components need to be restructured, which extends the concept of annotation presently reduced to a elementary link between geometric models and symbolic information.

At the level of assembly components, shape analysis [30] is particularly useful to generate dimensionally reduced models needed for structural mechanics. [3] shows that analyzing a B-Rep CAD model to derive a construction graph, i.e. a set of construction trees, can be a robust basis to generate dimensionally reduced models [18], [32].

5.2.3. Knowledge-based shape transfert

Participants: Marie-Paule Cani, Ali Dicko, Francois Faure, Olivier Palombi.

Characters with precise internal anatomy are important in film and visual effects, as well as in medical applications. We have proposed the first semi-automatic method for creating anatomical structures, such as bones, muscles, viscera and fat tissues [5], as illustrated in 6 . This is done by transferring a reference anatomical model from an input template to an arbitrary target character, only defined by its boundary representation (skin). The fat distribution of the target character needs to be specified. We can either infer this information from MRI data, or allow the users to express their creative intent through a new editing tool. The rest of our method runs automatically: it first transfers the bones to the target character, while maintaining their structure as much as possible. The bone layer, along with the target skin eroded using the fat thickness information, are then used to define a volume where we map the internal anatomy of the source model using harmonic (Laplacian) deformation. This way, we are able to quickly generate anatomical models for a large range of target characters, while maintaining anatomical constraints.

../../../../projets/imagine/IMG/Hahmann-Raweb2013.png

Figure 5. Computation of assembly interfaces in NURBS representation.

../../../../projets/imagine/IMG/anatomyTransfer.png

Figure 6. A reference anatomy (left) is automatically transferred to arbitrary humanoid characters. This is achieved by combining interpolated skin correspondences with anatomical rules.

5.3. Models for motion and animation

- **Scientist in charge**: François Faure
- **Other permanent researchers**: Marie-Paule Cani, Damien Rohmer, Rémi Ronfard.

5.3.1. *Physical models*

Participants: Marie-Paule Cani, François Faure, Pierre-Luc Manteaux.

Frame-based deformable solids Our frame-based deformable model was published as a book chapter [31]. It combines the realism of physically based continuum mechanics models and the usability of frame-based skinning methods, allowing the interactive simulation of objects with heterogeneous material properties and complex geometries. The degrees of freedom are coordinate frames. In contrast with traditional skinning, frame positions are not scripted but move in reaction to internal body forces. The deformation gradient and its derivatives are computed at each sample point of a deformed object and used in the equations of Lagrangian mechanics to achieve physical realism. We introduce novel material-aware shape functions in place of the traditional radial basis functions used in meshless frameworks, allowing coarse deformation functions to efficiently resolve non-uniform stiffnesses. Complex models can thus be simulated at high frame rates using a small number of control nodes.

Figure 7. A dam break simulation with 5000 particles simulated with WCSPH (on the left) and with our adaptive method (on the right). On the right image, blue corresponds to full-dynamics particles, green to transition particles and red to restrained particles.

Adaptive particle simulation In collaboration with the NANO-D Inria Team, we have explored the use of Adaptively Restrained (AR) particles for graphics simulations [25]. Contrary to previous methods, Adaptively Restrained Particle Simulations (ARPS) do not adapt time or space sampling, but rather switch the positional degrees of freedom of particles on and off, while letting their momenta evolve. Therefore, inter-particles forces do not have to be updated at each time step, in contrast with traditional methods that spend a lot of time there. We first adapted ARPS to particle-based fluid simulations, as illustrated in 7 and proposed an efficient incremental algorithm to update forces and scalar fields. We then introduced a new implicit integration scheme enabling to use ARPS for cloth simulation as well. Our experiments showed that this new, simple strategy for adaptive simulations can provide significant speedups more easily than traditional adaptive models.

5.3.2. *Skinning virtual characters*

Participants: Marie-Paule Cani, Damien Rohmer.

Skinning is a widely used technique to deform articulated virtual characters. It can be computed fastly and therefore can deliver real-time feedback at the opposite of physically based simulation. Still standard skinning approaches cannot handle well large deformations and may require manual corrections.

In collaboration with Loic Barthe and Rodolphe Vaillant from University of Toulouse, and collaborators from Victoria University, Inria Bordeaux and University of Bath, we develop a new automatic correction for skinning deformation that has been published in SIGGRAPH [14]. Based on the volumetric implicit representation paradigm, it adjust the mesh vertices and improves the visual appearance of the deformed surface. Moreover, it seamlessly handle skin contact ensuring that no self collision can occurs as seen in fig. 8 . Finally, the method can mimic muscular bulges controled by the implicit blending operators described in the work [7].

`../../../../projets/imagine/IMG/skinning.png`

Figure 8. Example of large skinning deformation obtained in [14].

5.3.3. Animating crowds

Participants: Marie-Paule Cani, Quentin Galvane, Kevin Jordao, Kim Lim.

Crowd animation is an interesting case, since it can be either computed by developing artificial intelligence methods, by using physically-based simulation of some extended particle systems, or by applying a kinematic texturing methodology, made possible by the repetitive nature of crowd animations. We launched this new topic in the group in 2013, enabling us to explore the two last crowd animation methods:

Firstly, in collaboration with the University SAINTS, Malaysia, we extended particle-based crowd simulation to the case when 4 different populations, with different goals and behaviors, are interacting within the same environment [24]. This as illustrated by a cultural heritage application, with the reconstruction of past life in a harbor in Malaysia in the 19th century: see Figure 9 .

../../../../projets/imagine/IMG/crowd.png

Figure 9. Crowd simulation with 4 different populations from [24].

Secondly, within the ANR project CHROME with Inria Rennes, we adopted the crowd-patches technique, i.e. the idea of combining patches carrying pre-computed crowd trajectories, for quickly populating very large environments [23]. We are currently developing novel methods for enabling the interactive space-time editing of these animations (a paper will be published at the next Eurographics conference).

5.4. Knowledge-based models for narrative design

- **Scientist in charge**: Rémi Ronfard
- **Other permanent researchers**: Marie-Paule Cani, François Faure, Jean-Claude Léon, Olivier Palombi

5.4.1. Cinematographic virtual camera control

Participants: Marie-Paule Cani, Quentin Galvane, Vineet Gandhi, Chen Kim Lim, Rémi Ronfard.

Steering Behaviors for Autonomous Cameras [21] : We proposed a new method for automatically filming crowd simulations with autonomous cameras, using specialized camera steering behaviors and forces. Experimental results show that the method provides a good coverage of events in moderately complex crowds simulations, with consistently correct image composition and event visibility.

The prose storyboard language [26] : We presented a formal language for describing movies shot by shot, where each shot is described with a unique sentence. The language uses a simple syntax and limited vocabulary borrowed from working practices in traditional movie-making, and is intended to be readable both by machines and humans. The language is designed to serve as a high-level user interface for intelligent cinematography and editing systems.

5.4.2. Virtual actors

Participants: Adela Barbulescu, Rémi Ronfard.

Audio-Visual Speaker Conversion using Prosody Features [17] : We presented a new approach towards speaker identity conversion using speech signals and 3D facial expressions. Audio prosodic features are extracted from time alignment information for a better conversion of speaking styles. A subjective evaluation was performed to illustrate that the converted sequences are perceived as belonging to the target speakers. We are working to extend that approach to visual prosody features and to apply it to the situation where a director controls the expressions of a virtual actor, while maintaining its personality traits.

5.4.3. Narrative analysis of video

Participants: Vineet Gandhi, Rémi Ronfard.

Naming and detecting actors in movies [22]: We proposed a generative model for localizing and naming actors in long video sequences. More specifically, the actor's head and shoulders are each represented as a constellation of optional color regions. Detection can proceed despite changes in view-point and partial occlusions. This work is being extended to the case of theatre actors during performances and rehearsals. It also opens the way to future work in automatic analysis of cinematographic and editing styles in real movie scenes. This was also presented as a poster at the International Conference on Computational Photography (ICCP).

Recording theatre rehearsals [29] : We presented a contribution to the International Federation for Theatre Research describing our ongoing collaboration with the Theatre des Celestins in Lyon, emphasising that high quality vídeo recordings make it possible to study the genetic evolution of a theatre performance, and make it an object of scientific study as well as an object of aesthestic appreciation.

5.5. Creating and interacting with virtual prototypes

- **Scientist in charge**: Jean-Claude Léon
- **Other permanent researchers**: Marie-Paule Cani, Olivier Palombi, Damien Rohmer, Rémi Ronfard.

5.5.1. Sketch-based modeling

Participants: Marie-Paule Cani, Martin Guay, Rémi Ronfard.

The Line of Action [8] : The line of action is a conceptual tool often used by cartoonists and illustrators to help make their figures more consistent and more dramatic. In this paper, we proposed a mathematical definition of the line of action (LOA), which allows us to automatically align a 3D virtual character to a user-specified LOA by solving an optimization problem. This woek is now being extended to the more challenging case of creating complete animations from storyboard-like hand-drawn sketches (see fig. 10).

../../../../projets/imagine/IMG/lineofactionteaser.png

Figure 10. The line of action defines the pose of the virtual character.

5.5.2. Sculpting methods

Participants: Marie-Paule Cani, Stefanie Hahmann, Damien Rohmer, Lucian Stanculescu.

Sculpting methods is a very powerful approach to design virtual models from an existing model. In the work of Lucian Stanculescu [13] we extend the standard sculpting paradigm of surfaces, to multi-dimensional nested structures. In this method, lower dimensional structures such as points and curves can be defined on the surface to defined a nested structure. Each part can follow a specific deformation behaviors. We therefore categorize the geometrical and topological behavior of the structure (such as rigidity or mutability) to develop a wider range of possible deformation. This method facilitate the persistence of sharp features that automatically split or merge with variable rigidity, even when the shape chages genus. This approach enable to deform a surface exhibiting typical behavior of both CAD model with sharp edges, and CG model with smooth surfaces as seen in fig. 11 .

../../../../projets/imagine/IMG/lucian.png

Figure 11. Deformed surface with nested structure from [13].

We also extends the sculpting approach to handle detailed surfaces. During sculpting deformation such as local stretching, the surface details should not extend as the global shape, but rather duplicates to ensure that the surface keeps his detailed appearance. We studied this question under two different approaches.

The first one, in collaboration with Max-Planck Institute focussed on the deformation of 1D-like parametric structure such as castle walls of centripede characters. The method enable to freely extend, compress, split and merges parts of the structures. The deformed structure is generated by an assembly of basic parts whose behaviors are encoded using a discrete shape grammar. During deformation, the system finds the most suitable collection of parts to assemble and ensure that the global shapes is coherent with the input rules. This work as been published in Eurographics 2014 [10].

The second approach consists in extending sculpting to continuous freeform deformation of a 2D surface with details. During the deformation gesture of stretching or compression, details on the surfaces should seamlessly appear or desappear continuously. In this work, we studied the simpler case of a planar surface with high field details and presented our result in GTMG [27]. We now work on the more general extention to this work as a collaboration with Max-Plack Institute and University of College London.

5.5.3. Interactive control of procedural models: terrains and waterfalls

Participants: Adrien Bernhardt, Marie-Paule Cani, Arnaud Emilien, Ulysse Vimont.

Procedural models, used for easily modeling large, natural environments, pose a specific challenge in terms of user control: how can these automatic methods, useful for quickly generating a huge number of self-similar details, be adapted to allow the coarse to fine level of control needed by the users?

This topic was first explored within Adrien Bernhard's PhD thesis [1], where we introduced a real-time terrain modeling tool using a fast GPU-based terrain solver with a lightweight CPU-based data structure. This tool was recently extended in collaboration with Cambridge University, to enable first-person sketch-based editing of terrains models.

Secondly, we have been working on interactive procedural modeling of plausible waterfalls, in collaboration with Montreal University. Offering interactive user control for this application is particularly challenging, since the shape taken by a fall heavily depends on the underlying terrain. Our solution, based on vectorial user-control, on a flow solver, and on procedural adaptation of the underlying terrain, enable users to quickly create plausible flowing water, while controlling which fall segments are in contact with the terrain (vs. in free fall), the topology of the network, and how much the flow should adapt to the current terrain, vs. the terrain to the user-designed trajectories (fig. 12). A paper is under review and a presentation has been made at the AFIG conference [16].

../../../../projets/imagine/IMG/trou_2.jpg ../../projets/imagine/IMG/trou.jpg

Figure 12. Waterfalls modeling using our approach developped in [16]

5.5.4. Interaction methods

Participants: Rémi Brouet, Marie-Paule Cani.

We are currently exploring the use of multi-touch tables for the interactive design and editing of 3D scenes, in collaboration with the human-computer interaction group of LIG laboratory. The main challenge here is to find out how to use a 2D interaction media for editing 3D content, hence how to intuitively control the third dimension (depth, non-planar rotations, 3D deformations, etc).

Our first work consisted in an user study where we analyzed all possible hand interactions on table-tops and explored the ways users would intuitively try to manipulate 3D environments, either for changing the camera position or for moving objects around [20]. We extracted a general interaction pattern from this study. Our implementation enables both seamless navigation and docking in 3D scenes, without the need for any menu or button to change mode. We are currently extending this work to object editing scenarios, where shapes are to be bent of twisted in 3D using 2D interaction.

IMARA Project-Team

6. New Results

6.1. ABV

Participants: Hao Li, Paulo Lopes Resende, Evangeline Pollard, Joshué Pérez Rastelli, Fawzi Nashashibi.

The ABV project builds on the HAVEit philosophy (a previous IMARA project for high speed automation) by 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. Real experiments on a prototype vehicle have been conducted. The experiment objective was to realize several use-cases: lane following, changing of lane, overtaking, ACC and emergency braking. All these maneuvers have been successfully conducted several times on the Satory tracks (cf. [46], [41] for more details) during the final event of the project which took place late March 2013.

6.2. Urban Autonomous Driving

Participants: Evangeline Pollard, Guillaume Tréhard, Fawzi Nashashibi.

Beyond low speed automation, IMARA is tackling a very important issue for autonomous driving on open roads, which is: dealing with intersections. In collaboration with Valeo, Imara wants to provide innovative way to safely cross any kind of intersections for an autonomous vehicle in a urban context and without communication. The goal is to deal with intersection with different shapes, (roundabout, T junctions , X junctions, *etc.*), with different rules, specific (traffic lights, main road...) or not ("priority to the right" in France), with different traffic (busy or empty).

6.3. Vehicle to pedestrian communications

Participants: Pierre Merdrignac, Oyunchimeg Shagdar, Evangeline Pollard, Fawzi Nashashibi.

Vehicle and pedestrian collisions often result in fatality and serious injury to the vulnerable road users. While vehicle to vehicle (V2V) communications have taken much attention in the academic and industrial sectors, very limited effort has been made for vehicle to pedestrian communications. Unlike the V2V cases, where antennas are often installed on the vehicle rooftop, pedestrian's handheld device can be carried in such a way e.g, in a bag or in a pocket, which results in poor and unpredictable communications quality. In this work, we seek to an answer to the questions of whether the Wi-Fi-based V2P communications meet the requirements of the pedestrian safety application. This year, we studied the performances of the V2P communications especially receive signal strength, packet inter-arrival time, and message delivery ratio. Moreover, in order to demonstrate the feasibility of pedestrian safety supported by the V2P communications, we developed a software tool, V2ProVu, which has the functionalities of Wi-Fi based V2P communications, collision risk calculations, and hazard alarming.

6.4. Visible light communications for platooning control

Participants: Mohammad Abualhoul, Oyunchimeg Shagdar, Mohamed Marouf, Fawzi Nashashibi.

While V2V communications is requisite for platooning stability, the existing radio communications technologies suffer from poor performance in highly dense road scenarios, which are exactly to be created for platooning. Targeting this issue, we study the applicability of visible light communications (VLC) for information exchange between the platoon members [20], [35]. Because the existing studies on VLC mainly focus on indoor applications or for communications from traffic light to vehicle, the performances of VLC for V2V is not clear. In this work, we develop a complete VLC channel and noise model by taking account of the key parameters including background noise and incidence angle. Our studies show that it is feasible to achieve up to 7 meters line of sight communication range even in the presence of optical noise at significant levels and with up to 60 degree of road curvature.

6.5. ITS-G5 for road safety and efficiency applications

Participants: Oyunchimeg Shagdar, Younes Bouchaala, Mohammad Abualhoul, Manabu Tsukada, Thierry Ernst.

To support V2V and V2I communications for road safety and efficiency applications, ETSI standardized ITS-G5 technology. One of key objectives of the SCORE@F project is to study the performance of ITS-G5 in real-world scenarios and demonstrate its applicability to road safety and efficiency applications. Under the scope of the SCORE@F project, we studied the performances of ITS-G5 for both the V2V and V2I communications based on field tests and theoretical studies with emphasis on the effects of channel in combination with MAC and some parameters of car traffic [32]. An important insight achieved from the study is that in addition to the distance dependent pathloss, the signal fading and road traffic characteristics provide significant impacts on the reliability of ITS-G5.

We also study the performance of the ITS-G5 medium access control protocol for realistic autonomous driving applications especially to seek answers to the questions of whether the IEEE 802.11p can support merging control and how the communications performance is translated into that of CACC (Cooperative Adaptive Cruise Control) [33]. The study discloses several useful insights including packet inter-arrival time and throughput but not packet delivery ratio, gives good indications of the CACC performance; the V2I communications structure is preferred over the V2V structure for CACC.

Finally, we demonstrate the low latency video streaming over ITS-G5 to support platoon and reverse parking maneuvers [21].

6.6. Cooperative driving

Participants: Joshué Pérez Rastelli, Fawzi Nashashibi.

In the scope of the French project *"Co-Drive"* one task assigned to Inria was the development of a smart controller capable of driving the vehicle, allowing it to perform optimal traversal of traffic lights in order to reduce vehicle accelerations and thus the gas emissions. This controller needs remote information regarding the traffic lights' status, the distance to it and the time needed to reach it.

Three input variables, which are the traffic light times, red light, green light and the distance to interception (DTI), were defined in fuzzy logic tool [37].

Two variables are used for the traffic light (Red and Green), where each of them has defined two completely symmetrical membership functions covering all the possible inputs. In this application the time cycle of the lights are 30 seconds for green and 20 seconds for red. The values of input membership functions were defined considering these times.

The DTI membership function (see Figure 1) gives more weight to the distance when the vehicle is closer at the intersection. In this situation, the vehicle can be inside the *short* or the *middle* label, because in these cases the response has to be faster than in the case where the vehicle is in the *long* label. The cross rule base, based on driver knowledge when the vehicle is arriving to an intersection, are defined using natural language.

Some Simulations were performed to validate the controller. However, the final implementation will be presented in 2014 during the final event of Co-Drive Project.

../../../../projets/imara/IMG/Figura5.jpg

Figure 1. Codrive: input variables for the speed reference fuzzy controller

6.7. Intelligent Planning algorithm using Bezier curves

Participants: Joshué Pérez Rastelli, Fawzi Nashashibi.

The Bezier curve is the heart of the Local Planning, which allows a fast trajectory computation in order to send the trajectory in real-time to the controller stage. This method has been recently used in robot mobile solutions due to its versatility and simplicity for intersections.

We have proposed a novel method for the generation of control points for two distinct road configurations: roundabouts and a standard intersections. If an intersection is being dealt with, the control points will be generated based on the reference path given by the Global Planner.

The experiments we made presented several urban intersections. Figure 2 shows the whole generated path with four intersections and a roundabout, using the global map. A comparison with different methods is drawn. The first one (thin line) is based on the static method used in [22], which sets the control points by hand. In this case we can see how sometimes the path passes over the sidewalk. The second experiment (dotted line) is using the same previous method, but modifying the distance used to position the control points, in order to obtain a path into the road. The third method (thick line) is the Intelligent Planning algorithm. As we can see in the figure, the automatic algorithm sets the control points of Bezier (based on the convex hull property) achieving a smooth path, without going over sidewalks or obstacles.

6.8. Ontologies

Participants: Evangeline Pollard, Philippe Morignot, Fawzi Nashashibi.

Full autonomy of ground vehicles is a major goal of the ITS (Intelligent Transportation Systems) community. However, reaching such highest autonomy level in all situations (weather, traffic, . . .) is seen as impossible in practice, despite recent results regarding driverless cars (e.g., Google Cars). In addition, an automated vehicle should also self-assess its own perception abilities, and not only perceive its environment. In this new research axis, we propose an intermediate approach towards full automation, by defining a spectrum of automation

Figure 2. Generated path using different methods

levels, from fully manual (the car is driven by a driver) to fully automated (the car is driven by a computer), based on an ontological model for representing knowledge. We also propose a second ontology for situation assessment (what does the automated car perceive?), including the sensors/actuators state, environmental conditions and driver's state. Finally, we also define inference rules to link the situation assessment ontology to the automation level one [24].

6.9. Communications and Management Control for Cooperative Vehicular Systems

Participants: Ines Ben Jemaa, Oyunchimeg Shagdar, Arnaud de La Fortelle.

One of the attractive applications of electric autonomous vehicles is electric automated Car Sharing service, where on-demand passenger transportation is provided by a set of automated vehicles and a control center, which is installed in the Internet. Data transmission from the control center to the set of vehicles requires an efficient multicast data delivery, i.e. multi-cast routing. The conventional multicast routing in the Internet is based on protocols such as Protocol Independent Multicast (PIM), which relies on a tree structure to deliver packets from the source to the destinations. Thanks to the fixed topology of the Internet, it is possible to build a large and stable multicast trees. However, due to the highly mobile nature of vehicular networks, it is not clear how stable and large can be such trees in vehicular environments. This year, we studied the stability of multicast trees for data flows from the Internet to a set of vehicles [38], [36]. Our study shows that the stability of multicast tree largely depends on the relative velocity (inter-vehicle) and the road density but not directly on the road shape or moving direction. Based on our study we are developing a mobility aware multicast routing protocol, which constructs its tree based on the vehicles' mobility dynamics and the road condition.

6.10. New urban transportation platforms: Inria's Cybus

Participants: François Charlot, Joshué Pérez Rastelli, Fawzi Nashashibi, Paulo Lopes Resende, Michel Parent, Armand Yvet.

Figure 3. The Cybus operated at La Rochelle City during 3 months as a free transport service.

Cybus is the best known 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, and Bezier curve in path planning generation. 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.

6.11. Real-time visual perception: detection and localization of static and moving objects from a moving stereo rig

Participants: Benjamin Lefaudeux, Fawzi Nashashibi.

Perception of the surrounding environment is one of the many tasks an automated vehicle has to achieve in complex and ever-changing surroundings. This, typically includes several distinct sub-tasks, such as map-building, localization, static obstacles and moving objects detection and identification. Some of these tasks are nowadays very well known, such as the map-building process which has been extensively investigated in the last decade ; whereas the perception, localization and classification of moving objects from an equally moving vehicle are in many aspects a work in progress. The objective of the PhD thesis of Benjamin Lefaudeux was to propose a vision-based approach built on the extensive tracking of numerous visual features over time, from a stereo-vision pair.

Through on-the-fly environment 3D reconstruction, based on visual clues, we proposed an integrated method to detect and localize static and moving obstacles, whose position, orientation and speed vector is estimated. Our implementation runs in real-time depending on the number of processed points, and should in the future be enclosed in a more complete, probabilistic pipeline. The complete achievements are described in the thesis of Benjamin Lefaudeux ([8] defended on September 30th) with very interesting and competitive results obtained with international benchmarks (cf. Figure 4) and on the real vehicles of IMARA.

6.12. Belief propagation inference for traffic prediction

Participants: Cyril Furtlehner, Jean-Marc Lasgouttes, Victorin Martin.

../../../../projets/imara/IMG/KITTI_CamView.jpg

../../../../projets/imara/IMG/KITTI_Estimated

Figure 4. Left: A single camera view from the KITTI sequence. Right: A bird view of the scene as modeled by the system: point cloud and estimated trajectory.

This work [55] deals with real-time prediction of traffic conditions in a setting where the only available information is floating car data (FCD) sent by probe vehicles. The main focus is on finding a good way to encode some coarse information (typically whether traffic on a segment is fluid or congested), and to decode it in the form of real-time traffic reconstruction and prediction. Our approach relies in particular on the belief propagation algorithm.

These studies have been done in particular in the framework of the projects Travesti and Pumas.

This year's highlights are

- Victorin Martin has defended his PhD thesis entitled "Modélisation probabiliste et inférence par l'algorithme Belief Propagation" [9] at Mines-ParisTech on May 23.
- The work about the theoretical aspects of encoding real valued variables into a binary Ising model has been published as a research report [44] and submitted for publication.

6.13. Sparse covariance inverse estimate for Gaussian Markov Random Field

Participants: Cyril Furtlehner, Jean-Marc Lasgouttes, Victorin Martin.

We investigate the problem of Gaussian Markov random field (GMRF) selection under the constraint that the model is suitable for Gaussian belief propagation (GaBP) inference. We develop a method based on iterative proportional scaling (IPS) to incrementally select optimal GMRF factors, while maintaining GaBP compatibility. Besides the intrinsic sparsity-inducing capability, the proposed method is indeed sufficiently flexible to incorporate various spectral constraints like e.g. walk summability (WS) to insure the compatibility of the solutions with Gaussian Belief Propagation inference. Experimental tests on various datasets with refined L_0 or L_1 regularized sparse inverse estimate indicate that this approach is competitive and provides us with useful alternatives to traditional sparsity-inducing penalizations norms, giving more freedom in the graph structure selection process with no additional computational cost.

6.14. 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 are developing a discrete events simulator for that purpose, which model relies on an adapted events/decision graph. 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 a partial rewrite of the simulator in order to make it more generic and handle the new setting of the CATS project with automated vehicles.

6.15. Herding behavior in a social game

Participants: Guy Fayolle, Jean-Marc Lasgouttes.

The system *Ma Micro Planète* belongs to the so-called *Massively Multi-Player online Role Playing game* (MMORPG), its main goal being to incite users to have a sustainable mobility. Two objectives have been pursued.

- Construct an experimental platform to collect data in order to prompt actors of the mobility to share information (open data system).
- See how various mechanisms of a game having an additive effect could modify the transportation requests.

At the heart of the game are community-driven *points of interest* (POIs), or *sites*, which have a score that depends on the players activity. The aim of this work is to understand the dynamics of the underlying stochastic process. We analyze in detail its stationary regime in the thermodynamic limit, when the number of players tends to infinity. In particular, for some classes of input sequences and selection policies, we provide necessary and sufficient conditions for the existence of a complete meanfield-like measure, showing off an interesting *condensation* phenomenon.

The work has been published this year in *Queueing Systems* [11].

6.16. Analytic properties 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 these last three years in two main directions.

6.16.1. The group and zero drift case

In several recent studies on random walks with small jumps in the quarter plane, it has been noticed that the so-called *group of the walk* governs the behavior of a number of quantities, in particular through its *order*. When the *drift* of the random walk is equal to 0, we have provided an effective criterion (see RA 2012) giving the order of this group. More generally, we showed that in all cases where the *genus* of the algebraic curve defined by the so-called *kernel* is 0, the group is infinite, except precisely for the zero drift case, where finiteness is quite possible.

This year, we investigated new proofs of this results, which could lead to an explicit tractable criterion for the finiteness of the group, which a priori, as shown in [2] involves a ratio of elliptic integrals.

6.16.2. Counting and asymptotics

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. In the case of small jumps (size at most one), 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

$$K(x,y,z)F(x,y,z) = c(x)F(x,0,z) + \widetilde{c}(y)F(0,y,z) + c_0(x,y),$$

where x, y, z are complex variables, $K(x,y,z)$ is a polynomial of degree 2 (both in x and y), and linear in the time variable z which plays somehow the role of a parameter. The question of the type of the associated counting generating functions, rational, algebraic, or holonomic (i.e. 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 open.

The nature of the singularities of the function F plays a key role for this classification. Starting from our study [54], we proved in various cases that the first singularities of F(1,0,z) are either polar or correspond to a value z_g for which the genus of the algebraic curve $K(x,y,z) = 0$ passes from 1 to 0 (i.e. a torus becomes a sphere).

6.16.3. Harmonic functions and more general jumps

The determination of Martin boundaries in the case of random walks is a longstanding problem, solved only in special situations. For homogeneous random walks in the quarter plane, stopped on the boundary (the axes), with upward jumps of size 1, and arbitrary downward jumps of size d, it turns out that the computation of harmonic functions is here plainly equivalent to find a positive function H satisfying a functional equation of the form

$$L(x,y)H(x,y) = L(x,0)H(x,0) + L(0,y)H(0,y) - L(0,0)H(0,0).$$

Here the chief difficulty to make the reduction to a boundary value problem is to analyze the algebraic curve $L(x,y) = 0$, which might be of arbitrary genus. Some examples lead us to conjecture the existence of a *single real cut* inside the unit disk, which should allow to get integral form solution.

6.16.4. Correction of papers

Guy Fayolle found important errors in several articles dealing with models involving random walks in the quarter plane. This is the object of the letter to the editors [10]. The Concerned authors are currently preparing corrected versions.

IN-SITU Project-Team

6. New Results

6.1. Interaction Techniques

Participants: Caroline Appert, Michel Beaudouin-Lafon, David Bonnet, Anastasia Bezerianos, Olivier Chapuis [correspondant], Cédric Fleury, Stéphane Huot, Can Liu, Wendy Mackay, Halla Olafsdottir, Cyprien Pindat, Theophanis Tsandilas.

We explore interaction techniques in a variety of contexts, including individual interaction techniques on different display surfaces that range from mobile devices to very large wall-sized displays through standard desktop and tabletops. This year, we investigated how people can use different body parts and limbs to convey information to interactive systems. *BodyScape* provides a framework for analysing and designing interaction techniques that involve the entire human body. Both *WristPointing*, which overcomes the limited range of motion of the wrist, and *HeadPad*, which takes the user's head orientation into account, are whole body techniques that facilitate target acquisition. *Arpege* can interpret a wide range of chord gestures, designed according to the range of motion and limitations of the human hand, and includes a dynamic guide with integrated feedforward/feedback to enhance learning by novices, without slowing down experts. On mobile devices, we designed novel interaction techniques that increase the expressivity of gestures by a single finger, including *ThumbRock*, based on movement dynamics, *SidePress*, which senses pressure on the device, and *Powerup*, which detects proximity. We also continued to develop advanced interactive visualization techniques, including *Gimlenses*, which supports focus+context representations for navigating within 3D scenes.

BodyScape – The entire human body plays a central role in interaction. The BodyScape design space [34] (honorable mention at CHI 2013) explores the relationship between users and their environment, specifically how different body parts enhance or restrict movement for specific interactions. BodyScape can be used to analyze existing techniques or suggest new ones. In particular, we used it to design and compare two free-hand techniques, on-body touch and mid-air pointing, first separately, then combined. We found that touching the torso is faster than touching the lower legs, since it affects the user's balance; and touching targets on the dominant arm is slower than targets on the torso because the user must compensate for the applied force.

HeadPad – Rich interaction with high-resolution wall displays is not limited to remotely pointing at targets. Other relevant types of interaction include virtual navigation, text entry, and direct manipulation of control widgets. However, most techniques for remotely acquiring targets with high precision have studied remote pointing in isolation, focusing on pointing efficiency and ignoring the need to support these other types of interaction. We investigated high-precision pointing techniques capable of acquiring targets as small as 4 millimeters on a 5.5 meters wide display while leaving up to 93 of a typical tablet device's screen space available for task-specific widgets [27]. We compared these techniques to state-of-the-art distant pointing techniques and have shown that two of our techniques, a purely relative one and one that uses head orientation, perform as well or better than the best pointing-only input techniques while using a fraction of the interaction resources.

WristPointing – Wrist movements are physically constrained and take place within a small range around the hand's rest position. We explored pointing techniques that deal with the physical constraints of the wrist and extend the range of its input without making use of explicit mode-switching mechanisms [33]. Taking into account elastic properties of the human joints, we investigated designs based on rate control. In addition to pure rate control, we examine a hybrid technique that combines position and rate-control and a technique that applies non-uniform position-control mappings. Our experimental results suggest that rate control is particularly effective under low-precision input and long target distances. Hybrid and non-uniform position-control mappings, on the other hand, result in higher precision and become more effective as input precision increases.

Arpege – While multi-touch input has become a standard for interacting with devices equipped with a touchscreen with simple techniques like pinch-to-zoom, the number of gestures systems are able to interpret remains rather small. Arpège [23] is a progressive multitouch input technique for learning chords, as well as a robust recognizer and guidelines for building large chord vocabularies. We conducted two experiments to evaluate our approach. Experiment one validated our design guidelines and suggests implications for designing vocabularies, i.e. users prefer relaxed to tense chords, chords with fewer fingers and chords with fewer tense fingers. Experiment two demonstrated that users can learn and remember a large chord vocabulary with both Arpège and cheat sheets, and Arpège encourages the creation of effective mnemonics.

ThumbRock – Compared with mouse-based interaction on a desktop interface, touch-based interaction on a mobile device is quite limited: most applications only support tapping and dragging to perform simple gestures. Finger rolling provides an alternative to tapping but uses a recognition process that relies on either per-user calibration, explicit delimiters or extra hardware, making it difficult to integrate into current touch-based mobile devices. We introduce ThumbRock [19], a ready-to-use micro gesture that consists in rolling the thumb back and forth on the touchscreen. Our algorithm recognizes ThumbRocks with more than 96% accuracy without calibration nor explicit delimiter by analyzing the data provided by the touch screen with a low computational cost. The full trace of the gesture is analyzed incrementally to ensure compatibility with other events and to support real-time feedback. This also makes it possible to create a continuous control space as we illustrate with our MicroSlider, a 1D slider manipulated with thumb rolling gestures.

SidePress – Virtual navigation on a mobile touchscreen is usually performed using finger gestures: drag and flick to scroll or pan, pinch to zoom. While easy to learn and perform, these gestures cause significant occlusion of the display. They also require users to explicitly switch between navigation mode and edit mode to either change the viewport's position in the document, or manipulate the actual content displayed in that viewport, respectively. SidePress [31] augments mobile devices with two continuous pressure sensors co-located on one of their sides (Figure 9 -(Left)). It provides users with generic bidirectional navigation capabilities at different levels of granularity, all seamlessly integrated to act as an alternative to traditional navigation techniques, including scrollbars, drag-and-flick, or pinch-to-zoom. We built a functional hardware prototype and developed an interaction vocabulary for different applications. We conducted two laboratory studies. The first one showed that users can precisely and efficiently control SidePress; the second, that SidePress can be more efficient than drag-and-flick touch gestures when scrolling large documents.

Powerup – Current technology like Arduino (http://arduino.cc/) opens a large space for designing new electronic device. We built the Power-up Button [30] by combining both pressure and proximity sensing to enable gestural interaction with one thumb (Figure 9 -(Right)). Combined with a gesture recognizer that takes the hand's anatomy into account, the Power-up Button can recognize six different mid-air gestures performed on the side of a mobile device. This gives it, for instance, enough expressive power to provide full one-handed control of interface widgets displayed on screen. This technology can complement touch input, and can be particularly useful when interacting eyes-free. It also opens up a larger design space for widget organization on screen: the button enables a more compact layout of interface components than what touch input alone would allow. This can be useful when, e.g., filling the numerous fields of a long Web form, or for very small devices.

Gimlenses – Complex 3D virtual scenes such as CAD models of airplanes and representations of the human body are notoriously hard to visualize. Those models are made of many parts, pieces and layers of varying size, that partially occlude or even fully surround one another. Gimlenses [28] provides a multi-view, detail-in-context visualization technique that enables users to navigate complex 3D models by interactively drilling holes into their outer layers to reveal objects that are buried, possibly deep, into the scene (see Figure 10). These holes are constantly adjusted so as to guarantee the visibility of objects of interest from the parent view. Gimlenses can be cascaded and constrained with respect to one another, providing synchronized, complementary viewpoints on the scene. Gimlenses enable users to quickly identify elements of interest, get detailed views of those elements, relate them, and put them in a broader spatial context.

Dashboard Exploration – Visual stories help us communicate knowledge, share and interpret experiences and have become a focus in visualization research in recent years. In this paper we discuss the use of storytelling

../../../../projets/in-situ/IMG/sidepress.png

../../../../projets/in-situ/IMG/powerup.png

Figure 9. (left) The SidePress prototype has two pressure sensors on one of its sides ; (right) The Power-Up button prototype is equipped with a button for sensing both pressure and proximity.

in Business Intelligence (BI) analysis [21] (Best Paper Award). We derive the actual practices in creating and sharing BI stories from in-depth interviews with expert BI analysts (both story “creators” and “readers”). These interviews revealed the need to extend current BI visual analysis applications to enable storytelling, as well as new requirements related to BI visual storytelling. Based on these requirements we designed and implemented a storytelling prototype tool with appropriate interaction techniques, that is integrated in an analysis tool used by our experts, and allows easy transition from analysis to story creation and sharing. We report experts' recommendations and reactions to the use of the prototype to create stories, as well as novices' reactions to reading these stories.

Hybrid-Image Visualizations – Data analysis scenarios often incorporate one or more displays with sufficiently large size and resolution to be comfortably viewed by different people from various distances. Hybrid-image visualizations [15] blend two different visual representations into a single static view, such that each representation can be perceived at a different viewing distance. They can thus be used to enhance overview tasks from a distance and detail-in-context tasks when standing close to the display. Viewers interact implicitly with these visualizations by walking around the space. By taking advantage of humans' perceptual capabilities, hybrid-image visualizations show different content to viewers depending on their placement, without requiring tracking of viewers in front of a display. Moreover, because hybrid-images use a perception-based blending approach, visualizations intended for different distances can each utilize the entire display.

Evolutionary Visual Exploration – In a high-dimensionality context, the visual exploration of information is challenging, as viewers are often faced with a large space of alternative views on the data. We present [14], a system that combines visual analytics with stochastic optimization to aid the exploration of multidimensional datasets characterized by a large number of possible views or projections. Starting from dimensions whose values are automatically calculated by a PCA, an interactive evolutionary algorithm progressively builds (or evolves) non-trivial viewpoints in the form of linear and non-linear dimension combinations, to help users discover new interesting views and relationships in their data. The system calibrates a fitness function (optimized by the evolutionary algorithm) to take into account the user interactions to calculate and propose new views. Our method leverages automatic tools to detect interesting visual features and human interpretation to derive meaning, validate the findings and guide the exploration without having to grasp advanced statistical concepts. Our prototype was evaluated through an observational study with five domain experts, and helped

../../../../projets/in-situ/IMG/gimlenses-teaser.png

Figure 10. Exploring the CAD drawing of a car engine. The three Gimlenses provide detailed views of different constituent parts of the engine, at different magnification levels and with varying orientation, while revealing their location inside the global 3D model. (a) Context view. (b) Magnified side view of a knot behind, and thus originally hidden by, the cylinder head cover. (c) View fully revealing a poppet valve in-context from a different angle than the main view, with (d) another Gimlens configured so as to provide a low-angled point of view on the valve.

them quantify qualitative hypotheses, try out different scenarios to dynamically transform their data, and to better formulate their research questions and build new hypotheses for further investigation.

6.2. Research Methods

Participants: Michel Beaudouin-Lafon, Anastasia Bezerianos, Jérémie Garcia, Stéphane Huot, Ilaria Liccardi, Wendy Mackay [correspondant].

Conducting empirical research is a fundamental part of InSitu's research activities, including observation of users in field and laboratory settings to discover problems faced by users, controlled laboratory experiments to evaluate the effectiveness of the technologies we develop, longitudinal field studies to determine how our technologies work in the real world, and participatory design, to explore design possibilities with users throughout the design process.

However, we not only use research methods, we also investigate and develop them. As organizers of the CHI'13 conference in Paris, which had record-breaking numbers of submissions (over 2000) and participants (3500), we instituted a number of innovations in both the process of creating the program and presenting information to conference participants. In collaboration with researchers at MIT, we introduced an "author-sourcing" process (with an 87% participation rate) for collecting affinity data. We then developed a collaborative, interactive, visualization system on the WILD wall display, combined with the *Cobi* interactive constraint-solving system, that enabled us to resolve all presenter conflicts and successfully place all 500+ papers and events in relevant sessions ([35], [26]). We also replaced the "CHI Madness" series of 25-second presentations with "Video Previews", in which each research paper, course, panel or other event has a 30-second video preview. These are now available on the CHI'13 website, the ACM/CHI YouTube channel and in the ACM Digital Library, before the paywall. We also developed and field-tested the Interactive Schedule on large, interactive displays, which allowed conference attendees to both view upcoming Video Previews and use their mobile phones to search for particular content and create customized playlists [29]. We also developed two interactive table-top applications that were presented at CHI'13 Interactivity, that allowed attendees to visualize and explore conference events as well as to create customized video playlists.

The RepliCHI workshop at CHI'13, co-organized by Wendy Mackay, examined issues with respect to encouraging replication of controlled experiments, and introduced the RepliCHI award to top research articles that offer strong empirical contributions that include replication. She also organized a session called *Interacting with CHI* in which participants explained the technologies and processes they developed to support the CHI conference design and execution.

In the context of our work with Interactive Paper to support music composition, we developed Paper Tonnetz, a paper-based interface to composing melodies and chords based on musical patterns expressed in Euler's Tonnetz, and demonstrated it at CHI'13 Interactivity ([22]). We also explored how to create an interactive event for the "Fête de la Science", called "Design Me a Sound Landscape", in which participants can create their own ways of expressing a landscape and add natural sounds, such as wind, rain, moving water, that another participant can experience as they move on an interactive floor. Finally, we explored the drawing process, with the *Drawing Assistant* ([25]) in which users receive guidance and feedback as they learn to draw from photographs.

6.3. Engineering of interactive systems

Participants: Caroline Appert, Michel Beaudouin-Lafon [correspondant], Olivier Chapuis, Stéphane Huot, Wendy Mackay.

InSitu has a long tradition of developing software tools and user interface toolkits to facilitate the creation of interactive systems. These tools allow us to better experiment with our ideas and are therefore an integral part of our research methodology. Most of them are freely available and some are used outside InSitu for research or teaching.

Our work has focused on developing middleware for the WILD platform, InSitu's experimental ultra-high-resolution interactive room for studying collaborative interaction and the visualization of very large datasets [2]. WILD features a wall-sized display with 32 monitors, a multitouch table, a motion-tracking system and various mobile devices. Running applications on WILD requires developing advanced distributed systems that coordinate, in real time, the 16 computers of the cluster driving the wall display with a variety of clients and servers running on other computers, including mobile devices.

We investigated the use of Web standards and protocols to develop and deploy such applications. Hydrascope [24] introduces the concept of *meta-application* that combines, adapts and/or repurposes existing web applications for an environment such as WILD. It uses a web browser (or even a web engine, e.g. WebKit) as a rendering and interaction toolkit and Web protocols (HTTP and WebSockets) for communication. We demonstrated how to control a wall-size presentation tool built on Google Present and a wall-size map built on Google Maps without modifying these applications but by taking advantage of the capability of web applications for introspection.

This approach was used to develop CHIWall, an application designed to help us schedule the CHI 2013 conference that InSitu chaired this year in Paris. The resulting tool combines a wall-size display of the full program with a constraint-detection and constraint-solving assistant called Cobi [26], which itself uses crowd-sourced information from the authors. The resulting application supports collaborative work to fine-tune the program (Figure 11) and features a flexible architecture that has been reused for other prototype applications.

In summary, InSitu has continued to make significant progress towards mature toolkits that support post-WIMP and distributed user interfaces. These toolkits, in turn, have enabled us to experiment with novel interaction techniques using rapid prototyping. Conversely, our work on novel interaction techniques has driven the development of software toolkits that embody their underlying principles, facilitating further exploration. This back-and-forth between techniques, methods and tools is a defining feature of InSitu, captured by the *Designeering Interaction* [11] framework. As the focus of our research on interaction techniques has shifted from on-the-desktop to off-the-desktop, this approach has proven more valid than ever: improving interaction in such environments requires more complex software architectures and tools; in turn, these tools and architectures are a key step to getting these technologies outside the lab.

../../../../projets/in-situ/IMG/CHIwall.png

Figure 11. A group of users interacting with the CHIwall application to fine tune the schedule of the CHI 2013 conference (500 events (400 research articles) in 200 sessions, in 16 parallel tracks over four days). The tablet interface provides additional detail about the content of a session, such as the abstract and the details of the affinity data crowd-sourced from the authors.

INDES Project-Team

6. New Results

6.1. Security

Participants: Ilaria Castellani, Bernard Serpette [correspondant], José Santos.

6.1.1. Stateful Declassification Policies for Event-Driven Programs

We propose a novel mechanism for enforcing information flow policies with support for declassification on event-driven programs. Declassification policies consist of two functions. First, a projection function specifies for each confidential event what information in the event can be declassified directly. Second, a stateful release function specifies the aggregate information about all confidential events seen so far that can be declassified. We provide evidence that such declassification policies are useful in the context of JavaScript web applications. An enforcement mechanism for our policies is presented and its soundness and precision are proven. Finally, we give evidence of practicality by implementing and evaluating the mechanism in a browser.
Report and mechanization can be found in http://people.cs.kuleuven.be/~mathy.vanhoef/declass.

6.1.2. A Monitor Inlining Compiler for Securing JavaScript Programs

JavaScript applications can include untrusted code dynamically loaded from third party code providers (such as online advertisements). This issue raises the need for enforcement mechanisms to ensure security properties for JavaScript programs. The dynamic nature of the JavaScript programming language makes it a hard target for static analysis. Hence, research on mechanisms for enforcing security properties for JavaScript programs has mostly focused on dynamic approaches, such as runtime monitoring and program instrumentation. We design and implement a novel compiler that inlines a security monitor and we formally prove it correct with respect to an information flow security property. To the best of our knowledge, it is the first proven correct information flow monitor inlining transformation for JavaScript programs.
Report can be found in http://www-sop.inria.fr/indes/ifJS. See also software section.

6.1.3. Modular Extensions of Security Monitors for Web APIs: The DOM API Case Study

JavaScript programs often interact with the web page on which they are included, as well as with the browser itself, through external APIs such as the DOM API, the XMLHttpRequest API, and the W3C Geolocation API. The continuous emergence and heterogeneity of different external APIs renders the problem of precisely reasoning about JavaScript security particularly challenging. To tackle this problem, we propose a methodology for extending arbitrary sound JavaScript monitors. The methodology allows us to prove noninterference for external APIs in a modular way. Thus, when considering new external APIs, the noninterference property of the security monitor still holds. We present two groups of DOM interfaces that illustrate how to extend a noninterferent monitor model with: (1) basic DOM methods, for which we have discovered new information leaks not explored in previous work; (2) live collections, which are special features of the DOM API with an unconventional semantics that can lead to several previously unknown information leaks. Finally, we inline an extensible noninterferent JavaScript monitor that handles (1) and (2), and we make it available online
Report can be found in http://www-sop.inria.fr/indes/ifJS.

6.1.4. A Certified Lightweight Non-Interference Java Bytecode Verifier

We propose a type system to verify the non-interference property in the Java Virtual Machine. We verify the system in the Coq theorem prover.

Noninterference guarantees the absence of illicit information flow throughout program execution. It can be enforced by appropriate information flow type systems. Much of the previous work on type systems for non-interference has focused on calculi or high-level programming languages, and existing type systems for low-level languages typically omit objects, exceptions and method calls. We define an information flow type system for a sequential JVM-like language that includes all these programming features, and we prove, in the Coq proof assistant, that it guarantees non-interference. An additional benefit of the formalisation is that we have extracted from our proof a certified lightweight bytecode verifier for information flow. Our work provides, to the best of our knowledge, the first sound and certified information flow type system for such an expressive fragment of the JVM.

This work appeared in the journal of Mathematical Structures in Computer Science [9].

6.1.5. Session types for liveness and security

Within the COST Action BETTY, we have started studying the interplay between liveness properties and secure information flow properties in session calculi, in collaboration with a colleague from Torino University. Recent developments in static analysis techniques have shown that behavioural types, and in particular session types, may be used to enforce liveness properties of communicating systems. Examples of such properties are deadlock freedom, eventual message delivery and session termination. Because secure information flow in communicating systems depends on the observation of messages, there is a clear connection between information flow analysis and the liveness properties of the systems under consideration. We have been examining the joint application of liveness enforcement and secure information flow analysis in session calculi. It appears that, by strengthening the assumptions on the liveness of systems, it is possible to relax the conditions under which a system satisfies secure information flow properties. This is ongoing work, which is expected to continue within the BETTY Action.

6.1.6. Noninterference in reactive synchronous languages

We defined two properties of Reactive Noninterference (RNI) for a core synchronous reactive language called CRL formalising secure information flow. Both properties are time-insensitive and termination-insensitive. Again, coarse-grained RNI is more abstract than fine-grained RNI.

Finally, a type system guaranteeing both security properties was presented. Thanks to a design choice of CRL, which offers two separate constructs for loops and iteration, and to refined typing rules, this type system allows for a precise treatment of termination leaks, which are an issue in parallel languages.

This work has been presented at the International Symposium on Trustworthy Global Computing (TGC 2013) [11]. It is also described in Attar's PhD thesis pejman:tel-00920152.

6.2. Models, semantics, and languages

Participants: Pejman Attar, Gérard Berry, Gérard Boudol, Ilaria Castellani, Johan Grande, Cyprien Nicolas, Tamara Rezk, Manuel Serrano [correspondant].

6.2.1. Formalization and Concretization of Ordered Networks

Overlay networks have been extensively studied as a solution to the dynamic nature, scale and heterogeneity of large computing platforms, and are a fundamental layers of most existing peer-to-peer networks. The basic mechanism offered by an overlay network, is routing, i.e., the mechanism enabling the delivery of messages from any node to any other node in the network. On top of routing are built crucial functionalities of peer-to-peer networks, such as networks maintenance (nodes joining and leaving the network) and information distribution and retrieval. Over the years, different topologies and routing mechanisms have been proposed in literature. However, there is a lack of formal works unifying these different designs and establishing their correctness. This paper proposes a formal common basis, partially validated with the Coq theorem prover, with the nice property of only requiring the definition of a total order on the nodes. We investigate how such a basic design can be used to build deadlock/livelock-free algorithms for routing, node insertion, and node deletion in the fault-free environment. The genericity of our design is then explored through the construction of orders on nodes corre- sponding to different topologies commonly encountered in the peer-to-peer domain. To validate the methodology proposed, a simulator tool was developed. This tool is able, given the definition of an order and the definition of shortcuts, to simulate the corresponding overlay network and to explore its performance.

6.2.2. Absence Prediction in Esterel

We have formally proved, with the Coq system, the correctness of an absence prediction of Esterel's signals. For this we have formalised in Coq the static analysis and the interpreter written in Scheme (see the previous activity report). With this formal specification, we prove the correctness of the analysis: if a signal is considered absent by the evaluator at one instant, then this signal will be not emitted during this instant. This work is described in a currently submitted paper.

6.2.3. Reactive Synchronous Languages

CRL: We have studied the security property of noninterference in a synchronous Core Reactive Language (CRL). In the synchronous reactive paradigm, programs communicate by means of broadcast events, and their parallel execution is regulated by a notion of instant.

We have first shown that CRL programs are indeed reactive, namely that they always converge to a state of termination or suspension ("end of instant") in a finite number of steps. This property is important as it also entails the reactivity of a program to its environment, namely its capacity to input events from the environment at the start of instants, and to output events to the environment at the end of instants. While classical in synchronous languages, this property required to be established afresh in CRL, since this language makes use of a new asymmetric parallel operator.

We defined two bisimulation equivalences on CRL programs, corresponding respectively to a fine-grained and to a coarse-grained observation of programs. We showed that coarse-grained bisimilarity is more abstract than fine-grained bisimilarity, as it is insensitive to the order of generation of events and to repeated emissions of the same event during an instant.

DSLM :

We have finalised our work on the language DSLM (Dynamic Synchronous Language with Memory), which is an extension of CRL with memory and distribution. There are now several sites, and agents may migrate between sites. Two main properties are established for DSLM: reactivity of each agent and absence of data-races between agents. Since DSLM uses the same asymmetric parallel operator as CRL, reactivity is proven in a similar way. Moreover, the language offers a way to benefit from multi-core and multi-processor architectures, by means of the notion of synchronized scheduler, which abstractly models a computing resource. Each site may be expanded and contracted dynamically by varying its number of synchronized schedulers. Moreover agents can be moved transparently from one scheduler to another one within the same site. In this way one can formally model the load-balancing of agents over a site. This work is part of Pejman Attar's PhD thesis, defended in December 2013.

6.2.4. Locking Fast

We have studied the integration of low-level locking mechanisms in programming language execution environments. We have shown that for a given low-level locking mechanism the performance of the applications may vary significantly according to decisions taken for integrating it in the runtime system. We have studied two different aspects. First, we have shown how to accelerate C IO locking by selecting at runtime the adequate implementation and by using spin locks instead of full-fledged mutexes. Second, we have presented a new schema for improving the slow path of Java-like synchronized blocks. It consists in lifting the exception handler that is installed on the stack and which is in charge of releasing a monitor up to the closest exception handler already installed on the stack. All these optimizations have been implemented in Hop, our Web programming language. We have conducted experiments that shows significant speed up (up to 30%) for applications using locks extensively.

The synchronization lifting technique could be generalized to all the exception handlers, not only the handlers of synchronized blocks. As lifting only modifies the interception of exceptions, not the way they are thrown, it is compatible with languages such as Java or JavaScript that store a description of the stack at the moment when the exception is thrown inside the exception handlers. The technique should thus be broadly applicable. Exploring this idea is left for future work.

This work is described in the paper that will be published in the proceedings of the SAC'14 conference [12].

6.2.5. JThread

The jthread library is a library for Hop offering threads and mutexes and whose main locking function implements deadlock avoidance. Our library offers structured locking (i.e., critical sections instead of explicit `lock`/`unlock` functions). It supports nested locking. Our library is implemented using the preexisting pthread library and is offered as an alternative to the latter.

Compared to usual locking functions, our primitive relies on the programmer to provide some supplementary information such as the set of mutexes that might be acquired while owning a first one. However, for this supplementary information we chose default values that limit the need for the programmer to actually write it to a minimum.

The syntax of our locking construct is as follows:

(synchronize* l [:prelock p]
expr1
expr2
...)

where `l` is the list of mutexes to lock and `p` is a list that contains (some of [1]) the mutexes that might be locked during the execution of the body of the construct.

The implementation of this function relies on the ability to lock n mutexes at once. We found an algorithm for this that is both deadlock-free and starvation-free. Our algorithm relies on a dynamic total ordering of threads; this is inspired by Lamport's bakery algorithm.

We wrote a starvation-freedom property that applies to our real-life language with dynamic thread creation and programs that run forever on purpose. To express the property we need to define the following relation over threads:

$t_1 prec t_2$ iff. $existsm.t_1$ owns m and t_2 is waiting to lock m.

Let $prec^*$ be the symmetric transitive closure of $prec$.

The property that we chose and proved for our algorithm is:

If each non-waiting thread eventually releases all the mutexes it owns **and if** for each waiting thread t the number of threads t' s.t. $t' prec^* t$ does not tend toward $+infty$ over time **then** each waiting thread eventually gets the mutexes it is waiting to lock.

We have implemented our library and integrated it to Hop. We haven't released it yet. An article is in preparation.

6.3. Web programming

Participants: Gérard Berry, Yoann Couillec, Ludovic Courtès, Cyprien Nicolas, Vincent Prunet, Tamara Rezk, Marcela Rivera, Bernard Serpette, Manuel Serrano [correspondant].

[1] According to a few rules that we impose

6.3.1. Colored λ-calculus

We have extended the bicolored λ-calculus to a polychromic one. With two colors, we were able to abstract the Hop language with its '`$`' and '`~`' annotations. With more than two colors, we can also modele embedded languages as a query based language, for example. As for the bicolored version, we have defined a static transformation aggregating expressions of the same color. We have formally proved, with the Coq system, the correctness, the confluence and the terminaison of the transformation. This work has been accepted for publication at the conference JFLA'14. [14].

6.3.2. Multitier Debugging

The distributed nature of Web applications makes debugging difficult. The programming languages and tools commonly used make it even more complex. Generally the server-side and the client-side are implemented in different settings and the debugging is treated as two separated tasks: on the one hand, the debugging of the server, on the other hand, the debugging of the client. Most studies and tools focus on this last aspect. They concentrate on the debugging of JavaScript in the browser. Although useful, this only addresses one half of the problem. Considering the debugging of Web applications as a whole raises the following difficulties:

- As the server-side and the client-side are generally implemented in different languages, debuggers for the Web do not capture the whole execution of the application. Programming the server and the client in the same language helps but is not sufficient to let the debugger expose a coherent view of the whole execution as this also demands a runtime environment that enforces consistent representations of data structures and execution traces.
- The JavaScript tolerant semantics tends to defer errors raising. For instance, calling a function with an insufficient number of arguments may lead to filling a data structure with the unexpected `undefined` value which, in turn, may raise a type error when accessed. The *distance* between the error and its actual cause may be arbitrarily long which can make the relation between the two difficult to establish.
- The JavaScript event loop used for the GUI splits the execution into unrelated callback procedures which get called upon event receipts. When an error occurs, the active stack trace only contains elements relative to the current callback invocation. It is oblivious of the context of the callback. Understanding the cause of the error is then not easy.

Pursuing our research on multitier programming for the Web, we have built a programming environment which eliminates most of these problems.

- When an error is raised, the full stack trace is reported. This stack trace might contain server stack frames, client stack frames, or both. We call this a *multitier stack trace*.
- When an error occurs, either on the client or on the server, its source location is reported by the debugger.
- In *debugging mode*, types, arities, and array bounds, are strictly enforced on the server and on the client. Hence, when the execution of the program deviates from the formal semantics of the language, an error is raised immediately.

A paper currently submitted presents this debugger and exposes the salient aspects of its implementation is under submission.

6.3.3. Hop and HipHop : Multitier Web Orchestration

Our aim is to help programming rich applications driven by computers, smartphones or tablets; since they interact with various external services and devices, such applications require orchestration techniques that merge classical computing, client-server concurrency, web-based interfaces, and event-based programming. To achieve this, we extend the Hop multitier web programming platform [5] by the new HipHop domain specific language (DSL), which is based on the synchronous language Esterel. HipHop orchestrates and synchronizes internal and external activities according to timers, events generated by the network, GUIs, sensors and devices, or internally computed conditions.

Like Esterel, Hiphop is a concurrent language based on the perfect synchrony hypothesis: a HipHop program repeatedly reacts in conceptual zero-delay to input events by generating output events; synchronization and communication between parallel statements is also performed in conceptual zero-delay. Perfect synchrony makes concurrent programs deterministic and deadlock-free, the only non-determinism left being that of the application environment. Its implementation is cycle-based, execution consisting of repeated atomic cycles "read inputs / compute reaction / generate outputs" in coroutine with the main Hop code. Concurrency is compiled away by static or dynamic sequential scheduling of code fragments. Cyclic execution atomicity avoids interference between computation and input-output, which is the usual source of unexpected non-determinism and synchronization problems for classical event-handler based programming.

While Esterel is limited to static applications, HipHop is designed for dynamicity. Its implementation on top of Hop makes it possible to dynamically build and run orchestration programs at any time using Hop's reflexivity facilities. It even makes it possible to modify a HipHop program between two execution cycles. It also simplifies the language by importing Hop's data definition facilities, expressions, modular structure, and higher-order programming features. It relies on the Web asynchronous concurrency and messaging already supported by Hop.

Using HipHop for real-life applications such as multimedia applications has been presented in an invited paper of the conference ICDCIT'14 [10].

This year, we extended the HipHop language with dynamic constructions, namely `genpar&` and `dyngenpar&`. These new constructs allow HipHop applications to parallelize treatment of event's values without knowing *a priori* the number of values carried by a given event. genpar& is alike a delayed parallel map execution, while dyngenpar& may create new parallel branches on-demand.

HipHop was also extended with listeners, alike HTML/DOM ones. The programmer can attach functions to any element of the HipHop programe that will be triggered when an instruction is started, suspended, resumed, terminated or aborted. These listeners enable us to trace a specific part of a program, easing its debugging.

6.3.4. Hop Programming Environment

In Linux-based environments, Hop is launched using the command line or using OS init scripts. This is inadequate for the Mac OS environment where graphical user interfaces are generally used to start, stop, and control applications. To fit the Mac OS users habits we have implemented a graphical front-end to Hop. It allows users to monitor and manage Hop processes. The implemented high level graphical interface is a XCode project which has been developed in objective-C for Mac OS X 10.7/10.8.

The main functionalities developed in this graphical front-end are:

- easily manage Hop processes. This GUI allows users to start, to stop and to restart the execution of Hop processes in a simplified manner. This process is executed in an independent thread in order to prevent impacts in the main program. Even though the process becomes independent, the main program can still control the execution of the Hop process by stopping or restarting it.
- Display messages in a user-frienly way. All messages as well as standard and error outputs, generated after the launch Hop server, are captured and redirected to be displayed in the main program. This allows the user to monitor the execution of the Hop process at any time. In this way, the user can search for a specific output or display pattern. All messages can be saved in external files for further analysis.
- Launch a Hop process with specific settings. It is possible to specify the port number on which the Hop server will accept connection. The verbosity and debugging level can also be specified according to the needs. At the same time the user can activate/deactivate specific options like Zeroconf and Webdav.
- Specify additional arguments to run a Hop process in a "command line like" way. In particular situations, advanced users could need to specify some options when launching the Hop server.

Additionally, a set of scripts have been developed to facilitate the generation and distribution of this work. A group of scripts allows one to compile and build the Hop GUI without needing a graphic Xcode interface. In this way, it is possible to generate an application bundle in a local machine as well as in a remote one without needing additional graphical interfaces.

The other group of scripts allows one to generate a “ready to use” dmg image containing Hop files. This dmg image can either include the graphical user interface (GUI) or not. Due to the continuous evolution of Hop based on new requirements and bugs fixed, the latter set of scripts provides a powerful tool to improve the releasing of new product versions.

This front-end has been integrated in the main Hop development tree. The MacOS pre-compiled version is publicly available on the Hop web site http://hop.inria.fr.

6.3.5. Web of Data

We are extending the Hop programming language in order to improve its data management: the amount of data it can access, the increasing number of sources of data and the heterogeneity of data it can accept. We have made an implementation of the SPARQL query language and the ORC orchestration language in Hop. We have written a configurable interpreter of the ORC language. The parallelism of the interpreter can be activated or not, for each operator of the ORC language. This specificity allows different executions of an ORC application, depending on the execution context or constraints, such as an execution on a client which disallows any parallelism. Within the X-Data project, we have participated in the development of a data intensive application in collaboration with Data Publica, the leading company of the project, and with the Inria Zenith research team. This application analyzed data sets provided by the French Insee institute to exhibit population commuting patterns. Our incentive for participating in this development was to acquire knowledge on programming data intensive applications. In the mid-term, we will rest on this expertise to create new data-aware programming languages or programming language extensions.

6.4. Web robotics

Participants: Ludovic Courtès, Cyprien Nicolas, Vincent Prunet [correspondant], Manuel Serrano.

6.4.1. Cable driven robots

The sound design of modern robotic applications demands for the configuration-time integration of various subsystems which together constitute a robot. APIs and protocols such as ROS (Robot Operating System) provide robot designers with tools to combine software and hardware components into a complete robot. In addition, more and more robots need to share information or interact with diffuse objects available in the robot neighborhood, and also with remote services, to log data (typically activity monitoring data in the case of an assistance robot), to send information messages (alarms or triggering events to some other infrastructure), to subscribe to services provided by objects or remote servers, to provide services that may help peer entities, to get new behaviors by downloading and installing applications within the robot. We develop tools and architectures to address these requirements using Hop as our main platform. We experiment software architectures involving robots, web objects, several integration models with third party components (hardware, software computation libraries for robotics, stand-alone robots), protocols, and libraries.

We pursued the joint work with Coprin Team about using Hop to coordinate a cable-driven robot. We changed the hardware on which Hop runs to a mini-PC instead of a standard laptop, plugged a wireless router, and used the wireless network from a tablet to move the robot. We also improved the robot hardware and software. The setup has been summarized in a paper [13] and presented at a national conference on robotics.

6.4.2. Web Robotics

Web Robotics is a two years Inria ADT project targeting the development of technical foundations (libraries and toolkits) and demos of web enabled robots. The project is led by Indes (Vincent Prunet, Manuel Serrano), software development is supported by Inria SED (Ludovic Courtès), robots are provided by Inria Coprin. A demonstrator and dissemination platform consisting of a cable robot and dedicated web services have been set up to enable people to interact with the robot through a web server (web http://webrobotics.inria.fr:8080/hop/welcome).

../../../../projets/indes/IMG/macosx.png

Figure 1. A screenshot of the MacOS X Hop graphical interface.

../../../../projets/indes/IMG/webrobotics.png

Figure 2. A web-controlled cable robot. The whole application is implemented with Hop.

The web robotics demo demonstrates:

- the programming of robot control functions within Hop
- simulation/real hardware abstraction
- hardware control (Phidget integration)
- integration of specialized robotics libraries
- multi server architecture
- management of multiple users and authenticated access to critical resources.

Hop added value is to provide:

- an unconstrained specification environment where data and services are easily shared among servers and clients;
- a seamless, plugin free, integration into standard web browsers.

Also in 2013, Indes has joined the PAL (Person Assisted Living) Inria project, to develop web enabled applications within the project.

IPSO Project-Team

5. New Results

5.1. Multi-revolution composition methods for highly oscillatory differential equations

In [45], we introduce a new class of multi-revolution composition methods (MRCM) for the approximation of the Nth-iterate of a given near-identity map. When applied to the numerical integration of highly oscillatory systems of differential equations, the technique benefits from the properties of standard composition methods: it is intrinsically geometric and well-suited for Hamiltonian or divergence-free equations for instance. We prove error estimates with error constants that are independent of the oscillatory frequency. Numerical experiments, in particular for the nonlinear Schrödinger equation, illustrate the theoretical results, as well as the efficiency and versatility of the methods.

5.2. Weak second order multi-revolution composition methods for highly oscillatory stochastic differential equations with additive or multiplicative noise

In [61], we introduce a class of numerical methods for highly oscillatory systems of stochastic differential equations with general noncommutative noise. We prove global weak error bounds of order two uniformly with respect to the stiffness of the oscillations, which permits to use large time steps. The approach is based on the micro-macro framework of multi-revolution composition methods recently introduced for deterministic problems and inherits its geometric features, in particular to design integrators preserving exactly quadratic first integral. Numerical experiments, including the stochastic nonlinear Schrödinger equation with space-time multiplicative noise, illustrate the performance and versatility of the approach.

5.3. High order numerical approximation of the invariant measure of ergodic SDEs

In [41], we introduce new sufficient conditions for a numerical method to approximate with high order of accuracy the invariant measure of an ergodic system of stochastic differential equations, independently of the weak order of accuracy of the method. We then present a systematic procedure based on the framework of modified differential equations for the construction of stochastic integrators that capture the invariant measure of a wide class of ergodic SDEs (Brownian and Langevin dynamics) with an accuracy independent of the weak order of the underlying method. Numerical experiments confirm our theoretical findings.

5.4. PIROCK: a swiss-knife partitioned implicit-explicit orthogonal Runge-Kutta Chebyshev integrator for stiff diffusion-advection-reaction problems with or without noise

In [13], a partitioned implicit-explicit orthogonal Runge-Kutta method (PIROCK) is proposed for the time integration of diffusion-advection-reaction problems with possibly severely stiff reaction terms and stiff stochastic terms. The diffusion terms are solved by the explicit second order orthogonal Chebyshev method (ROCK2), while the stiff reaction terms (solved implicitly) and the advection and noise terms (solved explicitly) are integrated in the algorithm as finishing procedures. It is shown that the various coupling (between diffusion, reaction, advection and noise) can be stabilized in the PIROCK method. The method, implemented in a single black-box code that is fully adaptive, provides error estimators for the various terms present in the problem, and requires from the user solely the right-hand side of the differential equation. Numerical experiments and comparisons with existing Chebyshev methods, IMEX methods and partitioned methods show the efficiency and flexibility of our new algorithm.

5.5. An offline-online homogenization strategy to solve quasilinear two-scale problems at the cost of one-scale problems

In [39], inspired by recent analyses of the finite element heterogeneous multiscale method and the reduced basis technique for nonlinear problems, we present a simple and concise finite element algorithm for the reliable and efficient resolution of elliptic or parabolic multiscale problems of nonmonotone type. Solutions of appropriate cell problems on sampling domains are selected by a greedy algorithm in an offline stage and assembled in a reduced basis (RB). This RB is then used in an online stage to solve two-scale problems at a computational cost comparable to the single-scale case. Both the offline and the online cost are independent of the smallest scale in the physical problem. The performance and accuracy of the algorithm are illustrated on 2D and 3D stationary and evolutionary nonlinear multiscale problems.

5.6. Reduced basis finite element heterogeneous multiscale method for quasilinear elliptic homogenization problems

In [40], a reduced basis nite element heterogeneous multiscale method (RB-FE-HMM) for a class of nonlinear homogenization elliptic problems of nonmonotone type is introduced. In this approach, the solutions of the micro problems needed to estimate the macroscopic data of the homogenized problem are selected by a Greedy algorithm and computed in an online stage. It is shown that the use of reduced basis (RB) for nonlinear numerical homogenization reduces considerably the computational cost of the nite element heterogeneous multiscale method (FE-HMM). As the precomputed microscopic functions depend nonlinearly on the macroscopic solution, we introduce a new a posteriori error estimator for the Greedy algorithm that guarantees the convergence of the online Newton method. A priori error estimates and uniqueness of the numerical solution are also established. Numerical experiments illustrate the e ciency of the proposed method.

5.7. Weak second order explicit stabilized methods for stiff stochastic differential equations

In [16], we introduce a new family of explicit integrators for stiff Itô stochastic differential equations (SDEs) of weak order two. These numerical methods belong to the class of one-step stabilized methods with extended stability domains and do not suffer from the stepsize reduction faced by standard explicit methods. The family is based on the standard second order orthogonal Runge-Kutta Chebyshev methods (ROCK2) for deterministic problems. The convergence, and the mean-square and asymptotic stability properties of the methods are analyzed. Numerical experiments, including applications to nonlinear SDEs and parabolic stochastic partial differential equations are presented and confirm the theoretical results.

5.8. Mean-square A-stable diagonally drift-implicit integrators of weak second order for stiff Itô stochastic differential equations

In [15], we introduce two drift-diagonally-implicit and derivative-free integrators for stiff systems of Itô stochastic differential equations with general non-commutative noise which have weak order 2 and deterministic order 2, 3, respectively. The methods are shown to be mean-square A-stable for the usual complex scalar linear test problem with multiplicative noise and improve significantly the stability properties of the drift-diagonally-implicit methods previously introduced [K. Debrabant and A. Rößler, Appl. Num. Math., 59, 2009].

5.9. Two-Scale Macro-Micro decomposition of the Vlasov equation with a strong magnetic field

In [25], we build a Two-Scale Macro-Micro decomposition of the Vlasov equation with a strong magnetic field. This consists in writing the solution of this equation as a sum of two oscillating functions with circonscribed oscillations. The first of these functions has a shape which is close to the shape of the Two-Scale limit of the solution and the second one is a correction built to offset this imposed shape. The aim of such a decomposition is to be the starting point for the construction of Two-Scale Asymptotic-Preserving Schemes.

5.10. A dynamic multi-scale model for transient radiative transfer calculations

In [33], a dynamic multi-scale model which couples the transient radiative transfer equation (RTE) and the diffusion equation (DE) is proposed and validated. It is based on a domain decomposition method where the system is divided into a mesoscopic subdomain, where the RTE is solved, and a macroscopic subdomain where the DE is solved. A buffer zone is introduced between the mesoscopic and the macroscopic subdomains, as proposed by Degond and Jin, who solve a coupled system of two equations, one at the mesoscopic and the other at the macroscopic scale. The DE and the RTE are coupled through the equations inside the buffer zone, instead of being coupled through a geometric interface like in standard domain decomposition methods. One main advantage is that no boundary or interface conditions are needed for the DE. The model is compared to Monte Carlo, finite volume and P1 solutions in one dimensional stationary and transient test cases, and presents promising results in terms of trade-off between accuracy and computational requirements.

5.11. Quasi-periodic solutions of the 2D Euler equation

In [24], we consider the two-dimensional Euler equation with periodic boundary conditions. We construct time quasi-periodic solutions of this equation made of localized travelling profiles with compact support propagating over a stationary state depending on only one variable. The direction of propagation is orthogonal to this variable, and the support is concentrated on flat strips of the stationary state. The frequencies of the solution are given by the locally constant velocities associated with the stationary state.

5.12. Optimization and parallelization of Emedge3D on shared memory architecture

In [38], a study of techniques used to speedup a scientific simulation code is presented. The techniques include sequential optimizations as well as the parallelization with OpenMP. This work is carried out on two different multicore shared memory architectures, namely a cutting edge 8x8 core CPU and a more common 2x6 core board. Our target application is representative of many memory bound codes, and the techniques we present show how to overcome the burden of the memory bandwidth limit, which is quickly reached on multi-core or many-core with shared memory architectures. To achieve efficient speedups, strategies are applied to lower the computation costs, and to maximize the use of processors caches. Optimizations are: minimizing memory accesses, simplifying and reordering computations, and tiling loops. On 12 cores processor Intel X5675, aggregation of these optimizations results in an execution time 21.6 faster, compared to the original version on one core.

5.13. Vlasov on GPU (VOG Project)

In [58], we are concerned with the numerical simulation of the Vlasov-Poisson set of equations using semi- Lagrangian methods on Graphical Processing Units (GPU). To accomplish this goal, modifications to traditional methods had to be implemented. First and foremost, a reformulation of semi-Lagrangian methods is performed, which enables us to rewrite the governing equations as a circulant matrix operating on the vector of unknowns. This product calculation can be performed efficiently using FFT routines. Second, to overcome the limitation of single precision inherent in GPU, a δf type method is adopted which only needs refinement in specialized areas of phase space but not throughout. Thus, a GPU Vlasov-Poisson solver can indeed perform high precision simulations (since it uses very high order reconstruction methods and a large number of grid points in phase space). We show results for rather academic test cases on Landau damping and also for physically relevant phenomena such as the bump on tail instability and the simulation of Kinetic Electrostatic Electron Nonlinear (KEEN) waves.

5.14. Uniformly accurate numerical schemes for highly oscillatory Klein-Gordon and nonlinear Schrödinger equations

In [37], we are interested in the numerical simulation of nonlinear Schrödinger and Klein-Gordon equations. We present a general strategy to construct numerical schemes which are uniformly accurate with respect to the oscillation frequency. This is a stronger feature than the usual so called "Asymptotic preserving" property, the last being also satisfied by our scheme in the highly oscillatory limit. Our strategy enables to simulate the oscillatory problem without using any mesh or time step refinement, and the orders of our schemes are preserved uniformly in all regimes. In other words, since our numerical method is not based on the derivation and the simulation of asymptotic models, it works in the regime where the solution does not oscillate rapidly, in the highly oscillatory limit regime, and in the intermediate regime with the same order of accuracy. The method is based on two main ingredients. First, we embed our problem in a suitable "two-scale" reformulation with the introduction of an additional variable. Then a link is made with classical strategies based on Chapman-Enskog expansions in kinetic theory despite the dispersive context of the targeted equations, allowing to separate the fast time scale from the slow one. Uniformly accurate (UA) schemes are eventually derived from this new formulation and their properties and performances are assessed both theoretically and numerically.

5.15. Asymptotic preserving schemes for the Wigner-Poisson-BGK equations in the diffusion limit

In [26], we focus on the numerical simulation of the Wigner-Poisson-BGK equation in the diffusion asymptotics. Our strategy is based on a "micro-macro" decomposition, which leads to a system of equations that couple the macroscopic evolution (diffusion) to a microscopic kinetic contribution for the fluctuations. A semi-implicit discretization provides a numerical scheme which is stable with respect to the small parameter ε (mean free path) and which possesses the following properties: (i) it enjoys the asymptotic preserving property in the diffusive limit; (ii) it recovers a standard discretization of the Wigner-Poisson equation in the collisionless regime. Numerical experiments confirm the good behaviour of the numerical scheme in both regimes. The case of a spatially dependent $\varepsilon(x)$ is also investigated.

5.16. Existence and stability of solitons for fully discrete approximations of the nonlinear Schrödinger equation

In [19], we study the long time behavior of a discrete approximation in time and space of the cubic nonlinear Schrödinger equation on the real line. More precisely, we consider a symplectic time splitting integrator applied to a discrete nonlinear Schrödinger equation with additional Dirichlet boundary conditions on a large interval. We give conditions ensuring the existence of a numerical soliton which is close in energy norm to the continuous soliton. Such result is valid under a CFL condition between the time and space stepsizes. Furthermore we prove that if the initial datum is symmetric and close to the continuous soliton, then the associated numerical solution remains close to the orbit of the continuous soliton for very long times.

5.17. Asymptotic preserving schemes for the Klein-Gordon equation in the non-relativistic limit regime

In [32], we consider the Klein-Gordon equation in the non-relativistic limit regime, i.e. the speed of light c tending to infinity. We construct an asymptotic expansion for the solution with respect to the small parameter depending on the inverse of the square of the speed of light. As the first terms of this asymptotic can easily be simulated our approach allows us to construct numerical algorithms that are robust with respect to the large parameter c producing high oscillations in the exact solution.

5.18. Sobolev stability of plane wave solutions to the cubic nonlinear Schrödinger equation on a torus

In [31], it is shown that plane wave solutions to the cubic nonlinear Schrödinger equation on a torus behave orbitally stable under generic perturbations of the initial data that are small in a high-order Sobolev norm, over long times that extend to arbitrary negative powers of the smallness parameter. The perturbation stays small in the same Sobolev norm over such long times. The proof uses a Hamiltonian reduction and transformation and, alternatively, Birkhoff normal forms or modulated Fourier expansions in time.

5.19. Weak backward error analysis for overdamped Langevin equation

In [57], we consider an overdamped Langevin stochastic differential equation and show a weak backward error analysis result for its numerical approximations defined by implicit methods. In particular, we prove that the generator associated with the numerical solution coincides with the solution of a modified Kolmogorov equation up to high order terms with respect to the stepsize. This implies that every measure of the numerical scheme is close to a modified invariant measure obtained by asymptotic expansion. Moreover, we prove that, up to negligible terms, the dynamic associated with the implicit scheme considered is exponentially mixing.

5.20. Weak backward error analysis for Langevin equation

In [56], We consider numerical approximations of stochastic Langevin equations by implicit methods. We show a weak backward error analysis result in the sense that the generator associated with the numerical solution coincides with the solution of a modified Kolmogorov equation up to high order terms with respect to the stepsize. This implies that every measure of the numerical scheme is close to a modified invariant measure obtained by asymptotic expansion. Moreover, we prove that, up to negligible terms, the dynamic associated with the implicit scheme considered is exponentially mixing.

5.21. Approximation of the invariant law of SPDEs: error analysis using a Poisson equation for a full-discretization scheme

In [44], we study the long-time behavior of fully discretized semilinear SPDEs with additive space-time white noise, which admit a unique invariant probability measure μ. We show that the average of regular enough test functions with respect to the (possibly non unique) invariant laws of the approximations are close to the corresponding quantity for μ.

More precisely, we analyze the rate of the convergence with respect to the different discretization parameters. Here we focus on the discretization in time thanks to a scheme of Euler type, and on a Finite Element discretization in space.

The results rely on the use of a Poisson equation; we obtain that the rates of convergence for the invariant laws are given by the weak order of the discretization on finite time intervals: order $1/2$ with respect to the time-step and order 1 with respect to the mesh-size.

5.22. An asymptotic preserving scheme based on a new formulation for NLS in the semiclassical limit

In [20], we consider the semiclassical limit for the nonlinear Schrodinger equation. We introduce a phase/amplitude representation given by a system similar to the hydrodynamical formulation, whose novelty consists in including some asymptotically vanishing viscosity. We prove that the system is always locally well-posed in a class of Sobolev spaces, and globally well-posed for a fixed positive Planck constant in the one-dimensional case. We propose a second order numerical scheme which is asymptotic preserving. Before singularities appear in the limiting Euler equation, we recover the quadratic physical observables as well as the wave function with mesh size and time step independent of the Planck constant. This approach is also well suited to the linear Schrodinger equation.

5.23. Asymptotic Preserving schemes for highly oscillatory Vlasov-Poisson equations

The work [28] is devoted to the numerical simulation of a Vlasov-Poisson model describing a charged particle beam under the action of a rapidly oscillating external field. We construct an Asymptotic Preserving numerical scheme for this kinetic equation in the highly oscillatory limit. This scheme enables to simulate the problem without using any time step refinement technique. Moreover, since our numerical method is not based on the derivation of the simulation of asymptotic models, it works in the regime where the solution does not oscillate rapidly, and in the highly oscillatory regime as well. Our method is based on a "two scale" reformulation of the initial equation, with the introduction of an additional periodic variable.

5.24. Uniformly accurate numerical schemes for highly oscillatory Klein-Gordon and nonlinear Schrödinger equations

The work [37] is devoted to the numerical simulation of nonlinear Schrödinger and Klein-Gordon equations. We present a general strategy to construct numerical schemes which are uniformly accurate with respect to the oscillation frequency. This is a stronger feature than the usual so called ”Asymptotic preserving" property, the last being also satisfied by our scheme in the highly oscillatory limit. Our strategy enables to simulate the oscillatory problem without using any mesh or time step refinement, and the orders of our schemes are preserved uniformly in all regimes. In other words, since our numerical method is not based on the derivation and the simulation of asymptotic models, it works in the regime where the solution does not oscillate rapidly, in the highly oscillatory limit regime, and in the intermediate regime with the same order of accuracy. In the same spirit as in [28], the method is based on two main ingredients. First, we embed our problem in a suitable ”two-scale" reformulation with the introduction of an additional variable. Then a link is made with classical strategies based on Chapman-Enskog expansions in kinetic theory despite the dispersive context of the targeted equations, allowing to separate the fast time scale from the slow one. Uniformly accurate (UA) schemes are eventually derived from this new formulation and their properties and performances are assessed both theoretically and numerically.

5.25. On the controllability of quantum transport in an electronic nanostructure

In [59], we investigate the controllability of quantum electrons trapped in a two-dimensional device, typically a MOS field-effect transistor. The problem is modeled by the Schrödinger equation in a bounded domain coupled to the Poisson equation for the electrical potential. The controller acts on the system through the boundary condition on the potential, on a part of the boundary modeling the gate. We prove that, generically with respect to the shape of the domain and boundary conditions on the gate, the device is controllable. We also consider control properties of a more realistic nonlinear version of the device, taking into account the self-consistent electrostatic Poisson potential.

5.26. The Interaction Picture method for solving the generalized nonlinear Schrödinger equation in optics

The ”interaction picture” (IP) method is a very promising alternative to Split-Step methods for solving certain type of partial differential equations such as the nonlinear Schrödinger equation involved in the simulation of wave propagation in optical fibers. The method exhibits interesting convergence properties and is likely to provide more accurate numerical results than cost comparable Split-Step methods such as the Symmetric Split-Step method. In [42] we investigate in detail the numerical properties of the IP method and carry out a precise comparison between the IP method and the Symmetric Split-Step method.

5.27. Solving highly-oscillatory NLS with SAM: numerical efficiency and geometric properties

In [46], we present the Stroboscopic Averaging Method (SAM), recently introduced in [7,8,10,12], which aims at numerically solving highly-oscillatory differential equations. More specifically, we first apply SAM to the Schrödinger equation on the 1-dimensional torus and on the real line with harmonic potential, with the aim of assessing its efficiency: as compared to the well-established standard splitting schemes, the stiffer the problem is, the larger the speed-up grows (up to a factor 100 in our tests). The geometric properties of SAM are also explored: on very long time intervals, symmetric implementations of the method show a very good preservation of the mass invariant and of the energy. In a second series of experiments on 2-dimensional equations, we demonstrate the ability of SAM to capture qualitatively the long-time evolution of the solution (without spurring high oscillations).

5.28. Analysis of models for quantum transport of electrons in graphene layers

In [51], we present and analyze two mathematical models for the self consistent quantum transport of electrons in a graphene layer. We treat two situations. First, when the particles can move in all the plane R^2, the model takes the form of a system of massless Dirac equations coupled together by a selfconsistent potential, which is the trace in the plane of the graphene of the 3D Poisson potential associated to surface densities. In this case, we prove local in time existence and uniqueness of a solution in $H^s(R^2)$, for $s > 3/8$ which includes in particular the energy space $H^{1/2}(R^2)$. The main tools that enable to reach $s \in (3/8, 1/2)$ are the dispersive Strichartz estimates that we generalized here for mixed quantum states. Second, we consider a situation where the particles are constrained in a regular bounded domain Ω. In order to take into account Dirichlet boundary conditions which are not compatible with the Dirac Hamiltonian H_0, we propose a different model built on a modified Hamiltonian displaying the same energy band diagram as H_0 near the Dirac points. The well-posedness of the system in this case is proved in H^s_A, the domain of the fractional order Dirichlet Laplacian operator, for $1/2 \leq s$.

5.29. Analysis of a large number of Markov chains competing for transitions

In [18], we consider the behavior of a stochastic system composed of several identically distributed, but non independent, discrete-time absorbing Markov chains competing at each instant for a transition. The competition consists in determining at each instant, using a given probability distribution, the only Markov chain allowed to make a transition. We analyze the first time at which one of the Markov chains reaches its absorbing state. When the number of Markov chains goes to infinity, we analyze the asymptotic behavior of the system for an arbitrary probability mass function governing the competition. We give conditions for the existence of the asymptotic distribution and we show how these results apply to cluster-based distributed systems when the competition between the Markov chains is handled by using a geometric distribution.

5.30. Markov Chains Competing for Transitions: Application to Large-Scale Distributed Systems

In [17], we consider the behavior of a stochastic system composed of several identically distributed, but non independent, discrete-time absorbing Markov chains competing at each instant for a transition. The competition consists in determining at each instant, using a given probability distribution, the only Markov chain allowed to make a transition. We analyze the first time at which one of the Markov chains reaches its absorbing state. We obtain its distribution and its expectation and we propose an algorithm to compute these quantities. We also exhibit the asymptotic behavior of the system when the number of Markov chains goes to infinity. Actually, this problem comes from the analysis of large-scale distributed systems and we show how our results apply to this domain.

5.31. Existence of densities for the 3D Navier–Stokes equations driven by Gaussian noise

In [30], we prove three results on the existence of densities for the laws of finite dimensional functionals of the solutions of the stochastic Navier-Stokes equations in dimension 3. In particular, under very mild assumptions on the noise, we prove that finite dimensional projections of the solutions have densi- ties with respect to the Lebesgue measure which have some smoothness when measured in a Besov space. This is proved thanks to a new argument inspired by an idea introduced by N. Fournier and J. Printems.

5.32. Invariant measure of scalar first-order conservation laws with stochastic forcing

In [50], we assume an hypothesis of non-degeneracy of the flux and study the long-time behaviour of periodic scalar first-order conservation laws with stochastic forcing in any space dimension. For sub-cubic fluxes, we show the existence of an invariant measure. Moreover for sub-quadratic fluxes we show uniqueness and ergodicity of the invariant measure. Also, since this invariant measure is supported by L^p for some p small, we are led to generalize to the stochastic case the theory of L^1 solutions developed by Chen and Perthame.

5.33. Degenerate Parabolic Stochastic Partial Differential Equations: Quasilinear case

In [49], we study the Cauchy problem for a quasilinear degenerate parabolic stochastic partial differential equation driven by a cylindrical Wiener process. In particular, we adapt the notion of kinetic formulation and kinetic solution and develop a well-posedness theory that includes also an L^1-contraction property. In comparison to the previous works of the authors concerning stochastic hyperbolic conservation laws and semilinear degenerate parabolic SPDEs, the present result contains two new ingredients that provide simpler and more effective method of the proof: a generalized Itô formula that permits a rigorous derivation of the kinetic formulation even in the case of weak solutions of certain nondegenerate approximations and a direct proof of strong convergence of these approximations to the desired kinetic solution of the degenerate problem.

5.34. Existence of densities for stable-like driven SDE's with Hölder continuous coefficients

In [29], we consider a multidimensional stochastic differential equation driven by a stable-like Lévy process. We prove that the law of the solution immediately has a density in some Besov space, under some non-degeneracy condition on the driving Lévy process and some very light Hölder-continuity assumptions on the drift and diffusion coefficients.

5.35. Ergodicity results for the stochastic Navier-Stokes equations: an introduction

In the chapter [36], we give an overview of the results on ergodicity for the stochastic Navier-Stokes equations. We first explain the basis on SPDEs and on the concept of invariant measures and ergodicity. Then, in the 2D case, we introduce progressively the various methods, finishing with a celebrated result due to M. Hairer and J. Mattingly on ergodicity with very degenerated noises. In the 3D case, the theory is much less complete. Nonetheles, we show that it is possible to construct Markov evolutions and, under some non degenary assumptions on the noise, to obtain ergodicity.

5.36. Weak truncation error estimates for elliptic PDEs with lognormal coefficients

In [22], we are interested in the weak error committed on the solution of an elliptic partial differential equation with a lognormal coefficient, resulting from the approximation of the lognormal coefficient through a Karhunen-Loéve expansion. We improve results of a previous work, in which L^p-estimates of the weak error are provided. Only small enough values of p (the corresponding values of p depend on the space dimension) could be considered and such bounds are not sufficient to be applied to practical cases. Moreover, the optimality of this weak order (which turns out to be twice the strong order) has not been studied numerically. Therefore, the aim of this paper is double. First we improve drastically the weak error estimate by providing a bound of the C^1-norm of the weak error. This requires regularity results in Hölder spaces, with explicit bounds for the constants. We also consider much more general test functions in the definition of the weak error. Finally, we show the optimality of the weak order and illustrate this weak convergence with numerical results.

5.37. Optimized high-order splitting methods for some classes of parabolic equations

In [21], we are concerned with the numerical solution obtained by splitting methods of certain parabolic partial differential equations. Splitting schemes of order higher than two with real coefficients necessarily involve negative coefficients. It has been demonstrated that this second-order barrier can be overcome by using splitting methods with complex-valued coefficients (with positive real parts). In this way, methods of orders 3 to 14 by using the Suzuki-Yoshida triple (and quadruple) jump composition procedure have been explicitly built. Here we reconsider this technique and show that it is inherently bounded to order 14 and clearly sub-optimal with respect to error constants. As an alternative, we solve directly the algebraic equations arising from the order conditions and construct methods of orders 6 and 8 that are the most accurate ones available at present time, even when low accuracies are desired. We also show that, in the general case, 14 is not an order barrier for splitting methods with complex coefficients with positive real part by building explicitly a method of order 16 as a composition of methods of order 8.

5.38. Higher-Order Averaging, Formal Series and Numerical Integration III: Error Bounds

In earlier papers, it has been shown how formal series like those used nowadays to investigate the properties of numerical integrators may be used to construct high-order averaged systems or formal first integrals of Hamiltonian problems. With the new approach the averaged system (or the formal first integral) may be written down immediately in terms of (i) suitable basis functions and (ii) scalar coefficients that are computed via simple recursions. In [23], we show how the coefficients/basis functions approach may be used advantageously to derive exponentially small error bounds for averaged systems and approximate first integrals.

KERDATA Project-Team

6. New Results

6.1. A-Brain and TomusBlobs

6.1.1. Experiments with TomusBlobs at large scale

Participants: Radu Tudoran, Alexandru Costan, Gabriel Antoniu.

Joint genetic and neuro-imaging data analysis may help identifying risk factors in target populations. Performing such studies on a large number of subjects is challenging as genotyping DNA chips can record several hundred thousands values per subject, while the fMRI images may contain 100k–1M volumetric picture elements. Determining statistically significant links between the two sets of data entails a massive amount of computation as one needs not only to compare all pair-wise relations but also to correct for family-wise multiple comparisons. These false positives are controlled by generating permutations of the input data set. The A-Brain initiative is such a data analysis application involving large cohorts of subjects and used to study and understand the variability that exists between individuals. Supposing that such an application could be executed on a single machine, the computation would take years. Cloud infrastructures have the potential to decrease this computation time to days, by parallellizing and scaling out the application. In order to execute this computation in parallel at a large scale, we noticed that the A-Brain application can be easily described as a MapReduce process. The problem was further divided into 28,000 computation tasks, which were executed as map jobs.

The experiment timespan was 14 days and was performed across 4 cloud deployments in 2 different US Azure data centers — North and West. The processing time for a map job is approximatively 2 hours and there are no notable time differences between the map execution time with respect to the geographical location. This is achieved due to the load balancing of the workload, the data locality within the deployments and to the geographical partition. The global result was aggregated using a MapIterativeReduce technique which was composed of 563 reduce jobs. This reduction process eliminates the implicit barrier between mappers and reducers, the reduction process happens in parallel with the map computation. During the period of the experiment the Azure services became temporary inaccessible, due to a failure of a secured certificate. Despite this problem, the framework was capable to handle the failure due to a safety mechanism that we implemented which suspended the computation until all Azure services became available again. Regarding the lost map/reduce enqueued jobs, the monitor mechanism, which supervises the computation progress, was able to restore them. The cost of the experiment was approximatively 210,000 compute hours, where 1 compute hour is equivalent to 1 CPU running for one hour. The monetary cost of the experiment adds up to almost 20,000 $. The total amount combines the cost of the compute resources, for which a value of 0.09 $/h was considered, the persistent Azure storage cost and the outbound traffic from the data centers. As a result of this experiment, we have confirmed that brain activation signals are a heritable feature.

6.1.2. Using dedicated compute nodes for data management on public clouds

Participants: Radu Tudoran, Alexandru Costan, Gabriel Antoniu.

A large spectrum of scientific applications, some generating data volumes exceeding petabytes, are currently being ported on clouds to build on their inherent elasticity and scalability. One of the critical needs in order to deal with this "data deluge" is an efficient, scalable and reliable storage. However, the storage services proposed by cloud providers suffer from high latencies, trading performance for availability. One alternative is to federate the local virtual disks on the compute nodes into a globally shared storage used for large intermediate or checkpoint data. This collocated storage supports a high throughput but it can be very intrusive and subject to failures that can stop the host node and degrade the application performance.

To deal with these limitations we proposed DataSteward [25], a data management system that provides a higher degree of reliability while remaining non-intrusive through the use of dedicated compute nodes. DataSteward harnesses the storage space of a set of dedicated VMs, selected using a topology-aware clustering algorithm, and has a lifetime dependent on the deployment lifetime. To capitalize on this separation, we introduced a set of scientific data processing services on top of the storage layer, that can overlap with the executing applications. We performed extensive experimentations on hundreds of cores in the Azure cloud: compared to state-of-the-art node selection algorithms, we show up to a 20 % higher throughput, which improves the overall performance of a real life scientific application by up to 45 %.

6.1.3. File transfers for workflows

Participants: Radu Tudoran, Alexandru Costan, Gabriel Antoniu.

Scientific workflows typically communicate data between tasks using files. Currently, on public clouds, this is achieved by using the cloud storage services, which are unable to exploit the workflow semantics and are subject to low throughput and high latencies. To overcome these limitations, we propose in [26] an alternative leveraging data locality through direct file transfers between the compute nodes. We rely on the observation that workflows generate a set of common data access patterns that our solution exploits in conjunction with context information to self-adapt, choose the most adequate transfer protocol and expose the data layout within the virtual machines to the workflow engines. This file management system was integrated within the Microsoft Generic Worker workflow engine and was validated using synthetic benchmarks and a real-life application on the Azure cloud. The results show it can bring significant performance gains: up to 5x file transfer speedup compared to solutions based on standard cloud storage and over 25 % application timespan reduction compared to Hadoop on Azure. This work was done in colaboration with Goetz Brasche and Ramin Rezai Rad from *Microsoft Advance Technology Lab Europe*.

6.2. Optimizing MapReduce Processing

6.2.1. Optimizing MapReduce in virtualized environments

Participant: Shadi Ibrahim.

As data-intensive applications become popular in the cloud, their performance on the virtualized platform calls for empirical evaluations and technical innovations. Virtualization has become a prominent tool in data centers and is extensively leveraged in cloud environments: it enables multiple virtual machines (VMs) — with multiple operating systems and applications — to run within a physical server. However, virtualization introduces the challenging issue of providing effective QoS to VMs and preserving the high disk utilization (i.e., reducing the seek delay and rotation overhead) when allocating disk resources to VMs. We addressed these challenges by developing two Disk I/O scheduling frameworks: *Flubber* and *Pregather*.

In [17], we developed a two-level scheduling framework that decouples throughput and latency allocation to provide QoS guarantees to VMs while maintaining high disk utilization. The high-level throughput control regulates the pending requests from the VMs with an adaptive credit-rate controller, in order to meet the throughput requirements of different VMs and ensure performance isolation. Meanwhile, the low-level latency control, by the virtue of the batch and delay earliest deadline first mechanism (BD-EDF), re-orders all pending requests from VMs based on their deadlines, and batches them to disk devices taking into account the locality of accesses across VMs.

In [24], we developed a novel disk I/O scheduling framework, named *Pregather*, to improve disk I/O efficiency through exposure and exploitation of the special spatial locality in the virtualized environment (regional and sub-regional spatial locality corresponds to the virtual disk space and applications' access patterns, respectively), thereby improving the performance of disk-intensive applications (e.g., MapReduce applications) without harming the transparency feature of virtualization (without a priori knowledge of the applications' access patterns). The key idea behind Pregather is to implement an intelligent model to predict the access regularity of sub-regional spatial locality for each VM.

We evaluated *Pregather* through extensive experiments that involve multiple simultaneous applications of both synthetic benchmarks and a MapReduce application (i.e., distributed sort) on Xen-based platforms. Our experiments indicate that *Pregather* results in high disk spatial locality, yields a significant improvement in disk throughput, and ends up with improved Hadoop performance. This work was done in collaboration with Hai Jin, Song Wu and Xiao Ling from Huazhong University of Science and Technology (HUST).

6.2.2. Investigating energy efficiency in MapReduce

Participants: Shadi Ibrahim, Houssem-Eddine Chihoub, Gabriel Antoniu, Luc Bougé.

A MapReduce system spans over a multitude of computing nodes that are frequency- and voltage-scalable. Furthermore, many MapReduce applications show significant variation in CPU load during their running time. Thus, there is a significant potential for energy saving by scaling down the CPU frequency. Some power-aware data-layout techniques have been proposed to save power, at the cost of a weaker performance. MapReduce applications range from CPU-Intensive to I/O-Intensive. More importantly, a typical MapReduce application comprises many subtasks, each subtask's workload being either a computation, a disk request or a bandwidth request. As a result, there is a high potential for power reduction by scaling down the CPU when the peak CPU performance is not used.

In this ongoing work, a series of experiments are conducted to explore the implications of *Dynamic Voltage Frequency scaling* (DVFS) settings on power consumption in Hadoop-clusters, by benefitting from the current maturity in DVFS research and of the introduction of governors (e.g., *performance*, *powersave*, *ondemand*, *conservative* and *userspace*). By adapting existing DVFS governors in Hadoop clusters, we observe significant variation in performance and power consumption of the cluster with different applications when applying these governors: the different DVFS settings are only sub-optimal for different MapReduce applications. Furthermore, our results reveal that current CPU governors do not exactly reflect their design goal and may even become ineffective to manage the power consumption. Based on this analysis, we are investigating a new approach to reduce the energy consumption in Hadoop through adaptively tuning the governors and/or the CPU frequencies during the execution of MapReduce jobs.

6.2.3. Hybrid infrastructures

Participants: Alexandru Costan, Ana-Ruxandra Ion, Gabriel Antoniu.

As Map-Reduce emerges as a leading programming paradigm for data-intensive computing, today's frameworks which support it still have substantial shortcomings that limit its potential scalability. At the core of Map-Reduce frameworks lies a key component with a huge impact on their performance: the storage layer. To enable scalable parallel data processing, this layer must meet a series of specific requirements. An important challenge regards the target execution infrastructures. While the Map-Reduce programming model has become very visible in the cloud computing area, it is also subject to active research efforts on other kinds of large-scale infrastructures, such as desktop grids. We claim that it is worth investigating how such efforts (currently done in parallel) could converge, in a context where large-scale distributed platforms become more and more connected together.

We investigated several directions where there is room for such progress: they concern storage efficiency under massive-data access concurrency, scheduling, volatility and fault-tolerance. We placed our discussion in the perspective of the current evolution towards an increasing integration of large-scale distributed platforms (clouds, cloud federations, enterprise desktop grids, etc.). We proposed an approach which aims to overcome the current limitations of existing Map-Reduce frameworks, in order to achieve scalable, concurrency-optimized, fault-tolerant Map-Reduce data processing on hybrid infrastructures. We are designing and implementing our approach through an original architecture for scalable data processing: it combines two approaches, BlobSeer and BitDew, which have shown their benefits separately (on clouds and desktop grids respectively) into a unified system. The global goal is to improve the behavior of Map-Reduce-based applications on the target large-scale infrastructures. The internship of Ana-Ruxandra Ion was dedicated to this topic and showed that for reliable hybrid Map-Reduce processing, one needs to first rely on public/private cloud resources, and then to scale them up using the local, yet volatile, desktop grid resources.

6.2.4. Key partitioning techniques

Participants: Shadi Ibrahim, Gabriel Antoniu.

Data locality is a key feature in MapReduce that is extensively leveraged in data-intensive cloud systems: it avoids network saturation when processing large amounts of data by co-allocating computation and data storage, particularly for the map phase. However, our studies with Hadoop, a widely used MapReduce implementation, demonstrate that the presence of partitioning skew (partitioning skew refers to the case when a variation in either the intermediate keys' frequencies or their distributions or both among different data nodes) causes a huge amount of data transfer during the shuffle phase and leads to significant unfairness on the reduce input among different data nodes. As a result, the applications suffer from severe performance degradation due to the long data transfer during the shuffle phase along with the computation skew, particularly in reduce phase. We addressed these problems by developing a new key/value partitioning called *LEEN*.

In [16], we develop a novel algorithm named *LEEN* for locality-aware and fairness-aware key partitioning in MapReduce. *LEEN* aims at saving the network bandwidth dissipation during the shuffle phase of the MapReduce job along with balancing the reducers' inputs. *LEEN* is conducive to improve the data locality of the MapReduce execution efficiency by the virtue of the asynchronous map and reduce scheme, thereby having more control on the keys distribution in each data node. *LEEN* keeps track of the frequencies of buffered keys hosted by each data node. In doing so, *LEEN* efficiently moves buffered intermediate keys to the destination considering the location of the high frequencies along with fair distribution of reducers' inputs.

To quantify the locality, data distribution and performance of *LEEN*, we conducted a comprehensive performance evaluation study using *LEEN* in Hadoop. Our experimental results demonstrate that *LEEN* interestingly can efficiently achieve higher locality, and balance data distribution after the shuffle phase. This work was done in collaboration with Hai Jin, Song Wu and Lu Lu from Huazhong University of Science and Technology (HUST) and Bingsheng He from Nanyang Technological University (NTU).

6.3. Cloud Storage Trade-Offs: Consistency and Self-Adaptiveness

6.3.1. Cost-aware consistency management in the cloud

Participants: Houssem-Eddine Chihoub, Shadi Ibrahim, Gabriel Antoniu.

With the emergence of cloud computing, many organizations have moved their data to the cloud in order to provide scalable, reliable and highly available services. To meet ever growing user needs, these services mainly rely on geographically-distributed data replication to guarantee good performance and high availability. However, with replication, consistency comes into question. Service providers in the cloud have the freedom to select the level of consistency according to the access patterns exhibited by the applications. Most optimizations efforts then concentrate on how to provide adequate trade-offs between consistency guarantees and performance. However, as the monetary cost completely relies on the service providers, in [20] we argue that monetary cost should be taken into consideration when evaluating or selecting a consistency level in the cloud. Accordingly, we define a new metric called *consistency-cost efficiency*. Based on this metric, we present a simple, yet efficient economical consistency model, called *Bismar*, that adaptively tunes the consistency level at run-time in order to reduce the monetary cost while simultaneously maintaining a low fraction of stale reads. Experimental evaluations with the Cassandra cloud storage on a Grid'5000 testbed show the validity of the metric and demonstrate the effectiveness of the proposed consistency model.

6.3.2. Analysis of the impact of consistency mangement on energy consumption

Participants: Houssem-Eddine Chihoub, Shadi Ibrahim, Gabriel Antoniu.

Energy consumption within datacenters has grown exponentially in recent years. In the era of Big Data, storage and data-intensive applications are one of the main causes of the high power usage. However, few studies have been dedicated to the analysis of the energy consumption of storage systems. Moreover, the impact of consistency management has never been investigated in spite of its high importance. In this work, we address this particular issue. We investigate the energy consumption of application workloads with different consistency models. Thereafter, we leverage the observations about power and the resource usage with every consistency level in order to provide insight into energy-saving practices. In this context, we introduce adaptive configurations of the storage cluster according to the used consistency level. Our experimental evaluations on Cassandra deployed on Grid'5000 demonstrate the existence of the inevitable tradeoff between consistency and energy saving. Moreover, they show how reconfiguring the storage cluster can lead to energy saving, enhanced performance, and better consistency.

6.3.3. Chameleon: customized consistency by means of application behavior modeling

Participants: Houssem-Eddine Chihoub, Gabriel Antoniu.

Multiple Big Data applications are being deployed worldwide to serve a very large number of clients nowadays. These applications vary in their performance and consistency requirements. Understanding such requirements at the storage system level is not possible. The high level semantics of an application is not exposed at the system level. In this context, the consequences of a stale read are not the same for all types of applications.

In [28] , we focus on managing consistency at the application level rather than at the system level. In order to achieve this goal, we propose an offline modeling approach of the application access behavior that considers its high–level consistency semantics. Furthermore, every application state is automatically associated with a consistency policy. At runtime, we introduce the *Chameleon* approach that leverages the application model to provide a customized consistency specific to that application. Experimental evaluations show the high accuracy of our modeling approach exceeding 96% of correct classification of the application states. Moreover, our experiments conducted on Grid'5000 show that *Chameleon* adapts, for every time period, according to the application behavior and requirements while providing best-effort performance.

6.4. Scalable I/O and Virtualization for Exascale Systems

6.4.1. Damaris/Viz

Participants: Matthieu Dorier, Gabriel Antoniu, Lokman Rahmani.

In the context of the Joint Inria/UIUC/ANL Laboratory for Petascale computing (JLCP), we are developing Damaris, which enables efficient I/O, data analysis and visualization at very large scale from SMP machines. The I/O bottlenecks already present on current petascale systems as well as the amount of data written by HPC applications force to consider new approaches to get insights from running simulations. Trying to bypass the need for storage or drastically reducing the amount of data generated will be of outmost importance for exascale. In-situ visualization has therefore been proposed to run analysis and visualization tasks closer to the simulation, as it runs.

We investigated the limitations of existing in-situ visualization software and proposed Damaris/Viz, a new version of Damaris that fills the gaps of these software by providing in-situ visualization support to Damaris. The use of Damaris/Viz on top of existing visualization packages allows us to:

- Reduce code instrumentation to a minimum in existing simulations,
- Gather the capabilities of several visualization tools to offer adaptability under a unified data management interface,
- Use dedicated cores to hide the run time impact of in-situ visualization and
- Efficiently use memory through a shared-memory-based communication model.

Experiments were conducted on Blue Waters (Cray XK6 at NCSA), Intrepid (BlueGene/P at ANL) and Grid'5000 with representative visualization scenarios for the CM1 [33] atmospheric simulation and the Nek5000 [35] CFD solver. Part of these experiments were carried by NCSA researcher Roberto Sisneros, who gave us important (and very positive) feedbacks on the usability of Damaris at scale (up to 6400 cores on Blue Waters) with real applications. The results of this work were presented as a poster in the PhD forum of IEEE IPDPS 2013 [22], published in a research report [29] and at the IEEE LDAV 2013 conference [23], and a demo of Damaris/Viz was presented at Inria's exhibition booth at the Supercomputing (SC 2013) conference.

This work enlightened the fact that interactive in-situ visualization, although greatly improved by Damaris/Viz, still lakes interactivity. Several meetings were organized with Tom Peterka (ANL) and Roberto Sisneros (NCSA) during the SC conference and during the 10th workshop of the JLPC. We started working on an approach that leverages information theory metrics to automatically find important features of the simulations' data and to reduce the visualization load accordingly.

6.4.2. CALCioM

Participants: Matthieu Dorier, Gabriel Antoniu.

Unmatched computation and storage performance in new HPC systems have led to a plethora of I/O optimizations ranging from application-side collective I/O to network and disk-level request scheduling on the file system side. As we deal with ever larger machines, the interference produced by multiple applications accessing a shared parallel file system in a concurrent manner becomes a major problem. Interference often breaks single-application I/O optimizations, dramatically degrading application I/O performance and, as a result, lowering machine wide efficiency.

Following discussions initiated in 2012 with ANL's Rob Ross and Dries Kimpe, a three month internship of Matthieu Dorier at Argonne National Lab during the summer 2013 led to the design and evaluation of CALCioM (Cross-Application Layer for Coordinated I/O Management), a framework that aims to mitigate I/O interference through the dynamic selection of appropriate scheduling policies. CALCioM allows several applications running on a supercomputer to communicate and coordinate their I/O strategy in order to avoid interfering with one another. Several I/O strategies were evaluated using this framework. Experiments on Argonne's BG/P Surveyor machine and on several clusters of Grid'5000 showed how CALCioM can be used to efficiently and transparently improve the scheduling strategy between several otherwise interfering applications, given specified metrics of machine wide efficiency.

Future work will explore approaches to automatically detect the temporal I/O patterns of simulations in order to further improve the scheduling decisions made by CALCioM.

6.4.3. Scalable metadata management for WAN

Participants: Rohit Saxena, Alexandru Costan, Gabriel Antoniu.

BlobSeer-WAN is a data management service specifically optimized for geographically distributed environments. It is an extension of BlobSeer, a large scale data management service. The metadata is replicated asynchronously for low latency. There is a version manager on each site and vector clocks are used to enable collision detection and resolution under highly concurrent access. It was developed within the framework of Viet-Trung Tran' s PhD thesis, in relation to the FP3C project.

BlobSeer-WAN is used as a storage backend for HGMDS, a multi master metadata server designed for a global distributed file system, developed at University of Tsukuba. Several experiments have been conducted with this setup on the Grid'5000 testbed which have shown scalable metadata performance under geographically distributed environments.

LAGADIC Project-Team

6. New Results

6.1. Visual tracking

6.1.1. 3D model-based tracking

Participants: Antoine Petit, Eric Marchand.

This study focused on the issue of estimating the complete 3D pose of the camera with respect to a potentially textureless object, through model-based tracking. We proposed to robustly combine complementary geometrical and color edge-based features in the minimization process, and to integrate a multiple-hypotheses framework in the geometrical edge-based registration phase [53], [52], [68], [11].

6.1.2. Pose estimation through multi-planes tracking

Participants: Bertrand Delabarre, Eric Marchand.

This study dealt with dense visual tracking robust towards scene perturbations using 3D information to provide a space-time coherency. The proposed method is based on a piecewise-planar scenes visual tracking algorithm which aims at minimizing an error between an observed image and reference templates by estimating the parameters of a rigid 3D transformation taking into acount the relative positions of the planes in the scene. Both the sum of conditional variance and mutual information have been considered[40] [67].

6.1.3. Pose estimation from spherical moments

Participant: François Chaumette.

This study has been realized in collaboration with Omar Tahri from ISR in Coimbra (Portugal) and Youcef Mezouar from Institut Pascal in Clermont-Ferrand. It was devoted to the classical PnP (Perspective-from-N-Points) problem whose goal is to estimate the pose between a camera and a set of known points from the image measurement of these points. We have developed a new method based on invariant properties of the spherical projection model, allowing us to decouple the pose estimation in two steps: the first one provides the translation by minimizing a criterium using an iterative Newton-like method, the second one directly provides the rotation by solving a Procrustes problem [65], [26].

6.1.4. Structure from motion

Participants: Riccardo Spica, Paolo Robuffo Giordano, François Chaumette.

Structure from motion (SfM) is a classical and well-studied problem in computer and robot vision, and many solutions have been proposed to treat it as a recursive filtering/estimation task. However, the issue of *actively* optimizing the transient response of the SfM estimation error has not received a comparable attention. In the work [64], we studied the problem of designing an online active SfM scheme characterized by an error transient response equivalent to that of a reference linear second-order system with desired poles. Indeed, in a nonlinear context, the observability properties of the states under consideration are not (in general) time-invariant but may depend on the current state and on the current inputs applied to the system. It is then possible to simultaneously act on the estimation gains and system inputs (i.e., the camera velocity for SfM) in order to optimize the observation process and impose a desired transient response to the estimation error. The theory developed in [64] has a general validity and can be applied to many different contexts: in [64] it is shown how to tailor the proposed machinery to two concrete SfM problems involving structure estimation for point features and for planar regions from measured image moments.

6.1.5. 3D reconstruction of transparent objects

Participant: Patrick Rives.

This work has been realized in collaboration with Nicolas Alt, Ph.D. student at the “Technische Universität München” (TUM).

Visual geometry reconstruction of unstructured domestic or industrial scenes is an important problem for applications in virtual reality, 3D video or robotics. With the advent of Kinect sensor, accurate and fast methods for 3D reconstruction have been proposed. However, transparent objects cannot be reconstructed with methods that assume a consistent appearance of the observed 3D structure for different viewpoints. We proposed an algorithm that searches the depth map acquired by a depth camera for inconsistency effects caused by transparent objects. Consistent scene parts are filtered out. The result of our method hence complements existing approaches for 3D reconstruction of Lambertian objects [30].

6.1.6. Pseudo-semantic segmentation

Participants: Rafik Sekkal, Marie Babel.

This study has been realized in collaboration with Ferran Marques from Image Processing Group of the Technical University of Catalonia (Barcelona). We designed a video segmentation framework based on contour projections. This 2D+t technique provides a joint hierarchical and multiresolution solution. Results obtained on state-of-the-art benchmarks have demonstrated the ability of our framework to insure the spatio-temporal consistency of the regions along the sequence.

6.1.7. Augmented reality

Participants: Pierre Martin, Eric Marchand.

Using Simultaneous Localization And Mapping (SLAM) methods becomes more and more common in Augmented Reality (AR). To achieve real-time requirement and to cope with scale factor and the lack of absolute positioning issue, we proposed to decouple the localization and the mapping step. This approach has been validated on an Android Smartphone through a collaboration with Orange Labs [46].

Dealing with AR, we have proposed a method named Depth-Assisted Rectification of Patches (DARP), which exploits depth information available in RGB-D consumer devices to improve keypoint matching of perspectively distorted images [44].

6.2. Visual servoing

6.2.1. Photometric moment-based visual servoing

Participants: Manikandan Bakthavatchalam, François Chaumette.

This goal of this work is to use a set of photometric moments as visual features for visual servoing. We first determined the analytical form of the interaction matrix related to these moments. From the results obtained in the past from binary moments, we then selected a set of four features to control four degrees of freedom (dof) with excellent decoupling and stability properties [35]. More recently, thanks to a collaboration with Omar Tahri from ISR Coimbra in Portugal, these results have been extended to the full six dof case.

6.2.2. Visual servoing of humanoid robot

Participant: François Chaumette.

This study has been realized in collaboration with the Pal robotics company located in Barcelona, Spain. It was devoted to the control of the arm of a humanoid robot by visual servoing for manipulation tasks [29].

6.2.3. Visual servoing of cable-driven parallel robot

Participant: François Chaumette.

This study is realized in collaboration with Rémy Ramadour and Jean-Pierre Merlet from Coprin group at Inria Sophia Antipolis. Its goal is to adapt visual servoing techniques for cable-driven parallel robot in order to achieve acurate manipulation tasks. This study is in the scope of the Inria large-scale initiative action PAL (see Section 8.2.6).

6.2.4. Nanomanipulation

Participants: Le Cui, Eric Marchand.

We began a work, within the ANR P2N Nanorobust project (see Section 8.2.1), on the development of micro- and nano-manipulation within SEM (Scanning Electron Microscope). Our goal is to provide visual servoing techniques for positioning and manipulation tasks with a nanometer precision. This year, we focused on the characterisation of the projection model of a SEM along with the approach required for its calibration.

6.3. Visual navigation of mobile robots

6.3.1. New RGB-D sensor design for indoor 3D mapping

Participants: Eduardo Fernandez Moral, Patrick Rives.

A multi-sensor device has been developed for omnidirectional RGB-D (color+depth) image acquisition (see Figure 5 .a). This device allows acquiring such omnidirectional images at high frame rate (30 Hz). This approach has advantages over other alternatives used nowadays in terms of accuracy and real-time spherical image construction of indoor environments, which are of particular interest for mobile robotics. This device has important prospective applications, such as fast 3D-reconstruction or simultaneous localization and mapping (SLAM). A novel calibration method for such device has been developed. It does not require any specific calibration pattern, taking into account the planar structure of the scene to cope with the fact that there is no overlapping between sensors. A method to perform image registration and visual odometry has also been developed. This method relies in the matching of planar primitives that can be efficiently obtained from the depth images. This technique performs considerably faster than previous registration approaches based on ICP.

6.3.2. Long term mapping

Participants: Tawsif Gokhool, Patrick Rives.

This work inscribes in the context of lifelong navigation and map building. The kind of representation that we focus on is made up of a topometric map consisting of a graph of spherical RGB-D views. Thanks to the use of a saliency map built from the photometric and geometric data, we are able to characterize the conditioning of the pose estimation algorithm and to keep as keyframes only a subset of the spherical RGB-D views acquired on the fly. Subsequently, a study on the spread of keyframes was made. The aim was to investigate ways of covering completely and optimally the explored environment in a pose graph representation. Again, over here, the benefits are twofold. Firstly, data acquisition at a throttle of 30 Hz induces many redundant information in the database, which may not necessarily contribute much to the registration phase. Therefore, intelligent selection of keyframes helped in the reduction of data redundancy. Furthermore, as pointed out in the literature, frame to keyframe alignment has the advantage of reducing trajectory drift since the propagation error is diminished as well (see Figure 5 .b)

6.3.3. Semantic mapping

Participants: Romain Drouilly, Patrick Rives.

Semantic mapping aims at building rich cognitive representations of the world in addition to classical topometric maps. A dense labeling has been achieved from high resolution outdoor images using an approach combining Random Forest (RF) and Conditional Random Field (CRF). A second development dealt with the use of semantic information for localization in indoor scenes. For this kind of scenes dense labeling is more difficult due to the large number of potential classes. Therefore algorithms developed for this task rely on a sparse representation of indoor environments called “pbmap”. It consists of a graph whose nodes are the planes present in a given scene. These planes are the only parts of the scene that are labeled. Very high labeling rates of planes has been reached (more than 90%) and it has been shown that these labeled planes could be useful for localization and navigation tasks.

6.3.4. Automous navigation of wheelchairs

Participants: Rafik Sekkal, François Pasteau, Marie Babel.

../../../../projets/lagadic/IMG/device.png

../../../../projets/lagadic/IMG/3Dmap.png

(a) (b)

Figure 5. a) Omnidirectional RGB-D sensor, b) Top view of dense visual SLAM with fusion of intensity and depth

The goal of this work is to design an autonomous navigation framework of a wheelchair by means of a single camera and visual servoing. We focused on a corridor following task where no prior knowledge of the environment is required. The servoing process matches the non-holonomic constraints of the wheelchair and relies on two visual features, namely the vanishing point location and the orientation of the median line formed by the straight lines related to the bottom of the walls [60]. This overcomes the initialization issue typically raised in the literature. The control scheme has been implemented onto a robotized wheelchair and results show that it can follow a corridor with an accuracy of ± 3 cm [50]. This study is in the scope of the Inria large-scale initiative action PAL (see Section 8.2.6) as well as of the Apash project (see Section 8.1.2).

6.3.5. *Semi-autonomous control of a wheelchair for navigation assistance along corridors*

Participants: Marie Babel, François Pasteau, Alexandre Krupa.

This study concerns a semi-autonomous control approach that we designed for safe wheelchair navigation along corridors. The control relies on the combination of a primary task of wall avoidance performed by a dedicated visual servoing framework and a manual steering task. A smooth transition from manual driving to assisted navigation is obtained thanks to a gradual visual servoing activation method that guarantees the continuity of the control law. Experimental results clearly show the ability of the approach to provide an efficient solution for wall avoiding purposes. This study is in the scope of the Inria large-scale initiative action PAL (see Section 8.2.6) as well as of the Apash project (see Section 8.1.2).

6.3.6. *Target tracking*

Participants: Ivan Markovic, François Chaumette.

This study was realized in the scope of the FP7 Regpot Across project (see Section 8.3.1.2) during the three-month visit of Ivan Markovic, Ph.D. student at the Unviersity of Zagreb. It consisted in developing a pedestrian visual tracking from an omni-directional fish-eye camera and a visual servoing control scheme so that a mobile robot is able to follow the pedestrian. This study has been validated on our Pioneer robot (see Section 5.5).

6.3.7. *Obstacle avoidance*

Participants: Fabien Spindler, François Chaumette.

This study was realized in collaboration with Andrea Cherubini who is now Assistant Prof. at Université de Montpellier. It is concerned with our long term researches about visual navigation from a visual memory without any accurate 3D localization [9]. In order to deal with obstacle avoidance while preserving the visibility in the visual memory, we have proposed a control scheme based on tentacles for fusing the data provided by a pan-tilt camera and a laser range sensor [14]. Recent progresses have been obtained by considering moving obstacles [39].

6.4. Medical robotics

6.4.1. *Needle detection and tracking in 3D ultrasound*

Participants: Pierre Chatelain, Alexandre Krupa.

We developed an algorithm for detecting and tracking a flexible needle in a sequence of 3D ultrasound volumes when it is manually inserted, without any a priori information on the insertion direction. Our approach is based on the combination of a RANSAC algorithm with Kalman filtering in a closed loop fashion and allows real-time tracking of the needle. In addition, a pose-based visual servoing was developed for automatically moving a robotized 3D ultrasound probe in order to keep the needle tip centered in the volume and to align its main axis with the central plane of the volume. This needle detection algorithm and probe automatic guidance were experimentally validated during the insertion of a needle in a gelatin phantom [38].

6.4.2. *Non-rigid target tracking in ultrasound images*

Participants: Marie Babel, Alexandre Krupa.

In order to robustly track the motion of a tumour or cyst during needle insertion, we developed a new approach to track a deformable target within a sequence of 2D ultrasound images. It is based on a dedicated hierarchical grid interpolation algorithm (HGI) that is typically used for real-time video compression purposes. This approach provides a continuous motion representation of the target by using a grid of control points that models both their global displacement and local deformations. The motion of each control point is estimated by a hierarchical and multi-resolution local search method in order to minimize the sum of squared difference of the target pixel intensity between successive images. This new approach was validated from 2D ultrasound images of real human tissues undergoing rigid and non-rigid deformations.

6.4.3. Adaptive arc-based path planning for robot-assisted needle 3D steering using duty-cycling control technique

Participant: Alexandre Krupa.

This study concerned the development of a method for three dimensional steering of a beveled-tip flexible needle that can be used in medical robotics for percutaneous assistance procedures. The proposed solution is the extension of an adaptive arc-based 2D planar approach. It combines the Rapidly-Exploring Random Tree (RRT) algorithm, the duty-cycling needle control technique and stop and turn phases to reorientate the needle in a new working plane each time it is necessary. Simulation results demonstrate the feasibility of this approach to reach a 3D target while avoiding obstacles and its robustness to needle kinematic model errors.

6.4.4. Gait analysis

Participants: Cyril Joly, Patrick Rives.

Clinical evaluation of frailty in the elderly is the first step to decide the degree of assistance they require. Advances in robotics make it possible to turn a standard assistance device into an augmented device that may enrich the existing tests with new sets of daily measured criteria. We designed an augmented 4-wheeled rollator, equipped with a Kinect and odometers, for daily biomechanical gait analysis. It allows to estimate on line legs and feet configurations during the walk. Preliminary results [43] obtained on four healthy persons show that relevant data can be extracted for gait analysis (*e.g.* foot orientation and tibia-foot angle, feet position) during an assisted walk.

This work has been realized in collaboration with Claire Dune from the University of Toulon and in the scope of the Inria large-scale initiative action PAL (see Section 8.2.6).

6.5. Control of single and multiple UAVs

6.5.1. State estimation and flight control of quadrotor UAVs

Participants: Riccardo Spica, Paolo Robuffo Giordano.

Over the last years the robotics community witnessed an increasing interest in the Unmanned Aerial Vehicle (UAV) field. In particular quadrotor UAVs have become more and more widespread in the community as experimental platform for, e.g., testing novel 3D planning, control and estimation schemes in real-world indoor and outdoor conditions. Indeed, in addition to being able to take-off and land vertically, quadrotors can reach high angular accelerations thanks to the relatively long lever arm between opposing motors. This makes them more agile than most standard helicopters or similar rotorcraft UAVs, and thus very suitable to realize complex tasks such as aerial mapping, air pollution monitoring, traffic management, inspection of damaged buildings and dangerous sites, as well as agricultural applications such as pesticide spraying.

Key components for the successful deployment of such systems are (i) a reliable state estimation module able to deal with highly unstructured and/or GPS-denied indoor environments, and (ii) a robust flight control algorithm able to cope with model uncertainties and external disturbances (e.g., adverse atmospheric conditions). The difficulty of these estimation and control problems is also increased by the limited amount of sensing and processing capabilities onboard standard quadrotors: this clearly imposes additional strict requirements on the complexity of the employed algorithms.

In the context of robust flight control of standard quadrotors, the works [31], [32] addressed the theoretical developments and experimental validation of a novel nonlinear adaptive flight controller able to estimate online the UAV dynamic parameters (such as the position of the center of mass when carrying unmodeled payloads), and to compensate for external wind gusts. In parallel, we also developed in [63] a high performance and open-source hardware/software control architecture for flight control of quadrotor UAVs made available to the general public on a open repository. This was achieved by combining state-of-the-art filtering and control techniques with a careful customization and calibration of a commercially available and low-cost quadrotor platform. Finally, still in the context of flight control, the work [58] reported a successful experimental validation of several flight tests for a novel overactuated quadrotor design with tilting propellers behaving as a fully-actuated rigid body in 3D space (thus, able to control its position and orientation in a fully decoupled way).

As for state estimation, the work [41] introduces a novel nonlinear estimation filter meant to obtain a metric measurement of the body-frame linear velocity from optical flow decomposition (thus, visual input) and concurrent fusion of the accelerometer/gyro readings from the onboard IMU. The peculiarity of this filtering technique is the possibility to both explicitly characterize and impose the transient response of the estimation error (thus, the filter performance) by acting on the estimation gains and UAV motion (acceleration). This is in contrast with the consolidated use of EKF schemes which, because of their inherent linearization of the system dynamics, do not typically allow to draw any conclusions about the stability/transient response of the estimation error.

These works were realized in collaboration with the robotics groups at the University of Cassino, Italy, and at the Max Planck Institute for Biological Cybernetics, Tübingen, Germany.

6.5.2. Collective control of multiple UAVs

Participant: Paolo Robuffo Giordano.

The challenge of coordinating the actions of multiple robots is inspired by the idea that proper coordination of many simple robots can lead to the fulfilment of arbitrarily complex tasks in a robust (to single robot failures) and highly flexible way. Teams of multi-robots can take advantage of their number to perform, for example, complex manipulation and assembly tasks, or to obtain rich spatial awareness by suitably distributing themselves in the environment. Within the scope of robotics, autonomous search and rescue, firefighting, exploration and intervention in dangerous or inaccessible areas are the most promising applications.

In the context of multi-robot (and multi-UAV) coordinated control, *connectivity* of the underlying graph is perhaps the most fundamental requirement in order to allow a group of robots accomplishing common goals by means of *decentralized* solutions. In fact, graph connectivity ensures the needed continuity in the data flow among all the robots in the group which, over time, makes it possible to share and distribute the needed information. In this respect, in [23] a fully decentralized strategy for continuous connectivity maintenance for a group of UAVs has been theoretically developed and experimentally validated on a team of 4 quadrotor UAVs. An extension for allowing an external planner (e.g., a human user) to vary online the minimum degree of connectivity of the group was also proposed in [59]. Finally, [48] dealt with the issue of coupling the purely reactive strategy for connectivity maintenance with an autonomous exploration algorithm in a cluttered 3D environment (still experimentally tested on a team of quadrotor UAVs). The complete software architecture developed for performing these and similar multi-UAV experiments was also published in [42].

These works were realized in collaboration with the robotics group at the Max Planck Institute for Biological Cybernetics, Tübingen, Germany.

LEAR Project-Team

6. New Results

6.1. Visual recognition in images

6.1.1. Label-Embedding for Attribute-Based Classification

Participants: Zeynep Akata, Florent Perronnin, Zaid Harchaoui, Cordelia Schmid.

Attributes are an intermediate representation, which enables parameter sharing between classes, a must when training data is scarce. We propose in [13] to view attribute-based image classification as a label-embedding problem: each class is embedded in the space of attribute vectors. We introduce a function which measures the compatibility between an image and a label embedding, as shown in Figure 1 . The parameters of this function are learned on a training set of labeled samples to ensure that, given an image, the correct classes rank higher than the incorrect ones. Results on the Animals With Attributes and Caltech-UCSD-Birds datasets show that the proposed framework outperforms the standard Direct Attribute Prediction baseline in a zero-shot learning scenario. The label embedding framework offers other advantages such as the ability to leverage alternative sources of information in addition to attributes (e.g. class hierarchies) or to transition smoothly from zero-shot learning to learning with large quantities of data.

6.1.2. Good Practice in Large-Scale Learning for Image Classification

Participants: Zeynep Akata, Florent Perronnin, Zaid Harchaoui, Cordelia Schmid.

In this paper [2], we benchmark several SVM objective functions for large-scale image classification. We consider one-vs-rest, multi-class, ranking, and weighted approximate ranking SVMs. A comparison of online and batch methods for optimizing the objectives shows that online methods perform as well as batch methods in terms of classification accuracy, but with a significant gain in training speed. Using stochastic gradient descent, we can scale the training to millions of images and thousands of classes. Our experimental evaluation shows that ranking-based algorithms do not outperform the one-vs-rest strategy when a large number of training examples are used. Furthermore, the gap in accuracy between the different algorithms shrinks as the dimension of the features increases. We also show that learning through cross-validation the optimal rebalancing of positive and negative examples can result in a significant improvement for the one-vs-rest strategy. Finally, early stopping can be used as an effective regularization strategy when training with online algorithms. Following these “good practices”, we were able to improve the state-of-the-art on a large subset of 10K classes and 9M images of ImageNet from 16.7% Top-1 accuracy to 19.1%.

6.1.3. Segmentation Driven Object Detection with Fisher Vectors

Participants: Ramazan Gokberk Cinbis, Jakob Verbeek, Cordelia Schmid.

In [18], we present an object detection system based on the Fisher vector (FV) image representation computed over SIFT and color descriptors. For computational and storage efficiency, we use a recent segmentation-based method to generate class-independent object detection hypotheses, in combination with data compression techniques. Our main contribution is a method to produce tentative object segmentation masks to suppress background clutter in the features. As illustrated in Figure 2 , re-weighting the local image features based on these masks is shown to improve object detection significantly. We also exploit contextual features in the form of a full-image FV descriptor, and an inter-category rescoring mechanism. Our experiments on the VOC 2007 and 2010 datasets show that our detector improves over the current state-of-the-art detection results.

6.1.4. Image Classification with the Fisher Vector: Theory and Practice

Participants: Jorge Sánchez, Florent Perronnin, Thomas Mensink, Jakob Verbeek.

../../../../projets/lear/IMG/akata2.png

Figure 1. Much work in computer vision has been devoted to image embedding (left): how to extract suitable features from an image? We focus on label embedding (right): how to embed class labels in a Euclidean space? We use attributes as side information for the label embedding and measure the "compatibility" between the embedded inputs and outputs with a function F

../../../../projets/lear/IMG/cinbis1.png

Figure 2. The image on the left and the one on the right show the top detection without and with using our segmentation-driven descriptors, respectively.

A standard approach to describe an image for classification and retrieval purposes is to extract a set of local patch descriptors, encode them into a high-dimensional vector and pool them into an image-level signature. The most common patch encoding strategy consists in quantizing the local descriptors into a finite set of prototypical elements. This leads to the popular Bag-of-Visual words (BOV) representation. In [10], we propose to use the Fisher Kernel framework as an alternative patch encoding strategy: we describe patches by their deviation from a "universal" generative Gaussian mixture model. This representation, which we call Fisher Vector (FV) has many advantages: it is efficient to compute, it leads to excellent results even with efficient linear classifiers, and it can be compressed with a minimal loss of accuracy using product quantization. We report experimental results on five standard datasets – PASCAL VOC 2007, Caltech 256, SUN 397, ILSVRC 2010 and ImageNet10K – with up to 9M images and 10K classes, showing that the FV framework is a state-of-the-art patch encoding technique. In figure 3 we show a representative benchmark performance comparison between BOV and FV representations.

Figure 3. Accuracy of the BOV and the FV as a function of the number of Gaussians (left) and feature dimensionality (right) on PASCAL VOC 2007 with SIFT descriptors only.

6.2. Learning and statistical models

6.2.1. Kernel-Based Methods for Hypothesis Testing: A Unified View

Participants: Zaid Harchaoui, Francis Bach, Olivier Cappe, Eric Moulines.

Kernel-based methods provide a rich and elegant framework for developing nonparametric detection procedures for signal processing. Several recently proposed procedures can be simply described using basic concepts of reproducing kernel Hilbert space embeddings of probability distributions, namely mean elements and covariance operators. In [5], we propose a unified view of these tools, and draw relationships with information divergences between distributions (see Figure 4).

6.2.2. Supervised Feature Selection in Graphs with Path Coding Penalties and Network Flows

Participants: Julien Mairal, Bin Yu.

../../../../projets/lear/IMG/zaid1.png

Figure 4. A schematic view of kernel embedding and mean element

In this paper [6], we consider supervised learning problems where the features are embedded in a graph, such as gene expressions in a gene network. In this context, it is of much interest to automatically select a subgraph with few connected components; by exploiting prior knowledge, one can indeed improve the prediction performance or obtain results that are easier to interpret. Regularization or penalty functions for selecting features in graphs have recently been proposed, but they raise new algorithmic challenges. For example, they typically require solving a combinatorially hard selection problem among all connected subgraphs. In this paper, we propose computationally feasible strategies to select a sparse and well-connected subset of features sitting on a directed acyclic graph (DAG), see Figure 5 . We introduce structured sparsity penalties over paths on a DAG called "path coding" penalties. Unlike existing regularization functions that model long-range interactions between features in a graph, path coding penalties are tractable. The penalties and their proximal operators involve path selection problems, which we efficiently solve by leveraging network flow optimization. We experimentally show on synthetic, image, and genomic data that our approach is scalable and leads to more connected subgraphs than other regularization functions for graphs.

../../../../projets/lear/IMG/julien1.png

Figure 5. Network Flow Model with Costs on Arcs for the Path Selection Problem

6.2.3. Structured Penalties for Log-linear Language Models

Participants: Anil Nelakanti, Cédric Archambeau, Julien Mairal, Francis Bach, Guillaume Bouchard.

Language models can be formalized as loglinear regression models where the input features represent previously observed contexts up to a certain length m. The complexity of existing algorithms to learn the parameters by maximum likelihood scale linearly in nd, where n is the length of the training corpus and d is the number of observed features. In this paper [26], we present a model that grows logarithmically in d, making it possible to efficiently leverage longer contexts (see Figure 6). We account for the sequential structure of natural language using tree-structured penalized objectives to avoid overfitting and achieve better generalization.

6.2.4. Optimization with First-Order Surrogate Functions

Participant: Julien Mairal.

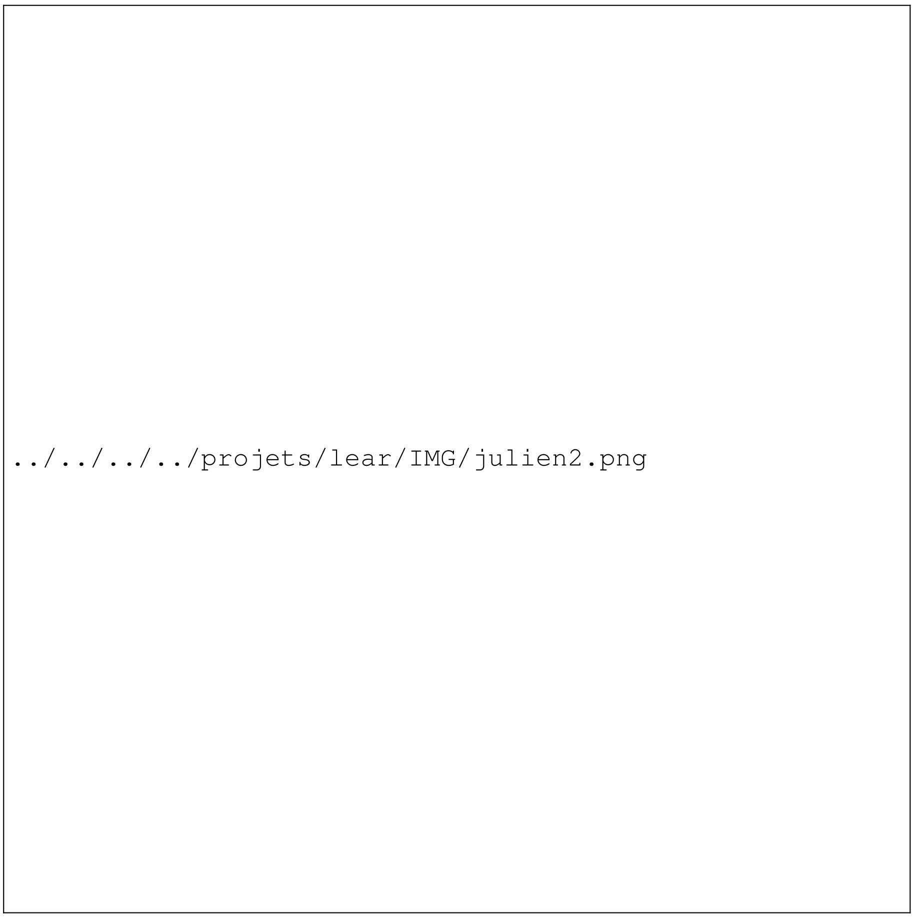


Figure 6. The classical measure of performance for natural language models is the perplexity (lower is better). Our models are denoted by ℓ_2^T and ℓ_{inf}^T.

In this paper [23], we study optimization methods consisting of iteratively minimizing surrogates of an objective function, as illustrated in Figure 7 . By proposing several algorithmic variants and simple convergence analyses, we make two main contributions. First, we provide a unified viewpoint for several first-order optimization techniques such as accelerated proximal gradient, block coordinate descent, or Frank-Wolfe algorithms. Second, we introduce a new incremental scheme that experimentally matches or outperforms state-of-the-art solvers for large-scale optimization problems typically arising in machine learning.

../../../../projets/lear/IMG/julien3.png

Figure 7. Illustration of the basic majorization-minimization principle. We compute a surrogate g_n of the objective function f around a current estimate θ_{n-1}. The new estimate θ_n is a minimizer of g_n. The approximation error h_n is smooth.

6.2.5. Stochastic Majorization-Minimization Algorithms for Large-Scale Optimization

Participant: Julien Mairal.

Majorization-minimization algorithms consist of iteratively minimizing amajorizing surrogate of an objective function. Because of its simplicity and its wide applicability, this principle has been very popular in statistics and in signal processing. In this paper [24], we intend to make this principle scalable. We introduce a stochastic majorization-minimization scheme which is able to deal with largescale or possibly infinite data sets. When applied to convex optimization problems under suitable assumptions, we show that it achieves an expected convergence rate of $O(1/\sqrt{n})$ after n iterations, and of $O(1/n)$ for strongly convex functions. Equally important, our scheme almost surely converges to stationary points for a large class of non-convex problems. We develop several efficient algorithms based on our framework. First, we propose a new stochastic proximal gradient method, which experimentally matches state-of-the-art solvers for large-scale ℓ_1- logistic regression.

Second, we develop an online DC programming algorithm for non-convex sparse estimation. Finally, we demonstrate the effectiveness of our approach for solving large-scale structured matrix factorization problems.

6.3. Recognition in video

6.3.1. Temporal Localization of Actions with Actoms

Participants: Adrien Gaidon, Zaid Harchaoui, Cordelia Schmid.

In this paper [4], we address the problem of localizing actions, such as opening a door, in hours of challenging video data. We propose a model based on a sequence of atomic action units, termed "actoms", that are semantically meaningful and characteristic for the action. Our Actom Sequence Model (ASM) represents an action as a sequence of histograms of actom-anchored visual features, which can be seen as a temporally structured extension of the bag-of-features. Training requires the annotation of actoms for action examples. At test time, actoms are localized automatically based on a non-parametric model of the distribution of actoms, which also acts as a prior on an action's temporal structure. We present experimental results on two recent benchmarks for action localization "Coffee and Cigarettes" and the "DLSBP" dataset. We also adapt our approach to a classification-by-localization set-up, and demonstrate its applicability on the challenging "Hollywood 2" dataset. We show that our ASM method outperforms the current state of the art in temporal action localization, as well as baselines that localize actions with a sliding window method (see Figure 8).

../../../../projets/lear/IMG/zaid3.png

Figure 8. Illustration of actoms-based decomposition of actions.

6.3.2. Activity representation with motion hierarchies

Participants: Adrien Gaidon, Zaid Harchaoui, Cordelia Schmid.

Complex activities, e.g., pole vaulting, are composed of a variable number of sub-events connected by complex spatio-temporal relations, whereas simple actions can be represented as sequences of short temporal parts. In [3], we learn hierarchical representations of activity videos in an unsupervised manner. These hierarchies of mid-level motion components are data-driven decompositions specific to each video. We introduce a spectral divisive clustering algorithm to efficiently extract a hierarchy over a large number of tracklets (i.e., local trajectories). We use this structure to represent a video as an unordered binary tree. We model this tree using nested histograms of local motion features. We provide an efficient positive definite kernel that computes the structural and visual similarity of two hierarchical decompositions by relying on models of their parent-child relations. We present experimental results on four recent challenging benchmarks: the High Five dataset, the Olympics Sports dataset, the Hollywood 2 dataset, and the HMDB dataset. We show that per-video hierarchies provide additional information for activity recognition. Our approach improves over unstructured activity models, baselines using other motion decomposition algorithms, and the state of the art (see Figure 9).

../../../../projets/lear/IMG/zaid2.png

Figure 9. Illustration of motion hierarchies for weight-lifting.

6.3.3. DeepFlow: Large displacement optical flow with deep matching

Participants: Philippe Weinzaepfel, Jerome Revaud, Zaid Harchaoui, Cordelia Schmid.

Optical flow computation is a key component in many computer vision systems designed for tasks such as action detection or activity recognition. However, despite several major advances over the last decade, handling large displacement in optical flow remains an open problem. Inspired by the large displacement optical flow of Brox and Malik, our approach, termed DeepFlow, blends a matching algorithm with a variational approach for optical flow. We propose in [31] a descriptor matching algorithm, tailored to the optical flow problem, that allows to boost performance on fast motions. The matching algorithm builds upon a multi-stage architecture with 6 layers, interleaving convolutions and max-pooling, a construction akin to deep convolutional nets. Figure 10 shows an outline of our approach. Using dense sampling, it allows to efficiently retrieve quasi-dense correspondences, and enjoys a built-in smoothing effect on descriptors matches, a valuable asset for integration into an energy minimization framework for optical flow estimation. DeepFlow efficiently handles large displacements occurring in realistic videos, and shows competitive performance on optical flow benchmarks. Furthermore, it sets a new state-of-the-art on the MPI-Sintel dataset.

../../../../projets/lear/IMG/weinzaepfel1.png

Figure 10. Outline of DeepFlow.

6.3.4. Event retrieval in large video collections with circulant temporal encoding

Participants: Jerome Revaud, Matthijs Douze, Cordelia Schmid, Hervé Jégou.

This paper [28] presents an approach for large-scale event retrieval. Given a video clip of a specific event, e.g., the wedding of Prince William and Kate Middleton, the goal is to retrieve other videos representing the same event from a dataset of over 100k videos. Our approach encodes the frame descriptors of a video to jointly represent their appearance and temporal order. It exploits the properties of circulant matrices to compare the videos in the frequency domain. This offers a significant gain in complexity and accurately localizes the matching parts of videos, see Figure 11 . Furthermore, we extend product quantization to complex vectors in order to compress our descriptors, and to compare them in the compressed domain. Our method outperforms the state of the art both in search quality and query time on two large-scale video benchmarks for copy detection, Trecvid and CCweb. Finally, we introduce a challenging dataset for event retrieval, EVVE, and report the performance on this dataset.

../../../../projets/lear/IMG/revaud1.png

Figure 11. Example of correctly aligned videos. Each row is a different video, and each column corresponds to temporally aligned frames from the videos.

6.3.5. Dense trajectories and motion boundary descriptors for action recognition

Participants: Heng Wang, Alexander Kläser, Cordelia Schmid, Cheng-Lin Liu.

../../../../projets/lear/IMG/Wang1.png

Figure 12. Illustration of the information captured by HOG, HOF, and MBH descriptors. Gradient/flow orientation is indicated by color (hue) and magnitude by saturation. The optical flow (top, middle) shows constant motion in the background, which is due to the camera movements. The motion boundaries (right) encode the relative motion between the person and the background.

This paper [11] introduces a video representation based on dense trajectories and motion boundary descriptors. Trajectories capture the local motion information of the video. A state-of-the-art optical flow algorithm enables a robust and efficient extraction of the dense trajectories. As descriptors we extract features aligned with the trajectories to characterize shape (point coordinates), appearance (histograms of oriented gradients) and motion (histograms of optical flow). Additionally, we introduce a descriptor based on motion boundary histograms (MBH) (see the visualization in Figure 12), which is shown to consistently outperform other state-of-the-art descriptors, in particular on real-world videos that contain a significant amount of camera motion.

We evaluate our video representation in the context of action classification on nine datasets, namely KTH, YouTube,Hollywood2, UCF sports, IXMAS, UIUC, Olympic Sports, UCF50 and HMDB51. On all datasets our approach outperforms current state-of-the-art results.

6.3.6. Action Recognition with Improved Trajectories

Participants: Heng Wang, Cordelia Schmid.

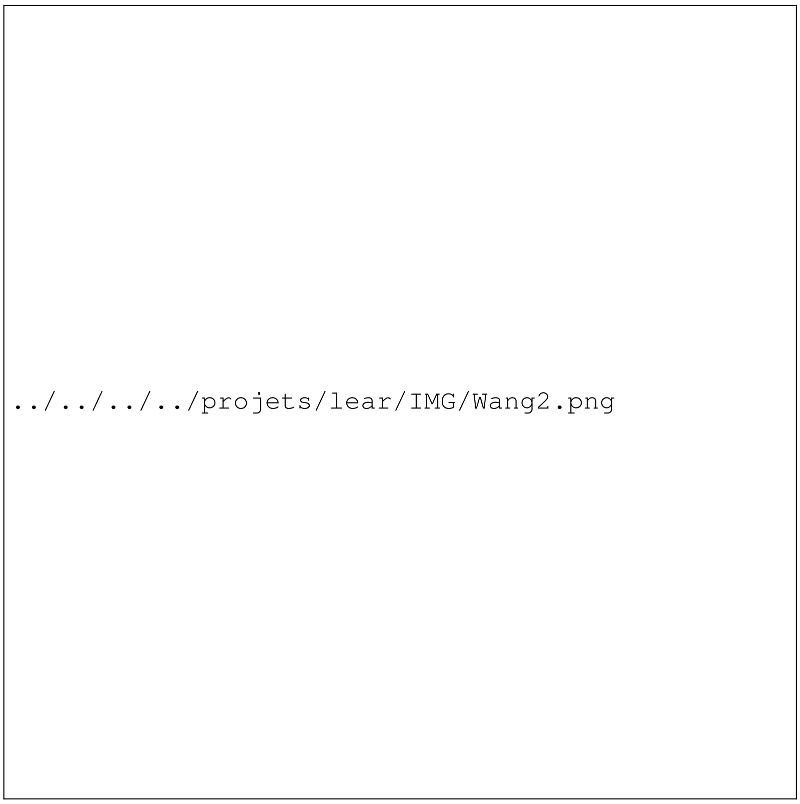


Figure 13. Visualization of human detection and inlier matches (top, left) as well as removed background trajectories, which are due to camera motion (top, right). The bottom row compares the original optical flow (bottom, left) and the warped version (bottom, right).

This paper [30] improves dense trajectories by taking into account camera motion to correct them. To estimate camera motion, we match feature points between frames using SURF descriptors and dense optical flow, which are shown to be complementary. These matches are, then, used to robustly estimate a homography with RANSAC. Human motion is in general different from camera motion and generates inconsistent matches. To improve the estimation, a human detector is employed to remove these matches. Given the estimated camera motion, we remove trajectories consistent with it. We also use this estimation to cancel out camera motion from the optical flow. This significantly improves motion-based descriptors, such as HOF and MBH (see Figure 13). Experimental results on four challenging action datasets (i.e., Hollywood2, HMDB51, Olympic Sports and UCF50) significantly outperform the current state of the art.

6.3.7. Action and event recognition with Fisher vectors on a compact feature set

Participants: Dan Oneaţă, Jakob Verbeek, Cordelia Schmid.

Action recognition in uncontrolled video is an important and challenging computer vision problem. Recent progress in this area is due to new local features and models that capture spatio-temporal structure between local features, or human-object interactions. Instead of working towards more complex models, we focus in this paper [27] on the low-level features and their encoding. We evaluate the use of Fisher vectors as an alternative to bag-of-word histograms to aggregate a small set of state-of-the-art low-level descriptors, in combination with linear classifiers. We present a large and varied set of evaluations, considering (i) classification of short actions in five datasets, (ii) localization of such actions in feature-length movies, and (iii) large-scale recognition of complex events. We find that for basic action recognition and localization MBH features alone are enough for state-of-the-art performance. For complex events we find that SIFT and MFCC features provide complementary cues. On all three problems we obtain state-of-the-art results, while using fewer features and less complex models.

6.3.8. Stable hyper-pooling and query expansion for event detection

Participants: Matthijs Douze, Jerome Revaud, Cordelia Schmid, Hervé Jégou.

This work [19] makes two complementary contributions to event retrieval in large collections of videos. First, we compare different ways of quantizing video frame descriptors in terms of temporal stability. Our best choices compare favorably with the standard pooling technique based on k-means quantization, see Figure 14 . Second, we introduce a technique to improve the ranking. It can be interpreted either as a query expansion method or as a similarity adaptation based on the local context of the query video descriptor. Experiments on public benchmarks show that our methods are complementary and improve event retrieval results, without sacrificing efficiency.

6.3.9. Finding Actors and Actions in Movies.

Participants: Piotr Bojanowski, Francis Bach, Ivan Laptev, Jean Ponce, Cordelia Schmid, Josef Sivic.

This work [16] addresses the problem of learning a joint model of actors and actions in movies using weak supervision provided by scripts. Specifically, we extract actor/action pairs from the script and use them as constraints in a discriminative clustering framework. The corresponding optimization problem is formulated as a quadratic program under linear constraints. People in video are represented by automatically extracted and tracked faces together with corresponding motion features. First, we apply the proposed framework to the task of learning names of characters in movies and demonstrate significant improvements over previous methods used for this task. Second, we explore joint actor/action constraints and show their advantage for weakly supervised action learning. We validate our method in the challenging setting of localizing and recognizing characters and their actions in the feature length movies Casablanca and American Beauty. Figure 15 shows an example of our results.

../../../../projets/lear/IMG/douze1.png

Figure 14. Several quantizations of video frame descriptors (left) to a color-coded index in $\{0, ..., 31\}$. *Leftmost column: standard k-means, right: the proposed SSC. Time runs vertically.*

../../../../projets/lear/IMG/bojanowski1.png

Figure 15. Automatic detection and annotation of characters and their actions in the movie Casablanca. The automatically resolved correspondence between video and script is color-coded.

LFANT Project-Team

6. New Results

6.1. Class groups and other invariants of number fields

Participants: Karim Belabas, Jean-Paul Cerri, Pierre Lezowski.

In collaboration with E. Friedman, K. Belabas presented in [22] a new algorithm to compute the residue at $s = 1$ of the Dedekind zeta function of a number field, conditional on GRH. This improves on previous results of Eric Bach [31] by a useful constant factor. Such an estimate is one of the two key analytic ingredients to Buchmann's class group algorithm, the other being the existence (under GRH) of an explicit set of small generators [33].

In collaboration with F. Thorne, H. Cohen worked on Dirichlet series associated to cubic and quartic fields with given resolvent. In [23] they give an explicit formula for the Dirichlet series $\sum_K |\Delta(K)|^{-s}$, where the sum is over isomorphism classes of all cubic fields whose quadratic resolvent field is isomorphic to a fixed quadratic field k. This is a sequel to previous work of Cohen and Morra, where such formulæ are proved in a more general setting, in terms of sums over characters of certain groups related to ray class groups. Here, the analysis is carried further and they prove explicit formulæ for these Dirichlet series over $\mathbb{Q}$. As an application, they compute tables of the number of S_3-sextic fields K with discriminant ranging up to 10^{23}. An accompanying PARI/GP implementation is available.

In [24], they give an explicit formula for the Dirichlet series $\sum_K |\Delta(K)|^{-s}$, where this time the sum is over isomorphism classes of all quartic fields whose cubic resolvent field is isomorphic to a fixed cubic field k. This work is a sequel to an unpublished preprint of Cohen, Diaz y Diaz, and Olivier.

The papers by H. Cohen on Haberland's formula and numerical computation of Petersson scalar products and by A. Angelakis and P. Stevenhagen on imaginary quadratic fields with isomorphic abelian Galois groups, which were presented at the ANTS-X conference, were published in [17], [16].

6.2. Number and function fields

Participants: Athanasios Angelakis, Jean-Marc Couveignes, Karim Belabas.

In collaboration with Reynald Lercier, Jean-Marc Couveignes presents in [12] a randomised algorithm that on input a finite field K with q elements and a positive integer d outputs a degree d irreducible polynomial in $K[x]$. The running time is $d^{1+o(1)} \times (\log q)^{5+o(1)}$ elementary operations. The $o(1)$ in $d^{1+o(1)}$ is a function of d that tends to zero when d tends to infinity. And the $o(1)$ in $(\log q)^{5+o(1)}$ is a function of q that tends to zero when q tends to infinity. In particular, the complexity is quasi-linear in the degree d.

The book of surveys "Explicit methods in number theory. Rational points and Diophantine equations" [19] edited by K. Belabas with contributions from K. Belabas, F. Beukers, P. Gaudry, W. McCallum, B. Poonen, S. Siksek, M. Stoll and M. Watkins presents the state of the art of the use of explicit methods in arithmetic geometry to solve diophantine problems.

6.3. Quaternion algebras

Participants: Jean-Paul Cerri, Pierre Lezowski, Aurel Page.

In a joint work with J. Chaubert ([11]), J.-P. Cerri and P. Lezowski have studied totally definite quaternion fields over number fields which are Euclidean, that is to say that they admit a left or right Euclidean order. In particular, they have established the complete list of totally definite and Euclidean quaternion fields over real quadratic number fields. In this list, all fields are in fact norm-Euclidean. The proofs are both theoretic and algorithmic.

A. Page uploaded a new version of his article [30] on the computation of arithmetic Kleinian groups, incorporating comments from the referee.

6.4. Complex multiplication and modularity

Participants: Jean-Marc Couveignes, Andreas Enge, Nicolas Mascot, Enea Milio, Aurel Page, Damien Robert.

H. Ivey-Law has been implementing efficient algorithms to compute Hilbert class polynomials and modular polynomials for various modular functions, as well as various supplementary algorithms required by, or based on, these two primary components. These algorithms form an important and time-critical part of algorithms used to select elliptic curves for use in cryptographic applications.

The implementation is based on algorithms for these tasks published by A. Sutherland and his collaborators. It includes, more specifically, algorithms to compute Hilbert class polynomials for various different modular functions over $\mathbb{Z}$ or $\mathbb{Z}/M\mathbb{Z}$, modular polynomials for various different modular functions over $\mathbb{Z}$, $\mathbb{Z}/M\mathbb{Z}$, and/or pre-instantiated at a particular point. The supplementary algorithms include functionality for computing equations for isogenies between elliptic curves and equations for their codomains, for manipulating, interrogating and traversing isogeny volcanoes, for computing minimal polycyclic presentations of abstract groups, for testing supersingularity of j-invariants, for accessing optimised equations of the modular curve $X_1(N)$ for $N \leq 50$, for finding elliptic curves with a given trace or a given endomorphism ring, for calculating the endomorphism ring of a given elliptic curve, for computing the action of the torsor $\mathrm{Cl}(\mathcal{O})$ on the set of elliptic curves with endomorphism ring $\mathcal{O}$ and for enumerating the kernel of the map $\mathrm{Cl}(\mathbb{Z} + N\mathcal{O}) \to \mathrm{Cl}(\mathcal{O})$.

These algorithms are implemented in an experimental branch of PARI/GP, and will be integrated in the public version soon.

A. Enge and R. Schertz determine in [13] under which conditions singular values of multiple η-quotients of square-free level, not necessarily prime to 6, yield class invariants, that is, algebraic numbers in ring class fields of imaginary-quadratic number fields. It turns out that the singular values lie in subfields of the ring class fields of index $2^{k'-1}$ when $k' \geq 2$ primes dividing the level are ramified in the imaginary-quadratic field, which leads to faster computations of elliptic curves with prescribed complex multiplication. The result is generalised to singular values of modular functions on $X_0^+(p)$ for p prime and ramified.

The paper of R. Cosset and D. Robert [25] presenting an algorithm for computing isogenies between principally polarised abelian surface has been accepted for publication in Mathematics of Computation. This paper explains, given the theta coordinates of the points of a maximal isotropic kernel of the ℓ-torsion, how to compute the corresponding isogeny. It also gives formulæ for the conversion between theta coordinates and Mumford coordinates.

The paper by K. Lauter and D. Robert on Improved CRT Algorithm for Class Polynomials in Genus 2, which was presented at the ANTS-X conference, was published in [18].

A. Enge and E. Thomé describe in [14] a quasi-linear algorithm for computing Igusa class polynomials of Jacobians of genus 2 curves via complex floating-point approximations of their roots. After providing an explicit treatment of the computations in quartic CM fields and their Galois closures, they pursue an approach due to Dupont for evaluating ϑ-constants in quasi-linear time using Newton iterations on the Borchardt mean. They report on experiments with the implementation CMH and present an example with class number 20016.

N. Mascot's article on computing modular Galois representations [15] has been published in Rendiconti del Circolo Matematico di Palermo. This article describes an algorithm to compute Galois representations attached to a newform, and to deduce the Fourier coefficients of this newform modulo a small prime.

E. Milio has implemented R. Dupont's algorithms [38] in PARI/GP. With them, he has calculated the three modular polynomials in genus 2 and level 2 defined by Streng's version of Igusa modular forms and a modular polynomial of genus 2 and level 3 coming from theta modular forms.

6.5. Elliptic curve cryptology

Participants: Jean-Marc Couveignes, Andreas Enge, Damien Robert.

Couveignes and Lercier study in [26] the problem of parameterisations by radicals of low genus algebraic curves. They prove that for q a prime power that is large enough and prime to 6, a fixed positive proportion of all genus 2 curves over the field with q elements can be parameterised by 3-radicals. This results in the existence of a deterministic encoding into these curves when q is congruent to 2 modulo 3. Deterministic encodings into curves are useful in numerous situations, for instance in discrete logarithm cryptography. The parameterisation found by Couveignes and Lercier is in some sense the first generic one for genus 2 curves.

A software for this method is in preparation.

The survey [21], published in the *Handbook of Finite Fields*, presents the state of the art of the use of elliptic curves in cryptography.

6.6. Pairings

Participants: Andreas Enge, Damien Robert.

In [27], A. Enge gives an elementary and self-contained introduction to pairings on elliptic curves over finite fields. For the first time in the literature, the three different definitions of the Weil pairing are stated correctly and proved to be equivalent using Weil reciprocity. Pairings with shorter loops, such as the ate, ate$_i$, R-ate and optimal pairings, together with their twisted variants, are presented with proofs of their bilinearity and non-degeneracy. Finally, different types of pairings are reviewed in a cryptographic context. The article can be seen as an update chapter to [40].

With D. Lubicz, D. Robert has worked on extending the algorithm to compute Weil and Tate pairings using theta functions from [42] to the ate and optimal ate pairings in [29]. The result includes how to compute the Miller functions with theta functions, but also how to generalise ate and optimal ate pairings to Kummer varieties. In contrast to preceding algorithms using Miller functions which needed a geometric interpretation of the addition law and worked with Jacobians, this new algorithm uses only the algebraic Riemann relations and works on any abelian variety (provided with a theta structure). This algorithm has been implemented using AVISOGENIES.

LINKS Team

5. New Results

5.1. Querying Heterogeneous Linked Data

Participants: Guillaume Bagan, Iovka Boneva, Angela Bonifati, Pierre Bourhis, Radu Ciucanu, Tom Sebastian, Slawomir Staworko, Sophie Tison.

Staworko, Ciucanu and Boneva presented a new class of schemas for unordered XML trees, which are based on unordered regular expressions, also called multiplicity schemas. They show that many static analysis problems become feasible when removing disjunctions there [6].

Ciucanu and Staworko [8] investigated the case of unordered XML, where the relative order among siblings is ignored, and focused on the problem of learning schemas from examples given by the user. They considered disjunctive multiplicity schemas (DMS) and their restrictions, disjunction-free multiplicity schemas (MS). For both DMS and MS, they prove the learnable cases.

Regular path queries in graphs have found much recent interest in the context of SPARQL queries for linked open data in the RDF format. Bagan, Bonifati and Groz (former PhD student of Mostrare, now PostDoc at Tel-Aviv University) have obtained a precise characterization of those regular path queries that can be answered with polynomial data complexity [5] leading to a trichotomy (AC0, NL-complete, or else NP-complete). Thereby, they have solved an open question (raised by W. Martens in PODS'12).

XPath query evaluation over compressed trees has been studied in [12]. They focused on a fragment of XPath, which is the downward, navigational XPath and presented precise bounds on the time complexity of XPath query execution over grammar-compressed trees. In particular, they focused on counting the nodes selected by an XPath expression, extracting and materializing their pre-order numbers and serializing the obtained subtrees.

In [2], Groz, Staworko, Caron, Roos and Tison studied query rewriting with views when the classes used to define queries and views are Regular XPath and MSO. Next, they investigated problems of static analysis of security access specifications (SAS) by introducing the novel class of interval-bounded SAS and they defined three different manners to compare views (i.e. queries), with a security point of view. Finally, they provided a systematic study of the complexity for deciding these three comparisons.

5.2. Managing Dynamic Linked Data

Participants: Angela Bonifati, Denis Debarbieux, Joachim Niehren, Tom Sebastian.

Bonifati, Goodfellow (former PhD student at the University of Strathclyde, UK), Manolescu and Sileo (former PhD student at the University of Basilicata, Italy, directed by Bonifati) studied XML view maintenance in the presence of updates [1]. Their approach relies on algebraic operators for propagating source updates to the target XML view, e.g. in a typical scenario of GAV (global-as-view) schema mappings. Their algebraic approach is set-oriented as opposed to tuple-oriented methods presented in the literature. Moreoever, it leverages structural identifiers and structural join algorithms. As such, it proved to be more efficient than existing methods for updating materialized XML views.

Debarbieux, Gauwin (former PhD student in the team, now Assistant Professor at the University of Bordeaux), Niehren, Sebastian and Zergaoui (CEO at Innovimax) focused on using early nested word automata in order to approximate earliest query answering algorithms for nested word automata in a highly efficient manner [9]. This approximation can be made tight in practice for automata obtained from XPath expressions. An XPath streaming algorithm based on early nested word automata has been implemented in the FXP tool. FXP outperforms most previous tools in efficiency, while covering more queries of the XPathMark benchmark.

5.3. Linking Data Graphs

Participants: Angela Bonifati, Radu Ciucanu, Joachim Niehren, Aurélien Lemay, Grégoire Laurence, Antoine Ndione, Slawomir Staworko.

In [7], Bonifati, Ciucanu and Staworko investigate the problem of inferring arbitrary n-ary join predicates across two relations via user interactions. The relations can be found on the Web, thus they lack integrity constraints. In such a scenario, the user is asked to label as positive or negative a few tuples depending on whether she would like them in the join result or not. Deciding whether the remaing tuples are uninformative, i.e. do not allow to infer the query goal, can be done in polynomial time.

The PhD thesis of Ndione focuses on probabilistic algorithms to decide approximate membership of words in a language by using property testing. In [3], Ndione, Lemay and Niehren presented an algorithm that tests the membership modulo the edit distance. Their algorithm run in polynomial time, as opposed to other property testing algorithms, leveraging the Hamming distance or the edit distance with moves, that are exponential.

In [11], Laurence, Lemay, Niehren, Staworko and Tommasi (project leader of the Magnet team) studied the problem of learning sequential top-down tree-to- word transducers (STWs). They present a Myhill-Nerode characterization of the corresponding class of sequential tree-to-word transformations (STW). Next, they investigate what learning of stws means, identify fundamental obstacles, and propose a learning model with abstain. Finally, they present a polynomial learning algorithm.

In [4], Niehren, Champavère (former PhD student in the team), Gilleron and Lemay addressed the problem of learnability of regular queries in unranked trees. The idea is that tree pruning strategies and the schemas (DTD in the specific case) can guide the learning process and lead to a class of queries that are learnable according to those. The obtained learning algorithm adds pruning heuristics to the traditional learning algorithm based on tree automata and exploiting positive and negative examples.

LOGNET Team

6. New Results

6.1. A Backward-Compatible Protocol for Inter-routing over Heterogeneous Overlay Networks

Participants: Giang Ngo Hoang [contact], Luigi Liquori, Hung Nguyen Chan [VIELINA, Vietnam].

../../../../projets/lognet/IMG/ogp.png

Figure 11. An Overlay Gateway Protocol Topology

Overlay networks are logical networks running on the highest level of the OSI stack: they are applicative networks used by millions of users everyday. In many scenarios, it would be desirable for peers belonging to overlays running different protocols to communicate with each other and exchange certain information. However, due to differences in their respective protocols, this communication is often difficult or even impossible to be achieved efficiently, even if the overlays are sharing common objectives and functionalities. In this paper, we address this problem by presenting a new overlay protocol, called OGP (Overlay Gateway Protocol), allowing different existing networks to route messages between each other in a backward-compatible fashion, by making use of specialized peers joined together into a super-overlay. Experimental results on a large scale Grid5000 infrastructure show that having only a small number of nodes running the OGP protocol is sufficient for achieving efficient routing between heterogeneous overlay networks.

The three scenarios in Figure 11 are shown to illustrate the routing of three lookup queries, in which full OGP peers, lightweight OGP peers and blind peers interact in order to reach across overlays represent requests, while dashed lines represent responses. using the OGP super-overlay. The three smaller ovals represent standard overlays, while the largest oval represents the OGP super-overlay, forwarding messages back and forth between standard overlays. The black squares B; C; G; N and P represent full OGP peers, the black circles A; D and F represent lightweight OGP peers, while the white circles E; H, and M represent blind peers. Solid lines requests, while dashed lines represent responses. The paper is the continuation of the work of HotPost 2011 [7] and Hets-Nets 2012 [8]: it has been also accepted to ACM SAC 2013 [36] and a long version has been accepted to the International Conference ICDCN 2014 [32].

6.2. Interconnection of large scale unstructured P2P networks: modeling and analysis

Participants: Rossano Gaeta [Univ. Turin], Vincenzo Ciancaglini, Riccardo Loti, Luigi Liquori.

Interconnection of multiple P2P networks has recently emerged as a viable solution to increase system reliability and fault-tolerance as well as to increase resource availability. In this paper we consider interconnection of large scale unstructured P2P networks by means of special nodes (called *Synapses*) that are co-located in more than one overlay. Synapses act as *trait d'union* by sending/forwarding a query to all the P2P networks they belong to. Modeling and analysis of the resulting interconnected system is crucial to design efficient and effective search algorithms and to control the cost of interconnection. To this end, we develop a generalized random graph based model that is validated against simulations and it is used to investigate the performance of search algorithms for different interconnection costs and to provide some insight in the characteristics of the interconnection of a large number of P2P networks. To overcome this strong limitation, we develop a generalized random graph based model to represent the topology of one unstructured P2P network, the partition of nodes into Synapses, the probabilistic flooding based search algorithms, and the resource popularity. We validate our model against simulations and prove that its predictions are reliable and accurate. We use the model to investigate the performance and the cost of different search strategies in terms of the probability of successfully locating at least one copy of the resource and the number of queries as well as the interconnection cost. We also gain interesting insights on the dependency between interconnection cost and statistical properties of the distribution of Synapses. Finally, we show that thanks to our model we can analyze the performance of a system composed of a large number of P2P networks.

To the best of our knowledge, this is the first paper on model-based analysis of interconnection of large scale unstructured P2P networks [11] and the full version has been accepted to the conference [30].

6.3. SIEVE: a distributed, accurate, and robust technique to identify malicious nodes in data dissemination on MANET

Participants: Rossano Gaeta [Univ. Turin], Riccardo Loti [contact], Marco Grangetto [Univ Turin].

We consider the following problem: nodes in a MANET must disseminate data chunks using rateless codes but some nodes are assumed to be malicious, i.e., before transmitting a coded packet they may modify its payload. Nodes receiving corrupted coded packets are prevented from correctly decoding the original chunk. We propose SIEVE, a fully distributed technique to identify malicious nodes.

SIEVE is based on special messages called *checks* that nodes periodically transmit. A check contains the list of nodes identifiers that provided coded packets of a chunk as well as a flag to signal if the chunk has been corrupted. SIEVE operates on top of an otherwise reliable architecture and it is based on the construction of a *factor graph* obtained from the collected checks on which an incremental belief propagation algorithm is run to compute the probability of a node being malicious. Analysis is carried out by detailed simulations using ns-3. We show that SIEVE is very accurate and discuss how nodes speed impacts on its accuracy. We also show SIEVE robustness under several attack scenarios and deceiving actions. The paper has been accepted to [12] and a journal version in [26].

6.4. CCN-TV: a data-centric approach to real-time video services

Participants: Luigi Liquori, Vincenzo Ciancaglini [contact], Riccardo Loti, Giuseppe Piro [Politech Bari], Alfredo Grieco [Politech Bari].

Content Centric Networking is a promising data-centric architecture, based on in-network caching, name-driven routing, and receiver-initiated sessions, which can greatly enhance the way Internet resources are currently used, thus making the support for a broader set of users with increasing traffic demands possible. The CCN vision is, currently, attracting the attention of many researchers across the world, because it has all the potential to become ready to the market, to be gradually deployed in the Internet of today, and to facilitate a graceful transition from a host-centric networking rationale to a more effective data-centric working behavior. At the same time, several issues have to be investigated before CCN can be safely deployed at the Internet scale. They include routing, congestion control, caching operations, name-space planning, and application design. With reference to application-related facets, it is worth to notice that the demand for TV services is growing at an exponential rate over the time, thus requiring a very careful analysis of their performance in CCN architectures. To this end, in the present contribution we deploy a CCN-TV system, able to deliver real-time streaming TV services and we evaluate its performance through a simulation campaign based on real topologies. The paper has been accepted to [31] and [28] and a full version has been invited and will appear as book chapter to [33].

6.5. Towards a Trust and Reputation Framework for Social Web Platforms and @-economy

Participants: Thao Nguyen [contact], Bruno Martin [Unice], Luigi Liquori, Karl Hanks.

Trust and reputation systems (TRSs) have recently seen as a vital asset for the safety of online interaction environment. They are present in many practical applications, e.g., e-commerce and social web. A lot of more complicated systems in numerous disciplines also have been studied and proposed in academia. They work as a decision support tool for participants in the system, helping them decide whom to trust and how trustworthy the person is in fulfilling a transaction. They are also an effective mechanism to encourage honesty and cooperation among users, resulting in healthy online markets or communities. The basic idea is to let parties rate each other so that new public knowledge can be created from personal experiences. The greatest challenge in designing a TRS is making it robust against malicious attacks. In this paper, we provide readers an overview on the research topic of TRSs, propose a consistent research agenda in studying and designing a robust TRS, and present an implemented reputation computing engine alongside simulation results, which is our preliminary work to acquire the target of a trust and reputation framework for social web applications.

Information concerning the reputation of individuals has always been spread by word-of-mouth and has been used as an enabler of numerous economic and social activities. Especially now, with the development of technology and, in particular, the Internet, reputation information can be broadcast more easily and faster than ever before. Trust and Reputation Systems (TRSs) have gained the attention of many information and computer scientists since the early 2000s. TRSs have a wide range of applications and are domain specific. The multiple areas where they are applied, include social web platforms, e-commerce, peer-to-peer networks, sensor networks, ad-hoc network routing, and so on. Among these, we are most interested in social web platforms. We observe that trust and reputation is used in many online systems, such as online auction and shopping websites, including eBay, where people buy and sell a broad variety of goods and services, and Amazon, which is a world famous online retailer. Online services with TRSs provide a better safety to their users. A good TRS can also create incentives for good behavior and penalize damaging actions. Markets with the support of TRSs will be healthier, with a variety of prices and quality of service. TRSs are very important for an online community, with respect to the safety of participants, robustness of the network against malicious behavior and for fostering a healthy market.

../../../../projets/lognet/IMG/process.png

Figure 12. Process of designing a robust trust and reputation system

From a functional point of view, a TRS can be split into three components The first component gathers feedback on participants' past behavior from the transactions that they were involved in. This component includes storing feedback from users after each transaction they take part in. The second component computes reputation scores for participants through a Reputation Computing Engine (RCE), based on the gathered information. The third component processes the reputation scores, implementing appropriate reward and punishment policies if needed, and representing reputation scores in a way which gives as much support as possible to users' decision-making. A TRS can be centralized or distributed. In centralized TRSs, there is a central authority responsible for collecting ratings and computing reputation scores for users. Most of the TRSs currently on the Internet are centralized, for example the feedback system on eBay and customer reviews on Amazon. On the other hand, a distributed TRS has no central authority. Each user has to collect ratings and compute reputation scores for other users himself. Almost all proposed TRSs in the literature are distributed.

Some of the main unwanted behaviors of users that might appear in TRSs are: *free riding* (people are usually not willing to give feedback if they are not given an incentive to do so), *untruthful rating* (users give incorrect feedback either because of malicious intent or because of unintended and uncontrolled variables), *colluding* (a group of users coordinate their behavior to inflate each other's reputation scores or bad-mouth other competitors. Colluding motives are only clear in a specific application), *whitewashing* (a user creates a new identity in the system to replace his old one when the reputation of the old one has gone bad), *milking reputation* (at first, a participant behaves correctly to get a high reputation and then turns bad to make a profit from their high reputation score). The milking reputation behavior is more harmful to social network services and e-commerce than to the others.

This research aims to build on these studies and systematize the process of designing a TRS in general as depicted in Fig. 12 . First, we characterize the application system into which we want to integrate a TRS, and find and identify new elements of information which substitute for traditional signs of trust and reputation in the physical world. Second, based on the characteristics of the application, we find suitable working mechanisms and processes for each component of the TRS. This step should answer the following questions: “What kind of information do we need to collect and how?", “How should the reputation scores be computed using the collected information?", and “How should they be represented and processed to lead users to a correct decision?". To answer the first question, which corresponds to the information gathering component, we should take advantage of information technology to collect the vast amounts of necessary data. An RCE should meet these criteria: *accuracy* for long-term performance (distinguishing a newcomer with unknown quality from a low-quality participant who has stayed in the system for a long time), *weighting* towards recent behavior, *smoothness* (adding any single rating should not change the score significantly), and *robustness* against attacks. Third, we study the tentative design obtained after the second step in the presence of selfish behaviors. During the third step, we can repeatedly return to Step 2 whenever appropriate until the system reaches a desired performance. The fourth step will refine the TRS and make it more robust against malicious attacks. If a modification is made, we should return to Step 2 and check all the conditions in steps 2 and 3 before accepting the modification. The paper has been accepted to [22] and an improved software and a full paper are in preparation in 2014.

6.6. A Scalable Communication Architecture for Advanced Metering Infrastructure

Participants: Giang Ngo Hoang [contact], Luigi Liquori, Hung Nguyen Chan [VIELINA, Vietnam].

Advanced Metering Infrastructure (AMI), seen as foundation for overall grid modernization, is an integration of many technologies that provides an intelligent connection between consumers and system operators. One of the biggest challenge that AMI faces is to scalable collect and manage a huge amount of data from a large number of customers. In our paper, we address this challenge by introducing a mixed peer-to-peer (P2P) and client-server communication architecture for AMI in which metering data is aggregated and processed distributedly at multiple levels and in a tree-like manner. Through analysis we show that the architecture is featured with load scalability, resiliency with failure and partly self-organization. The experiments performed in large scale French Grid5000 platform [G5k] shows the communication efficiency in the proposed architecture. A technical report will be submitted to an international conference [37].

6.7. An Open Logical Framework

Participants: Luigi Liquori [contact], Marina Lenisa [Univ. Udine], Furio Honsell [Univ. Udine], Petar Maksimovic, Ivan Scagnetto [Univ. Udine].

The LFP Framework is an extension of the Harper-Honsell-Plotkin's Edinburgh Logical Framework LF with external predicates, hence the name Open Logical Framework. This is accomplished by defining lock type constructors, which are a sort of "diamond"-modality constructors, releasing their argument under the condition that a possibly external predicate is satisfied on an appropriate typed judgement. Lock types are defined using the standard pattern of constructive type theory, i.e. via introduction, elimination, and equality rules. Using LFP, one can factor out the complexity of encoding specific features of logical systems which would otherwise be awkwardly encoded in LF, e.g. side-conditions in the application of rules in Modal Logics, and sub-structural rules, as in non-commutative Linear Logic. The idea of LFP is that these conditions need only to be specified, while their verification can be delegated to an external proof engine, in the style of the Poincaré Principle or Deduction Modulo. Indeed such paradigms can be adequately formalized in LFP. We investigate and characterize the meta-theoretical properties of the calculus underpinning LFP: strong normalization, confluence, and subject reduction. This latter property holds under the assumption that the predicates are well-behaved, i.e. closed under weakening, permutation, substitution, and reduction in the arguments. Moreover, we provide a canonical presentation of LFP, based on a suitable extension of the notion of $\beta\eta$-long normal form, allowing for smooth formulations of adequacy statements.

LFP is parametric over a potentially unlimited set of (well-behaved) predicates P, which are defined on derivable typing judgements of the form $\Gamma \vdash_{\Sigma} N : \sigma$, see Fig 13 .

The syntax of LFP predicates is not specified, with the main idea being that their truth is to be verified via a call to an external validation tool; one can view this externalization as an oracle call. Thus, LFP allows for the invocation of external "modules" which, in principle, can be executed elsewhere, and whose successful verification can be acknowledged in the system via L-reduction. Pragmatically, lock types allow for the factoring out of the complexity of derivations by delegating the {checking, verification, computation} of such predicates to an external proof engine or tool. The proof terms themselves do not contain explicit evidence for external predicates, but just record that a verification {has to be (lock), has been successfully (unlock)} carried out. In this manner, we combine the reliability of formal proof systems based on constructive type theory with the efficiency of other computer tools, in the style of the Poincaré Principle. In this paper, we develop the meta-theory of LFP. Strong normalization and confluence are proven without any additional assumptions on predicates. For subject reduction, we require the predicates to be well-behaved, i.e. closed under weakening, permutation, substitution, and $\beta\mathcal{L}$-reduction in the arguments. LFP is decidable, if the external predicates are decidable. We also provide a canonical presentation of LFP, based on a suitable extension of the notion of $\beta\eta$-long normal form. This allows for simple proofs of adequacy of the encodings. In particular, we encode in LFP the call-by-value λ-calculus and discuss a possible extension which supports the design-by-contract paradigm. We provide smooth encodings of side conditions in the rules of Modal Logics, both in Hilbert and Natural Deduction styles. We also encode sub-structural logics, i.e. non-commutative Linear Logic. We also illustrate how LFP can naturally support program correctness systems and Hoare-like logics. In our encodings, we utilize a library of *external predicates*. As far as expressiveness is concerned, LFP is a stepping stone towards a general theory of shallow vs deep encodings, with our encodings being shallow by definition. Clearly, by Church's thesis, all external decidable predicates in LFP can be encoded, possibly with very deep encodings, in standard LF. It would be interesting to state in a precise categorical setting the relationship between such deep internal encodings and the encodings in LFP. LFP can also be viewed as a neat methodology for separating the logical-deductive contents from, on one hand, the verification of structural and syntactical properties, which are often needlessly cumbersome but ultimately computable, or, on the other hand, from more general means of validation. This work has been published in the ACM workshops [13] and [29] and a long version has been invited and appear in the Journal of Logic and Computation [27].

../../../../projets/lognet/IMG/OLF.png

Figure 13. Some rule of the Open Logical Framework

M3DISIM Team

6. New Results

6.1. Modeling

6.1.1. Collective effect in molecular motors assembly

Participants: Matthieu Caruel, Jean-Marc Allain [Ecole Polytechnique], Lev Truskinovsky [Ecole Polytechnique].

Skeletal muscles consist of active material capable of producing force. At the microscale, force is the result of complex interactions between two types of proteins named actin and myosin that work coherently in very large assemblies ($\sim 10^9$). The passive mechanical response of skeletal muscles at fast time scales is dominated by long range interactions inducing cooperative behavior without breaking the detailed balance. This leads to such unusual "material properties" as negative equilibrium stiffness and different behavior in force and displacement controlled loading conditions. Our fitting of experimental data suggests that "muscle material" is finely tuned to perform close to a critical point which explains large fluctuations observed in muscles close to the stall force. See paper [28].

6.1.2. Dimensional reductions of a cardiac model for effective validation and calibration

Participants: Matthieu Caruel, Alexandre Imperiale, Radomir Chabiniok [King's College London], Philippe Moireau, Dominique Chapelle, Yves Lecarpentier [Institut du Cœur].

Complex 3D beating heart models are now available, but their complexity makes calibration and validation very difficult tasks. We thus propose a systematic approach of deriving simplified reduced-dimensional models, in "0D"— typically, to represent a cardiac cavity, or several coupled cavities and in "1D"—to model elongated structures such as muscle samples or myocytes. We apply this approach with an earlier-proposed 3D cardiac model designed to capture length-dependence effects in contraction, which we here complement by an additional modeling component devised to represent length-dependent relaxation. We then present experimental data produced with rat papillary muscle samples when varying preload and afterload conditions, and we achieve some detailed validations of the 1D model with these data, including for the length-dependence effects that are accurately captured. Finally, when running simulations of the 0D model pre-calibrated with the 1D model parameters, we obtain pressure–volume indicators of the left ventricle in good agreement with some important features of cardiac physiology, including the so-called Frank–Starling mechanism, the End-Systolic Pressure–Volume Relationship, as well as varying elastance properties. This integrated multi-dimensional modeling approach thus sheds new light on the relations between the phenomena observed at different scales and at the local versus organ levels. See papers [13], [22].

6.1.2.1. Surface-based electrophysiology modeling and assessment of physiological simulations in atria

Participants: Dominique Chapelle, Annabelle Collin, Jean-Frédéric Gerbeau [Reo Project-Team], Mélèze Hocini [Institut LIRYC - IHU Bordeaux], Michel Haïssaguerre [Institut LIRYC - IHU Bordeaux].

The objective of this work is to assess a previously-proposed surface-based electrophysiology model with detailed atrial simulations. This model – derived and substantiated by mathematical arguments – is specifically designed to address thin structures such as atria, and to take into account strong anisotropy effects related to fiber directions with possibly rapid variations across the wall thickness. The simulation results are in excellent adequacy with previous studies, and confirm the importance of anisotropy effects and variations thereof, see Figure 1 . Furthermore, this surface-based model provides dramatic computational benefits over 3D models with preserved accuracy. See paper [23].

6.1.3. Strong convergence results in the asymptotic behavior of the 3D-shell model

Participants: Dominique Chapelle, Annabelle Collin.

../../../../projets/m3disim/IMG/atria.png

Figure 1. Surface-based modeling of atrial electrophysiology

The objective of this work is to revisit the asymptotic convergence properties – with respect to the thickness parameter – of the earlier-proposed 3D-shell model. This shell model is very attractive for engineering applications, in particular due to the possibility of directly using a general 3D constitutive law in the corresponding finite element formulations. We establish strong convergence results for the 3D-shell model in the two main types of asymptotic regimes, namely, bending- and membrane-dominated behavior. This is an important achievement, as it completely substantiates the asymptotic consistency of the 3D-shell model with 3D linearized isotropic elasticity. See paper [14].

6.1.4. Mechanical modeling and numerical methods for poromechanics: Applications to cardiac perfusion

Participants: Bruno Burtschell, Dominique Chapelle, Philippe Moireau.

We have previously formulated a rather general modeling framework of poromechanics – formulations that combine solid and fluid components to represent the behavior of a porous medium – to take into account large deformations and rapid fluid flows, see [29]. This allows to consider, in particular, the application of blood perfusion within the cardiac tissue, which – indeed – features these specific complex phenomena, out of the scope of classical poromechanical models. One of our major objectives now, within the PhD of Bruno Burtschell, is to propose and assess some associated relevant numerical schemes, which requires special care regarding both space and time discretizations. In a first stage, in order to ease numerical prototyping and assessments, an axisymmetric reduction of the model has been formulated, and some existing algorithms of fluid-structure interaction have been implemented within this axisymmetric framework (in FreeFEM++). The rationale is that our poromechanics formulations feature some rather deep similarities to so-called Arbitrary-Lagrangian-Eulerian (ALE) fluid-structure formulations, hence the latter are considered as a natural starting point for further extensions.

6.1.5. Personalized modeling for cardiac amyloidosis diagnosis

Participants: Alessandro Felder, Dominique Chapelle, Philippe Moireau, Jean-François Deux [Hôpital Henri Mondor], Thibault Damy [Hôpital Henri Mondor].

Cardiac amyloidosis is a condition induced by pathological deposition of amyloid proteins within muscle tissue and nerves, thus severely impairing the cardiac function and often requiring cardiac transplantation as the only available treatment. Our objective here in a first stage is to use our previously developed patient-specific modeling methodologies to analyse some clinical cases – based on actual patient data – to better apprehend the impact of the pathology on biomechanical properties. Further perspectives include the modeling of the protein deposition and associated tissue remodeling in order to predict the disease evolution in a patient-specific context. This work is performed in collaboration with medical doctors from Hôpital Henri Mondor (Créteil).

6.2. Model-Data Interaction

6.2.1. State observers of a vascular fluid-structure interaction model through measurements in the solid

Participants: Cristobal Bertoglio [Reo Project-Team], Dominique Chapelle, Miguel Angel Fernández [Reo Project-Team], Jean-Frédéric Gerbeau [Reo Project-Team], Philippe Moireau.

We analyze the performances of two types of Luenberger observers – namely, the so-called Direct Velocity Feedback and Schur Displacement Feedback procedures, originally devised for elasto-dynamics – 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. We first assess the observers using hemodynamics-inspired test problems with the complete model, including the Navier-Stokes equations in Arbitrary Lagrangian-Eulerian formulation, in particular. Then, in order to obtain more detailed insight we consider several well-chosen simplified models, each of which allowing a thorough analysis – emphasizing spectral considerations – while illustrating a major phenomenon of interest for the observer performance, namely, the added mass effect for the structure, the coupling with a lumped-parameter boundary condition model for the fluid flow, and the fluid dynamics effect *per se*. Whereas improvements can be sought when additional measurements are available in the fluid domain in order to more effectively deal with strong uncertainties in the fluid state, in the present framework this establishes Luenberger observer methods as very attractive strategies – compared, e.g., to classical variational techniques – to perform state estimation, and more generally for uncertainty estimation since other observer procedures can be conveniently combined to estimate uncertain parameters. See paper [11].

6.2.2. Improving Efficiency of Data Assimilation Procedure for a Biomechanical Heart Model by Representing Surfaces as Currents

Participants: Alexandre Imperiale, Alexandre Routier [Aramis Team], Philippe Moireau, Stanley Durrleman [Aramis Team].

We adapt the formalism of currents to compare data surfaces and surfaces of a mechanical model and we use this discrepancy measure to feed a data assimilation procedure. We apply our methodology to perform parameter estimation in a biomechanical model of the heart using synthetic observations of the endo- and epicardium surfaces of an infarcted left ventricle. We compare this formalism with a more classical signed distance operator between surfaces and we numerically show that we have improved the efficiency of our estimation justifying the use of state-of-the-art computational geometry formalism in the data assimilation measurements processing. See paper [24].

6.2.3. Optimal observer for parabolic problems

Participants: Karine Mauffrey, Philippe Moireau.

../../../../projets/m3disim/IMG/cropped_mesh_img_4.png

Figure 2. Heart model immersed in MR image data

We aim at proposing optimal observers strategies for reconstructing the solution of general systems of PDEs using avalaible observations, including both wave-type equations and heat-like equations. The main objective of this work is to present a self-contained analysis. For a general parabolic system, we have established the exponential stability of the operator occurring in the equation satisfied by the error between the target and the optimal observer. The proof relies on two major hypotheses: an observability inequality satisfied by the observation operator and a controllability property for the modeling error operator by which model noises enter the dynamics (controllability property which is related to the invertibility of the solution of the associated infinite dimensional Riccati equation). The next questions we want to tackle are the discretisation of the model and the construction of a reduced Kalman filter.

MADYNES Project-Team

6. New Results

6.1. Android Security

Participants: Olivier Festor, Abdelkader Lahmadi [contact], Eric Finickel.

Android-based devices include smart phones and tablets that are now widely adopted by users because they offer a huge set of services via a wide range of access networks (WiFi, GPRS/EDGE, 3G/4G). Android provides the core platform for developing and running applications. Those applications are available to the users over numerous online marketplaces. These applications are posted by developers, with little or no review process in place, leaving the market self-regulated by users. This policy generates a side-effect where users are becoming targets of different malicious applications which the goal is to steal their private information, collect all kind of sensitive data via sensors or abusing granted permissions to make surtaxed calls or messages. To address this security issue, monitoring the behaviour of running applications is a key technique enabling the identification of malicious activities.

During 2013, we have designed and extended a monitoring framework integrating observed network and system activities of running Android applications. We extended and enhanced our modular NetFlow probe [48] running on android devices to export observed network flow records to a collection point for their processing and analysis. Our embedded probe includes a new set of IPFIX information elements that we have designed [41] to encapsulate geographic location information within exported flows. This work was done in collaboration with the Univerisity of Twente, where they developed the flow collector and the analysis application.

We have also developed an embedded logging probe that exports available logs generated by an Android device to a big data enabled store [25]. We have analyzed the collected logs using TreeMapping visualization technique [46] to display behavioral graphs of Android applications. The generated graphs are able to provide an aggregated view of the different components of a running application. This view is useful to improve the understanding of the behaviour of an application.

6.2. Sensor networks monitoring

Participants: Rémi Badonnel, Alexandre Boeglin, Isabelle Chrisment, Olivier Festor, Abdelkader Lahmadi [contact], Anthea Mayzaud, Bilel Saadallah.

Low Power and Lossy Networks (LLNs) are made of interconnected wireless devices with limited resources in terms of energy, computing and communication. The communication channels are low-bandwidth, high loss rate and volatile wireless links subject to failure over time. They are dynamic and the connectivity is limited and fluctuant over time. Each node may loss frequently its connectivity with its neighborhood nodes. In addition, link layer frames have high constrains on their size and throughput is limited. These networks are used for many different applications including industrial automation, smart metering, environmental monitoring, homeland security, weather and climate analysis and prediction. The main issue in those networks is optimal operation combined with strong energy preservation. Monitoring, i.e the process of measuring sampled properties of nodes and links in a network, is a key technique in operational LLNs where devices need to be constantly or temporally monitored to assure their functioning and detect relevant problems which will result in an alarm being for- warded to the enterprise network for analysis and remediation.

We developed and designed a novel algorithm and a supporting framework [16] that improves a distributed poller-pollee based monitoring architecture. We empower the poller-pollee placement decision process and operation by exploiting available routing data to monitor nodes status. In addition, monitoring data is efficiently embedded in any messages flowing through the network, drastically reducing monitoring overhead. Our approach is validated through both simulation, implementation and deployment on a 6LoWPAN-enabled network. Results demonstrate that our approach is less aggressive and less resource consuming than its competitors.

In a previous work, we developed a first fully operational content centric networking protocol stack (CCNx) dedicated to a wireless sensor network. During this year, we have extended this implementation and designed a novel monitoring service [32] to efficiently aggregate data in a WSN. The developed solution has been implemented in the Contiki operating system and evaluated using the Cooja simulator. We have compared the performance of our proposed solution with the SPIN protocol in terms of the number of exchanged messages and response times. Our results show that our solution provides better performance for collecting and aggregating data inside the network using operators such as maximum or average.

This year, we also analyzed security attacks against LLN networks, and more specifically those targeting the RPL routing protocol. In that context, we introduced a taxonomy in order to classify these attacks into three main categories. The attacks against resources, such as DIS flooding attacks and increased rank attacks, permit to reduce the network lifetime through the generation of fake control messages or the building of RPL loops. The attacks against the topology, such as wormhole attacks or DAO inconsistency attacks, permit the network to converge to a sub-optimal configuration or to isolate one or several nodes. Finally, attacks against network traffic, such as eaves-dropping attacks and decreased rank attacks, permit to capture and analyse large part of the RPL traffic.

Based on this taxonomy, we compared the properties of attacks and discussed methods and techniques for monitoring them. In particular, we are investigating efficient solutions for supporting security monitoring in these resource-constrained environments [17]. We considered DODAG inconsistency attacks as a first case study. Scenarios were constructed to evaluate the performance of the RPL network when such attacks are carried out. Via an implementation in Contiki, it was identified that the internal mechanism proposed by RPL, which involves ignoring packets with the appropriate IPv6 header after a fixed threshold is reached, uses an arbitrary value for the threshold. A new function that dynamically scales this threshold was developed to improve performance of the network while under attack. In addition, a comparative study between the (1) no threshold, (2) fixed threshold and (3) dynamic threshold scenarios has been performed.

6.3. Monitoring of anonymous networks

Participants: Isabelle Chrisment [contact], Olivier Festor, Juan Pablo Timpanaro.

Anonymous networks have emerged to protect the privacy of network users and to secure the data exchange over the Internet. Nevertheless, the monitoring of these networks has not been investigated very much and only few networks have been studied. Large scale monitoring on these systems allows us to understand how they behave and which type of data which is shared among users.

In 2013, we continued our research about anonymous systems, with a special focus on the I2P network [3]. The I2P network provides an abstraction layer to permit two parties to communicate in an anonymous and secure manner. This network is optimized for anonymous web hosting and anonymous file-sharing. I2P's file-sharing community is highly active, where users deploy their file-sharing applications on top of the network. I2P uses a variation of Onion routing, thus assuring the unlinkability between a user and its file-sharing application.

Current statistics service for the I2P network do not provide values about the type of applications deployed in the network nor the geographical localization of users. We conducted the first large-sale monitoring on the I2P anonymous system, characterizing users and services running on top of the network. We first designed and implemented a distributed monitoring architecture based on probes placed in the I2P's distributed hash table (I2P's netDB), which allows us to collect a vast amount of network metadata. So, our distributed monitoring architecture provides us with different insights about the I2P network.

We were able to detect the behavior of particular applications, notably their period of activity. By considering the behavior of a particular anonymous service along with a particular set of I2P users, we determined in which measure this set of users was responsible for the activity of the anonymous service. We thus conducted a correlation analysis between the behavior of I2P users from two top cities along with the behavior of anonymous file-sharing clients (I2PSnark clients) throughout a particular period of time. By applying

[3] http://i2p2.de

Pearson's correlation coefficient, we achieved a group-based characterization and we determined that the activity of users from those cities explained 38% of all detected file-sharing activity [22], [2].

Starting from our limitations to de-anonymise a particular I2P user, we studied I2P's unidirectional tunnels and the mechanism used to create these tunnels. We discovered a vulnerability in this mechanism, vulnerability which allows an attacker to detect whether a user is the last participant in an inbound tunnel. With this knowledge, we showed that it would be possible to attack an I2P's eepsite in order to de-anonymise the eepsite's operator [39].

6.4. Configuration security automation

Participants: Rémi Badonnel [contact], Martin Barrere, Olivier Festor.

The main research challenge addressed in this work is focused on enabling configuration security automation in autonomic networks and services. In particular our objective is to increase vulnerability awareness in the autonomic management plane in order to prevent configuration vulnerabilities. The continuous growth of networking significantly increases the complexity of management. It requires autonomic networks and services that are capable of taking in charge their own management by optimizing their parameters, adapting their configurations and ensuring their protection against security attacks. However, the operations and changes executed during these self-management activities may generate vulnerable configurations.

A first part of our work in the year 2013 has been dedicated to the issue of past hidden vulnerable states [8]. Even though a known vulnerability may not be present on a current system, it could have been unknowingly active in the past providing an entry point for attacks that may still constitute a potential security threat in the present. Indeed, vulnerabilities can survive within active systems for a long period of time without being known. During this period, attackers may perform well-planned and clean attacks (e.g., stealing information) without being noticed by security entities (e.g., system administrators, intrusion detection systems, self-protection modules). Changes on the system or even its normal activity can alter or erase the remaining evidence on the current configuration. In that context, we have defined a new strategy for assessing past hidden vulnerable states. This solution is based on a mathematical model for describing and detecting unknown past security exposures and on an OVAL-based framework able to autonomously build and monitor the evolution of network devices and to outsource the assessment of their exposure in an automatic manner. We also have developed an implementation prototype that efficiently performs assessment activities over an SVN repository of IOS system images. Experimental results have confirmed the feasibility and scalability of our solution.

A second part aimed at light-weighting the vulnerability assessment process in the context of mobile devices [9]. Security activities imply a consumption of resources that should be taken to a minimum in order to maximize the performance and responsiveness of such critical environments. Sometimes users may prefer to deactivate security processes such as antivirus software instead of having a short battery lifetime. The proposed approach centralizes main logistic vulnerability assessment aspects as a service while mobile clients only need to provide the server with required data to analyze known vulnerabilities described with the OVAL language. By configuring the analysis frequency as well as the percentage of vulnerabilities to evaluate at each security assessment, our probabilistic solution permits to bound client resource allocation and also to outsource the assessment process. The strategy consists in distributing evaluation activities across time thus alleviating the workload on mobile devices, and simultaneously ensuring a complete and accurate coverage of the vulnerability dataset. This technique results in a faster assessment process, typically done in the cloud, and considerably reduces the resource allocation on the client side. A prototype of our vulnerability assessment framework for Android has been selected and presented during the demonstration session of the IEEE/IFIP IM'2013 international conference [10].

We are currently investigating new methods for remediating known vulnerabilities, formalizing the change decision problem as a satisfiability or SAT problem [27]. By specifying our vulnerability knowledge source as a logical formula, fixing those system properties we can not change and freeing those variables for which changes are available, our objective is to use a SAT solving engine for determining what changes have to be made so as to secure the system. In order to provide proactive and reactive solutions, we are interested in the concept of future state descriptions to specify how a system will look like after applying a specific change.

6.5. Cache Management in CCN

Participants: Thomas Silverston [contact], César Bernardini, Olivier Festor.

The Internet is currently mostly used for accessing content. Indeed, ranging from P2P file sharing to current video streaming services such as Youtube, it is expected that content will count for approximately 86% of the global consumer traffic by 2016.

While the Internet was designed for -and still focuses on- host-to-host communication (IP), users are only interested in actual content rather than source location. Hence, new Information-Centric Networking architectures (ICN) such as CCN, NetInf, Pursuit have been proposed giving high priority to efficient content distribution at large scale. Among all these new architectures, Content Centric Networking (CCN) has attracted considerable attention from the research community [4].

CCN is a network architecture based on named data where a packet address names content, not location. The notion of host as defined into IP does not exist anymore. In CCN, the content is not retrieved from a dedicated server, as it is the case for the current Internet. The premise is that content delivery can be enhanced by including per-node-caching as content traverses the network. Content is therefore replicated and located at different points of the network, increasing availability for incoming requests.

As content is cached along the path, it is crucial to investigate the caching strategy for CCN Networks and to propose new schemes adapted to CCN. We therefore designed *Most Popular Content* (MPC), a new caching strategy for CCN network [12], [11].

Instead of storing all the content at every nodes on the path, MPC strategy caches only popular content. With MPC, each node counts all the requests for a content and when it has been requested a large amount of time, the content will be cached at each node along the path. Otherwise, the content is not popular; it is transmitted but it is not cached into the network.

We implemented MPC into the ccnSim simulator and evaluate it through extensive simulations.

Our results demonstrate that using MPC strategy allow to achieve a higher Cache Hit in CCN networks and still reduces drastically the number of replicas. By caching only popular content, MPC helps at reducing the cache load at each node and the network resource consumption.

We expect that our strategy could serve as a base for studying name-based routing protocols. Being a suggestion based mechanism, it is feasible to adapt it to manage content among nodes, to predict popularity and to route content to destination. In addition, we are currently investigating the social relationship between users to improve our caching strategy for CCN networks.

Besides, Online Social Networks (OSN) have gained tremendous popularity on the Internet. Millions of users interact with each other through OSN such as Facebook or Twitter. New ubiquitous devices (smartphones, tablets) appeared and include functionalities to instantaneously share information through OSN. As a central component of CCN is in-network caching, the content's availability depends on several criteria such as cache strategies and replacement policies, cache size or content popularity. OSN carry extremely valuable information about users and their relationships. This knowledge can help to drastically improve the efficiency of Content Centric Networks. Thus, we propose to include social information in the design of a new caching strategy for Content Centric Networking. We designed *SACS*, a novel caching strategy for CCN based on the social information of users [28]. Our socially-aware caching strategy gives priority to content issued by Influential users and cache it pro-actively into the CCN network. We performed simulations of our caching strategy and show its ability to improve the cache performances of CCN. In addition, we implemented a prototype on PlanetLab and performed large-scale experiments. Our solution improves the caching performances of CCN by 2.5 times on real testbed.

6.6. QoS in Wireless Sensor Networks

Participants: François Despaux, Abdelkader Lahmadi, Evangelia Tsiontsiou, Kévin Roussel, Moutie Chehaider, Ye-Qiong Song [contact].

[4] http://www.ccnx.org

WSN research focus has progressively been moved from the energy issue to the QoS issue. Typical example is the MAC protocol design, which cares about not only low duty-cycle, but also high throughput with self-adaptation to dynamic traffic changes. Our research on WSN QoS is thoroughly organized in four topics:

- self-adaptive MAC protocol for both QoS and energy efficiency

  By combining our two previous MAC protocols called Queue-MAC and CoSenS, we extended Queue-MAC to iQueue-MAC to support multi-hop transmission [23], [6]. iQueue-MAC provides immediate yet energy-efficient throughput enhancement for dealing with burst or heavy traffic. Combined with CSMA/CA, iQueue-MAC makes use of queue length of each sensor node and allocates suitable TDMA slots to them for packets transmission. During light traffic period, no extra slots will be allocated; iQueue-MAC acts like other low duty-cycle MACs to conserve power. While in burst or heavy traffic period, iQueue-MAC senses the build up of packet queues and dynamically schedules adequate number of slots for packet transmission. Within ANR QUASIMODO project, we have implemented iQueue-MAC on STM32W108 chips that offer IEEE 802.15.4 standard communication. We set up several real-world experimental scenarios, including a 46 nodes multi-hop test-bed for simulating a general application, and conducted numerous experiments to evaluate iQueue-MAC, in comparison with other traffic adaptive duty-cycle protocols, such as multi- channel version RI-MAC and CoSenS. Results clearly show that iQueue-MAC outperforms multi-channel version of RI-MAC and CoSenS in terms of packet delay and throughput.

- QoS routing

  For supporting different QoS requirements, routing in WSN must simultaneously consider several criteria (e.g., minimizing energy consumption, hop counts or delay, packet loss probability, etc.). When multiple routing metrics are considered, the problem becomes a multi-constrained optimal path problem (MCOP), which is known as NP-complete. In practice, the complexity of the existing routing algorithms is too high to be implemented on the low cost and power constrained sensor nodes. Recently, Operator calculus (OC) has been developed by Schott and Staples with whom we collaborate. OC can be applied to solving MCOP problem with lower complexity and can deal with dynamic topology changes (which is the case in duty-cycled WSN). Through intensive numerical experiments, we have shown that OC has much less complexity compared with SAMCRA, known as one of the best existing algorithms. Sub-optimal paths can be obtained with a distributed version of OC, and following this principle, a first OC-based routing protocol is implemented over Contiki rime stack on TelosB motes. Its improvement and performance evaluation, as well as its integration to uIP/RPL stack is our ongoing work.

- Systems and middleware for supporting QoS in wireless sensor networks

  For supporting new protocols implementation which require to interact with low level services (MAC, Radio drivers, hardware timers) and integration to the Internet of Things approach, we focused on the OS for WSN. Several contributions have been made available for both ContikiOS (https://github.com/contiki-os/contiki/pull/519) and RiotOS (https://github.com/RIOT-OS/RIOT/pull/408, https://github.com/RIOT-OS/RIOT/pull/459). This allows to preparing for the next step towards the implementation of iQueue-MAC on both ContikiOS and RiotOS and compare experimentally with other protocols. In parallel and as part of LAR project, we also investigated the integration of different types of WSN using a gateway to make the data access transparent following RESTful webservice through CoAP/UPD/6loWPAN [24].

- End-to-end performance in multi-hop networks

  Probabilistic end-to-end performance guarantee may be required when dealing with real-time applications. As part of ANR QUASIMODO project, we are dealing with Markov modeling of multi-hop networks running slotted CSMA/CA (beacon enabled mode of IEEE 802.15.4). One of the problem of the existing models resides in their strong assumptions that may not be directly used to assess the end-to-end delay in practice. In particular, realistic radio channel, capture effect and OS-related implementation factors are not taken into account [15], [14]. We proposed to explore a new approach which is based on process mining to extract the Markov chain model from the execution of the protocol code.

6.7. Routing in Wireless Sensor Networks

Participants: Emmanuel Nataf [contact], Patrick-Olivier Kamgueu.

Our work on the estimation of the remaining energy inside a sensor is published in [18]. We have integrated this model in the standard routing protocol for wireless sensors networks (RPL) and compared our energy based routing against a routing plane based on the quality of transmission between sensors [30].

We have built a new model to combine together several criteria, as the remaining energy, the expected transmission rate and the hop count into one quality indicator. To achieve this, we propose to use fuzzy logic either because it is a recognized mathematical tool for combining heterogeneous data and because it can be implemented with a small memory footprint. Our work is fully integrated in the standard protocol and does not need additional messages or new protocol states.

We bought 35 sensors and deployed them in the Loria building. The goal of this deployment is manyfold :

- to build and observe a real network in a real environment;
- to provide the team with a demonstrative tool to help the understanding of our work;
- to provide the team with a testbed for other works on IoT, like the security monitoring or the QoS.

6.8. Online Risk Management

Participants: Rémi Badonnel [contact], Oussema Dabbebi, Olivier Festor.

Telephony over IP has known a large scale deployment and has been supported by the standardization of dedicated signaling protocols. This service is however exposed to multiple attacks due to a lower confinement in comparison to traditional PSTN networks. While a large variety of methods and techniques has been proposed for protecting VoIP networks, their activation may seriously impact on the quality of such a critical service. Risk management provides new opportunities for addressing this challenge. In particular our work aims at performing online risk management for VoIP networks and services. The objective is to dynamically adapt the service exposure with respect to the threat potentiality, while maintaining a low security overhead.

In the year 2013, these efforts on VoIP risk management have led the PhD defense of Oussema Dabbebi. This work has been structured into three axes [1]. The first axis concerns the automation of the risk management process in VoIP enterprise network. In this context, we have developed a mathematical model for assessing risk, a set of progressive countermeasures to counter attackers and mitigation algorithms that evaluate the risk level and takes the decision to activate a subset of countermeasures [4]. To improve our strategy, we have coupled it with an anomaly detection system based on SVM and a self-configuration mechanism which provides feedback about countermeasure efficiency. The second axis deals with the extension of our adaptive risk strategy to P2PSIP infrastructures. We have implemented a specific risk model and a dedicated set of countermeasures with respect to its peer-to-peer nature. For that, we have identified attack sources and established different threat scenarios. We have analysed the RELOAD framework and proposed trust mechanisms to address its residual attacks. Finally, the third axis focuses on VoIP services in the cloud where we have proposed a risk strategy and several strategies to deploy and apply countermeasures [5].

6.9. Pervasive Computing

Participants: Laurent Ciarletta [contact], Olivier Festor, Ye-Qiong Song, Yannick Presse, Emmanuel Nataf.

Vincent Chevrier, Thomas Navarrete Gutierrez and Julien Vaubourg (MAIA team) did contribute to part of this activity.

In Pervasive or Ubiquitous Computing, a growing number of communicating/computing devices are collaborating to provide users with enhanced and ubiquitous services in a seamless way. In a related field, Cyber Physical Systems also are technological systems that have to be considered within a physical world and its contraints. They are complex systems where several inter-related phenomena have to be considered. In order to be studied, modeled and evaluated, we propose the use of co-simulation and multimodeling.

Pervasive Computing is about interconnected and situated computing resources providing us(ers) with contextual services. These systems, embedded in the fabric of our daily lives, are complex: numerous interconnected and heterogeneous entities are exhibiting a global behavior impossible to forecast by merely observing individual properties. Firstly, users physical interactions and behaviors have to be considered. They are influenced and influence the environment. Secondly, the potential multiplicity and heterogeneity of devices, services, communication protocols, and the constant mobility and reorganization also need to be addressed. Our research on this field is going towards both closing the loop between humans and systems, physical and computing systems, and taming the complexity, using multi-modeling (to combine the best of each domain specific model) and co-simulation (to design, develop and evaluate) as part of a global conceptual and practical toolbox. We're applying this work on UAvs, dynamic networks (adhoc, mesh, P2P, wireless sensors and actuators), energy-constrained / location aware services, smart grids etc.

Such systems can be seen as complex and are present everywhere in our environment: internet, electricity distribution networks, transport networks. This systems have as characteristics: a large number of autonomous entities, dynamic structures, different time and space scales and emergent phenomena.

Application domains such as Smart Spaces, Smart Cities, Smart Trasportation Systems and Smart Grid makes us sometimes use Smart* or SmartX as a generic word. Madynes is focusing on the networking aspects of such systems and on the tools to develop and assess them. We cooperate with other teams and most notably the Maia team to be able to encompass issues and research questions that combine both networking and cognitive aspects.

In 2013 we worked on the following research topics :

- Assessment and evaluation of complex systems. Continuing the work on multi-modeling and co-simulation, we have participated with the MAIA team on the development of an architecture for the control of complex systems based on multi-agent simulation, a CPS co-simulation (next item) and a Smart grid simulation tool (last item), and continue working on the AA4MM framework (Agents and artefacts for Multiple heterogeneous Models).

  A control architecture has been proposed by Tomas Navarrete, based on an “equation-free” approach. We use a multi-agent model to evaluate the global impact of local control actions before applying the most pertinent set of actions. Associated to our architecture, an experimental platform has been developed to confront the basic ideas or the architecture within the context of simulated “free-riding” phenomenon in peer to peer file exchange networks. We have demonstrated that our approach allows to drive the system to a state where most peers share files, despite given initial conditions that are supposed to drive the system to a state where no peer shares. We have also executed experiments with different configurations of the architecture to identify the different means to improve the performance of the architecture.

  This work helped us to identify [13] the key issues related to the usage of the multi-agent paradigm in the context of control of complex systems.

- In Cyper Physical Systems, we have lead the design and implementation of the Aetournos (Airborne Embedded auTonomOUs Robust Network of Objects and Sensors) platform at Loria. The idea of AETOURNOS is to build a platform which can be at the same time a demonstrator of scientific realizations and an evaluation environment for research works of various teams of our laboratory. It is also its own research domain : building a completely autonomous and robust flock of collaborating UAVs.

  In Madynes, we focus on the CPS and their networks and applications. Those systems consist of numerous autonomous elements in sharp interaction which functioning require a tight coupling between software implementations and technical devices. The collective movements of a flock of flying communicating robots / UAVs, evolving in potentially perturbed environment constitute a good example of such a system. Indeed, if we look at the level of each of the elements playing a role into this system, a certain number of challenges and scientific questions can be studied: respect of real-time constraints of calculations for every autonomous UAV and for the communication between the robots, conception of individual, embedded, distributed or global management systems, development of self-adaptative mechanisms, conception of algorithms of collective movement etc... Furthermore, the answers to each of these questions have to finally contribute to the global functioning of the system. Applying co-simulation technique we plan to develop a hybrid "network-aware flocking behavior" / "behavior aware routing protocol". The platform is composed of several high-grade research UAVs (Pelican quadcopters and Firefly hexacopters) and lighter models (AR.Drone quacopters). We have provided a working set of tools : multi-simulation behavior / network / physics and generic software development using ROS (Robot Operating System). The UAVs carry a set of sensor for location awareness, their own computing capabilities and several wireless networks.

This work is discribed in a position paper where a first implementation of a formation flight is detailed [29].

- Smart grids and Smart spaces are another application domain. MS4SG (cf. has given us the opportunity to link multi-simulations tools such as HLA (High Level Architecture) and FMI (Functional Mockup Interface) thanks to our AA4MM framework. We've so far successfully applied our solution to the simulation of smart appartment complex and to combining the electrical and networking part of a Smart Grid (first deliverable and first workshop with EDF R&D, Supélec and SIANI were in september 2013). A paper has also been accepted to Simutools 2014. In 2014, we will continue working on the hybrid protocols and on the UAV platform, and apply our co-simulation work to Smart Grids and other Smart* [13].

6.10. SCADA Systems Security

Participants: Olivier Festor, Abdelkader Lahmadi [contact], Bilel Saadallah.

SCADA is a term used in several industries and it stands for *Supervisory Control and Data Acquisitions*. It refers to a centralized control and monitoring system for a variety of machinery and equipment involved with many industrial activities including: power generation and distribution, transportation, nuclear plants, manufacturing processes, etc. SCADA systems use a family of network protocols (PROFINET, MODBUS, DNP3) to monitor and control these industrial activities or even our homes. SCADA systems are becoming target to different attacks exploiting traditional IT vulnerabilities, e.g. buffer overflows, script crossing, crafted network packets, or specific vulnerabilities related to control and estimation algorithms employed by control processes. Several of them are daily discovered and disclosed or remain still unknown. The most threaten accidents in SCADA networks are caused by targeted attacks, where adversaries exploit those vulnerabilities available in software or network protocols components to disturb and make damage to the physical process. Therefore, it is important to provide new methods and tools for protecting SCADA network from malicious cyber attacks targeting physical processes and infrastructures.

During the year 2013, we have firstly designed and setup a SCADA test bed [31] to be able to analyze and develop security methods for several controlled physical systems. The testbed uses a Profinet based network to control experimental real-time simulated physical processes through hardware programmable logic controllers (PLCs). Secondly, we have developed a novel methodology to automatically discover a pattern of behaviour of a running controlled system through the analysis of communication messages traveling in its control loop network. The method applies process mining techniques on the exchanged communication packets between control and feedback devices to infer a model of the controlled running system. The extracted model will be then used to build a tailored anomaly-based intrusion detection module for the studied system.

6.11. Dynamic resource allocation for network virtualization

Participants: Said Seddiki, Bilel Nefzi, Mounir Frikha, Ye-Qiong Song [contact].

The objective of this research topic is to develop different resource allocation mechanisms in Network Virtualization, for creating multiple virtual networks (VNs) from a single physical network. It is accomplished by logical segmentation of the network nodes and their physical links. Sharing resources and improving utilization are the main idea of virtualization. Finding effective solutions for the needs expressed by users without deteriorating the performance of different VNs is a research challenge. In addition, solutions should meet different performance criteria required by network infrastructure.

We proposed several approaches that aim to select substrate nodes [21] with sufficient CPU, disk, and other resources, as well as substrate links with enough spare bandwidth [19], [20]. These dynamic approaches, where online monitoring of the VN is required, allow adaptively changing the resource allocations. We have shown through simulations that the proposed approaches offer higher utilization of physical network and better managing the satisfaction of virtual networks by minimizing the packet delays inside the physical node. They also provide a fair and efficient allocation of link capacity and avoid bottlenecks. The next step is the implementation of these propositions using OPENFLOW in a software defined network.

6.12. Crowdsourcing Services

Participants: Thomas Silverston [contact], Olivier Festor, Abdelkader Lahmadi, Elian Aubry.

Nowadays cities invest more in their public services, and particularly digital ones, to improve their resident's quality of life and attract more people. Thus, new crowdsourcing services appear and they are based on contributions made by mobile users equipped with smartphones. For example, the respect of the traffic code is essential to ensure citizens' security and welfare in their city. We therefore designed CrowdOut, a new mobile crowdsourcing service for improving road safety in cities. CrowdOut allows users to report traffic offense they witness in real time and to map them on a city plan. CrowdOut has been implemented and experiments and demonstrations have been performed in the urban environment of the Grand Nancy, in France. This service allows users appropriating their urban environment with an active participation regarding the collectivity. This service also represents a tool for city administrators to help for decisions and improve their urbanization policy, or to check the impact of their policy in the city environment.

MAESTRO Project-Team

5. New Results

5.1. Network Science

Participants: Eitan Altman, Konstantin Avrachenkov, Mahmoud El Chamie, Julien Gaillard, Philippe Nain, Giovanni Neglia, Marina Sokol.

5.1.1. Epidemic models of propagation of content

In [15], E. Altman and P. Nain, in collaboration with Y. Xu (MAESTRO member at the time of submission) and A. Shwartz (Technion, Israel), focus on the propagation of content in peer-to-peer (P2P) networks. They first study the transient behavior of some P2P networks whenever information is replicated and disseminated according to epidemic-like dynamics. They then use the insight gained from the previous analysis in order to predict how efficient could measures taken against P2P networks be. They first introduce a stochastic model which extends a classical epidemic model, and characterize the P2P swarm behavior in presence of free riding peers. They then study a second model in which a peer initiates a contact with another peer chosen randomly. In both cases the network is shown to exhibit phase transitions: a small change in the parameters causes a large change in the behavior of the network. The authors show, in particular, how phase transitions affect measures of content providers against P2P networks that distribute non-authorized music, books or articles, and what is the efficiency of counter-measures. In addition, this analytic framework can be generalized to characterize the heterogeneity of cooperative peers.

5.1.2. The design of recommendation systems (RS) for social networks

Recommendation systems take advantage of products and users information in order to propose items to targeted consumers. In [50], J. Gaillard, E. Altman, M. El Bèze and E. Ethis (both from Univ. Avignon) propose a framework to overcome the usual scalability issues of nowadays systems. The system includes a dynamic adaptation to enhance the accuracy of rating predictions by applying a new similarity measure. They perform several experiments on films data from Vodkaster, showing that systems incorporating dynamic adaptation improve significantly the quality of recommendations compared to static ones.

In [51] the same authors propose new modifications of the recommendation algorithm that allow not only to present a recommendation but also to propose a list of words which appeared frequently in recommendations of other people who watched that film and who have been identified to have similar preferences, according to their opinions on common movies.

5.1.3. Network centrality measures

A class of centrality measures called betweenness centralities reflects degree of participation of edges or nodes in communication between different parts of the network. The original shortest-path betweenness centrality is based on counting shortest paths which go through a node or an edge. One of shortcomings of this metric is that it ignores the paths that might be one or two hops longer than the shortest paths, while the edges on such paths can be important for communication processes in the network. To rectify this shortcoming a current flow betweenness centrality has been proposed. Similarly to the shortest-path betweenness, it has prohibitive complexity for large size networks. In [42] K. Avrachenkov, N. Litvak (Univ. of Twente, the Netherlands), V. Medyanikov (St. Petersburg State Univ., Russia) and M. Sokol propose and analyze two regularizations of the current flow betweenness centrality, α-current flow betweenness and truncated α-current flow betweenness, which can be computed fast and correlate well with the original current flow betweenness. In particular, the new centrality measures indicate well vulnerability of a network.

5.1.4. Average consensus protocols

Information can flow in a network through communication links connecting the nodes. Not all the links have the same importance and it is common in complex networks to distinguish “weak” links/ties and “strong” ones. Depending on the specific network, the strength of a link connecting two nodes can be quantified by its transmission capacity, the inter-meeting rate between the two nodes, the level of mutual trust of the two nodes, etc.. The topology of connections and the strength of the links are two factors that affect the speed of spread of information in the network. In [63], M. El Chamie and G. Neglia in collaboration with L. Severini (student at Univ. of Nice Sophia Antipolis, France) have shown that the topology can have stronger effect on the information spread than the strength of the links. In particular, they have considered an iterative belief propagation process as in average consensus protocols where each node in the network has a certain belief (a real number) that is updated iteratively by the weighted average of the nodes’ belief and the ones they connected to. They have shown by simulations on random graphs that a topological optimization can have a significant faster spread of beliefs than any weight selection optimization techniques. They have also given a 2-hop message averaging that performs faster convergence than standard algorithms.

The activity on “Reducing communication overhead of average consensus protocols”, described in MAESTRO’s 2012 activity report has lead to the publication [49].

5.2. Wireless Networks

Participants: Eitan Altman, Philippe Nain, Giovanni Neglia, Oussama Habachi.

5.2.1. Delay Tolerant Networks

We have pursued our study of optimal control in delay tolerant network. We studied the trade-off between delivery delay and energy consumption in a delay tolerant network in which a message (or a file) has to be delivered to each of several destinations by epidemic relaying. In addition to the destinations, there are several other nodes in the network that can assist in relaying the message. The optimal control policy was obtained in the mean-field limit of large number of mobiles by C. Singh, E. Altman, A. Kumar and R. Sundaresan in [33].

Our analysis of DTNs so far was done with mobility models in which all individuals move independently of each other. In [61]. S. Patil, M. Kumar and E. Altman have studied through simulations the multicast time in DTNs where the mobility of individuals follow dependent movement such as the one of flocking birds. This model is typical to cooperative movement and could be useful to describe a rescue team in an area hit by a disaster. We showed the impact of the parameters defining the mobility on the multicast time. If instead of broadcasting packets one first codes them (using network coding) then one can obtain substantial gain in the performance. This is shown in the case that all packets that are to be sent are available for coding before transmission. In [16], E. Altman studies in collaboration with F. de Pellegrini (CREATE-NET) and L. Sassatelli how to optimally decide on the amount of coded packets to create as a function of time in the case that the information to be coded is not available before transmission. This allows to optimize the system performance for the case of real-time traffic.

In [11], A. Ali, M. Panda, T. Chahed and E. Altman design and study a reliable transport protocol for DTNs consisting of both unicast and multicast flows. The improvement in reliability is brought in by a novel Global Selective ACKnowledgment (G-SACK) scheme and random linear network coding (RLC). The motivation for using network coding and G-SACKs comes from the observation that one should take the maximum advantage of the contact opportunities which occur quite infrequently in DTNs. Network coding and G-SACKs perform “mixing” of packet and acknowledgment information, respectively, at the contact opportunities and essentially solve the randomness and finite capacity limitations of DTNs. In contrast to earlier work on network coding in DTNs, we observe and explain the gains due to network coding even under an inter-session setting. Our results from extensive simulations of appropriately chosen “minimal” topologies quantify the gains due to each enhancement feature. In a related publication [67], A. Ali, L. Sassatelli, E. Altman and T. Chahed present an overview of theoretical background that is used for evaluating transport protocols in DTNs.

In [13], E. Altman formulates in collaboration with A. P. Azad, T. Başar (Univ. Illinois at Urbana Champain) and F. De Pellegrini (CREATE-NET) a problem where both transmission and activation of mobile terminals are controlled as a linear optimal control problem. They solve the problem by making use of this linearity in order to obtain explicit expressions for the objective function as a function of the control actions trajectories (rather than as a function of both actions and state trajectories). This allows them to compute the optimal strategies explicitly.

In [26], E. Altman studies in collaboration with D. Fiems (Ghent Univ.) a class of Markov-modulated stochastic recursive equations. This class includes multi-type branching processes with immigration as well as linear stochastic equations. Conditions are established for the existence of a stationary solution and expressions for the first two moments of this solution are found. Furthermore, the transient characteristics of the stochastic recursion are investigated: the first two moments of the transient solution are obtained as well. Finally, to illustrate the approach, the results are applied to the performance evaluation of packet forwarding in delay-tolerant mobile ad-hoc networks.

In [34], G. Neglia in collaboration with X. Zhang, H. Wang (both from Fordham Univ., Bronx, USA), J. Kurose and D. Towsley (both from Univ. of Massachusetts at Amherst, USA) has also investigated the benefits of applying Random Linear Coding (RLC) to unicast application in DTNs. Under RLC, nodes store and forward random linear combinations of packets as they encounter each other. For the case of a single group of packets originating from the same source and destined for the same destination, they have proved a lower bound on the probability that the RLC scheme achieves the minimum time to deliver the group of packets. Although RLC achieves a significant reduction in group delivery delay, it fares worse in terms of average packet delivery delay and network transmissions. When replication control is employed, RLC schemes reduce the group delivery delay without increasing the number of transmissions. In general, the benefit achieved by RLC is more significant under stringent resource (bandwidth and buffer) constraints, limited signaling, highly dynamic networks, and when it is applied to packets from same flow. For more practical settings with multiple continuous flows in the network, the researchers have shown the importance of deploying RLC schemes with a carefully tuned replication control in order to achieve reduction in average delay.

In [60], the same authors investigated the problem of determining the routing that minimizes the maximum/average delivery time or the maximum/average delivery delay for a set of packets in a deterministic Delay Tolerant Network, i.e. in a network for which all the nodes' transmission opportunities are known in advance. While the general problem with multiple sources and multiple destinations is NP-hard, they have presented a polynomial-time algorithm that can efficiently compute the optimal routing in the case of a single destination or of a single packet that needs to be routed to multiple destinations.

In [59], P. Nain in collaboration with D. Towsley (Univ. of Massachusetts at Amherst, USA), A. Bar-Noy and F. Yu (both from City Univ. of New York, USA), P. Basu (Raytheon BBN Technologies, USA), and M. P. Johnson (Univ. of California, Los Angeles, USA) consider the problem of estimating the end-to-end latency of intermittently connected paths in disruption/delay tolerant networks. While computing the time to traverse such a path may be straightforward in fixed, static networks, doing so becomes much more challenging in dynamic networks, in which the state of an edge in one timeslot (i.e., its presence or absence) is random, and may depend on its state in the previous timeslot. The authors compute the expected traversal time (ETT) for a dynamic path in a number of special cases of stochastic edge dynamics models, and for three different edge failure models, culminating in a surprisingly nontrivial yet realistic "hybrid network" setting in which the initial configuration of edge states for the entire path is known. The ETT for this "initial configuration" setting can be computed in quadratic time (as a function of path length), by an algorithm based on probability generating functions. Several linear-time upper and lower bounds on the ETT are provided and evaluated using numerical simulations.

5.2.2. Interference coordination in wireless networks

In [47], R. Combes, E. Altman and Z. Altman (Orange Labs, Issy les Moulineaux) model a LTE wireless network with Inter-Cell Interference Coordination (ICIC) at the flow level where users arrive and depart dynamically, in order to optimize quality of service indicators perceivable by users such as file transfer time for elastic traffic. They propose an algorithm to tune the parameters of ICIC schemes automatically based on measurements. The convergence of the algorithm to a local optimum is proven, and a heuristic to improve its convergence speed is given. Numerical experiments show that the distance between local optima and the global optimum is very small, and that the algorithm is fast enough to track changes in traffic on the time scale of hours. The proposed algorithm can be implemented in a distributed way with very small signaling load.

In [46], the same authors introduce self-organizing mechanisms as control loops, and study the conditions for stability when running control loops in parallel. Based on control theory, they propose a distributed coordination mechanism to stabilize the system. In certain cases, coordination can be achieved without any exchange of information between control loops. The mechanism remains valid in the presence of noise via stochastic approximation. Instability and coordination in the context of wireless networks are illustrated with two examples. The paper is essentially concerned with linear systems, and the applicability of our results for non-linear systems is discussed.

5.2.3. Streaming over wireless

The Quality of Experience (QoE) of streaming service is often degraded by playback interruptions. To mitigate these, the media player prefetches streaming contents before starting playback, at a cost of delay. In [66], Y. Xu, S. E. Elayoubi, E. Altman and R. El-Azouzi study the QoE of streaming from the perspective of flow dynamics. First, a framework is developed for QoE when streaming users join the network randomly and leave after downloading completion. They compute the distribution of prefetching delay using partial differential equations, and the probability generating function of playout buffer starvation using ordinary differential equations. Second, they extend the framework to characterize the throughput variation caused by opportunistic scheduling at the base station in the presence of fast fading. This study reveals that the flow dynamics is the fundamental reason of playback starvation. The QoE of streaming service is dominated by the average throughput of opportunistic scheduling, while the variance of throughput has very limited impact on starvation behavior.

5.2.4. Dynamic coverage of mobile sensor networks

B. Liu (Univ. of Massachusetts at Lowell, USA), O. Dousse (Nokia Research Center, Switzerland), P. Nain, and D. Towsley (Univ. of Massachusetts at Amherst, USA) study in [30] the dynamic aspects of the coverage of a mobile sensor network resulting from continuous movement of sensors. As sensors move around, initially uncovered locations may be covered at a later time, and intruders that might never be detected in a stationary sensor network can now be detected by moving sensors. However, this improvement in coverage is achieved at the cost that a location is covered only part of the time, alternating between covered and not covered. The authors characterize area coverage at specific time instants and during time intervals, as well as the time durations that a location is covered and uncovered. They further consider the time it takes to detect a randomly located intruder and prove that the detection time is exponentially distributed. For mobile intruders, a game theoretic approach allows to derive optimal mobility strategies for both sensors and intruders. The optimal sensor strategy is to choose the direction uniformly at random between 0 and 2π. The optimal intruder strategy is to remain stationary. This solution represents a mixed strategy which is a Nash equilibrium of the zero-sum game between mobile sensors and intruders.

5.2.5. Wireless network security

The activity on "Fast and secure rendezvous protocols for mitigating control channel DoS attacks" described in MAESTRO's 2012 activity report has lead to the publication [35].

5.3. Network Engineering Games

Participants: Eitan Altman, Konstantin Avrachenkov, Ilaria Brunetti, Julien Gaillard, Majed Haddad, Manjesh Kumar Hanawal, Alexandre Reiffers.

5.3.1. Association problem

In [32], A. Silva, in collaboration with H. Tembine, E. Altman and M. Debbah, study a non-cooperative association game where mobiles associate to Base Stations. The authors solve the problem using the theory of optimal transportation after incorporating in it the effect of network congestion. They are able to find a closed form expression for its solution. The authors also solve a global optimization problem for minimizing the total power needed by the mobile terminals over the whole network.

5.3.2. Cognitive radio

In [52] O. Habachi considers a non-cooperative Opportunistic Spectrum Access (OSA) where Secondary Users (SUs) access opportunistically the spectrum licensed for Primary Users (PUs) in TV white spaces (TVWS). As sensing licensed channels is time and energy consuming, the author considers a hierarchical Cognitive Radio (CR) architecture, where CR base stations sense a subset of the spectrum in order to locate some free frequencies. Thereafter, a SU that needs to communicate through TVWS sends a request to a CR base station for a free channel. The author models the problem using a Partially Observable Stochastic Game (POSG), and he takes into consideration the energy consumption of CR base stations and the Quality of Service of SUs. Since solving POSG optimally may require a significant amount of time and computational complexity, the author then models the OSA problem using a game theoretical approach, and proposes a symmetric Nash equilibrium solution concept. Finally, the simulations that validate the theoretical findings are provided.

In [24], J. Elias (Univ. Paris Descartes), F. Martignon (Univ. Paris Sud), L. Chen and E. Altman address the joint pricing and network selection problem in cognitive radio networks. The problem is formulated as a Stackelberg game where first the Primary and Secondary operators set the network subscription price to maximize their revenue. Then, users perform the network selection process, deciding whether to pay more for a guaranteed service, or use a cheaper, best-effort secondary network, where congestion and low throughput may be experienced. They use the Nash equilibrium concept to characterize the equilibria for the price setting game. On the other hand, a Wardrop equilibrium is used in the network selection game.

5.3.3. Cooperative games in wireless networks

We have pursued this year our new activity on cooperative games in wireless communications. We have pursued our work on coalition games and started working on the area of matching games. In [56], E. Altman, C. Hasan and J.-M. Gorce (both from Inria project-team SOCRATE) have addressed the problem of association of mobiles to base stations which can be viewed as a coalition game. They formulated the game using a stochastic geometric approach (one Poisson point process representing the base stations and another one representing the mobiles) and studied the impact of switching off base stations (for energy efficient operation).

An important class of games within cooperative games is the matching games. They have been used in stable marriage games (in which a bi-partite graph called matching is to be proposed between a group of men and women based on mutual ranking between this group). A second well-known application of matching games is the college admission problem in which students are assigned to colleges based on their preferences as well as on the preferences of the colleges. We introduced and solved two matching games in wireless communication using the theory of matching games. In [55] the same authors study a game similar to the above ones to match pairs of mobiles where one mobile serves as a relay for the other in the absence of a good direct channel to the base station. The utilities studied here are the outage probabilities. In [65], R. Vaca-Ramirez, E. Altman, J. S. Thompson and V. Ramos-Ramos propose a distributed algorithm for energy efficient virtual Multiple-input/Multiple-output coalition formation. They model cooperation as a game derived from the concept of stable marriage with incomplete lists. Single antenna devices such as mobile and relay stations cooperate in order to improve the user's and system's energy efficiency. In both problems above, the performance of the equilibrium is shown to be close to the social optimum and yet the complexity for achieving the equilibrium is only polynomial (whereas that of computing a global optimal matching is NP hard).

In [40] K. Avrachenkov, L. Cottatellucci (EURECOM) and L. Maggi (CREATE-NET, Italy) study multiple access channels whose channel coefficients follow a quasi-static Markov process on a finite set of states. The authors address the issue of allocating transmission rates to users in each time interval, such that optimality and fairness of an allocation are preserved throughout a communication, and moreover all the users are consistently satisfied with it. First, it is shown how to allocate the rates in a global optimal fashion. The authors provide a sufficient condition for the optimal rates to fulfill some fairness criteria in a time-consistent way. Then the authors utilize the game-theoretical concepts of time consistent Core and Cooperation Maintenance. It is demonstrated that in the model the sets of rates fulfilling these properties coincide and they also coincide with the set of global optimal rate allocations. The relevance of the presented dynamic rate allocation to LTE systems is also shown.

5.3.4. Bayesian games in networking

K. Veeraruna, E. Altman, R. El-Azouzi and S. Rajesh have studied in [29] a power control problem in which a base station allocates power according to the channel state as reported by the mobiles. The paper addresses the question of how to allocate the power, given that the channel reported by some non-cooperative mobile may be unreliable. They obtain the equilibrium allocation after formulating the problem as a Bayesian game.

In [38], E. Altman and T. Jiménez consider both a cooperative as well as non-cooperative admission into an M/M/1 queue. The only information available is a signal that says whether the queue size is smaller than some value L or not. They first compute the globally optimal and the Nash equilibrium stationary policy as a function of L. They compare the performance to that of full information and of no information on the queue size. They identify the value of L that optimizes the equilibrium performance.

In [58], K. Ibrahimi, E. Altman and M. Haddad introduce a signaling game approach to power control. They consider two players named player I and player II. They assume that player I only knows his channel state without any information about the channel state of player II and vice-versa. Player I moves first and sends a signal to player II which can be accurate or distorted. Player II chooses his power control strategy based on this information and his belief about the nature of the informed player's information. In order to analyze such a model, the proposed scheme game is transformed into an equivalent 4x4 matrix game. The authors establish the existence of Nash equilibria and then derive it numerically and study its properties.

In [53], M. Haddad and E. Altman, in collaboration with P. Wiecek and H. Sidi, present a Bayesian game theoretic framework for determining the decision to which cell a given mobile user should associate in LTE two-tier Heterogeneous Networks. Users are assumed to compete to maximize their throughput by picking the best locally serving cell with respect to their own measurement, their demand and a partial statistical channel state information of other users. In particular, the authors investigate the properties of a hierarchical game, in which the macro-cell BS is a player on its own. They derive analytically the utilities related to the channel quality perceived by users to obtain the equilibria. They show in the Stackelberg formulation, how the operator, by dynamically choosing the offset about the state of the channel, can optimize its global utility while end-users maximize their individual utilities.

5.3.5. Network neutrality and collusion

Representatives of several Internet access providers have expressed their wish to see a substantial change in the pricing policies of the Internet. In particular, they would like to see content providers pay for use of the network, given the large amount of resources they use. This would be in clear violation of the "network neutrality" principle that had characterized the development of the wireline Internet. We proposed and studied possible ways of implementing such payments and of regulating their amount. M. K. Hanawal and E. Altman have pursued in [54] working on network neutrality studying various ways of collusion between an ISP and a content provider and in particular, another form of non-neutrality in which a content provider signals to an ISP information on the popularity of its content and hides this information from other ISPs. They define and compute the price of collusion and study the impact of such signalling on the ISP that is in collusion as well as on the other ones.

In the situation just described, the demand is modelled to be elastic. In contrast, in [62], A. Reiffers and E. Altman study in collaboration with Y. Hayel pricing issues in non-neutral network with non-elastic traffic. A Stackelberg equilibrium is derived and the price of collusion is computed.

Our research on network neutrality started already on 2010 with a research report [83] that has now been published in [14]. We already reported on this publication in 2011 when it became available electronically.

5.3.6. Competition over popularity in social networks

In [39] E. Altman, P. Kumar, S. Venkatramanan and A. Kumar consider a situation where several content producers send their content to some subscriber of a social network. These posts appear on the subscriber's timeline which is assumed to have finite capacity. Whenever a new post arrives to the timeline, an older post leaves it. Therefore to be visible, a source has to keep sending contents from time to time. Each source is modelled as a player in a non-cooperative game in which one trades between the utility for being visible on the timeline and the cost (or effort) for keeping sending content. This game is solved in a Markovian setting the performance measures of interest are computed.

In [37], E. Altman in cooperation with F. De Pellegrini (CREATE-NET), D. Miorandi, T. Jiménez and R. El-Azouzi study situations in which subscribers of a social network take the decision whether to access or not some content, based on the number of views that the content has. Their analysis aims at understanding the way in which information about the quality of a given content can be deduced from view counts when only part of the viewers that access the content are informed about its quality. In this paper they present a game formulation for the behavior of individuals using a mean-field model: the number of individuals is approximated by a continuum of atomless players and for which the Wardrop equilibrium is the solution concept. They derive conditions on the problem's parameters that result in the emergence of threshold equilibria policies. But they also identify some parameters in which other structures are obtained for the equilibrium behavior of individuals.

5.3.7. Evolutionary games

Evolutionary game theory is a relatively young mathematical theory that aims at formalizing in mathematical terms evolution models in biology. In recent years this paradigm has penetrated more and more into other areas such as the linguistics, economics and engineering. The current theory of evolutionary game makes an implicit assumption that the evolution is driven by selfishness of individuals who interact with each other. In mathematical terms this can be stated as "an individual equals a player in a non-cooperative game model". This assumption turns out to be quite restrictive in modeling evolution in biology. It is now more and more accepted among biologist that the evolution is driven by the selfish interests of large groups of individuals; a group may correspond for example to a whole beehive or to an ants' nest. In [43] and [71], I. Brunetti and E. Altman propose an alternative paradigm for modeling evolution where a player does not necessarily represent an interacting individual but a whole class of such individuals. In [71] in particular, they use Markov Decision Evolutionary Games (MDEG) to allow a parent and a child represent the same individual at different states. This is yet another enhancement in what we understand as a player. An important contribution is in the study of the Hawk and Dove game in these new frameworks.

In [27], M. Haddad, J. Gaillard, E. Altman and D. Fiems (Ghent Univ.) study an evolutionary game in the MDEG framework of power control. Aging is taken into account by assuming that as the battery of the mobile becomes empty, high power is not available anymore. The goal of a mobile is to use power that maximizes the amount of traffic it can transmit during its lifetime. We restrict in this work to policies that are state independent and compute the equilibrium.

5.4. Green Networking and Smart Grids

Participants: Sara Alouf, Eitan Altman, Nicaise Choungmo Fofack, Delia Ciullo, Alain Jean-Marie, Giovanni Neglia.

5.4.1. Stochastic geometry methods for wireless design issues

In [64] the issue of energy efficiency in Orthogonal Frequency-Division Multiple Access (OFDMA) wireless networks is discussed by D. Tsilimantos, J.-M. Gorce (Inria project-team SOCRATE) and E. Altman. Their interest is focused on the promising concept of base station (BS) sleep mode, introduced recently as a key feature in order to dramatically reduce network energy consumption. The proposed technical approach fully exploits the properties of stochastic geometry, where the number of active cells is reduced in a way that the outage probability, or equivalently the signal to interference plus noise (SINR) distribution, remains the same. The optimal energy efficiency gains are then specified with the help of a simplified but yet realistic BS power consumption model. Furthermore, the authors extend their initial work by studying a non-singular path loss model in order to verify the validity of the analysis and finally, the impact on the achieved user capacity is investigated. In this context, the significant contribution of this paper is the evaluation of the theoretically optimal energy savings of sleep mode, with respect to the decisive role that the BS power profile plays.

5.4.2. Analysis of base stations with autonomous energy supply

S. Alouf, A. Jean-Marie and D. Ciullo have started the modeling of wireless communication base stations with autonomous energy supply (solar, wind). One challenge is to account for the random and non-stationary input of energy. A second challenge is to find the correct time and space granularity of the model, so as to ensure both the practical relevance of the model and numerical tractability. The activity will be backed up by a measurement campaign on the Com4Innov platform (http://www.com4innov.com/), that will provide information on energy consumption of different traffic patterns.

5.4.3. Demand-response system

Energy demand aggregators are new actors in the energy scenario: they gather a group of energy consumers and implement a demand-response paradigm. When the energy provider needs to reduce the current energy demand on the grid, it can pay the energy demand aggregator to reduce the load by turning off some of its consumers loads or postponing their activation. Currently this operation involves only greedy energy consumers like industrial plants. In [48], [78] A. Jean-Marie and G. Neglia in collaboration with G. Di Bella, L. Giarré, M. Ippolito and I. Tinnirello (all from Univ. of Palermo, Italy) have studied the potential of aggregating a large number of small energy consumers like home users as it may happen in smart grids. In particular they have addressed the feasibility of such approach by considering which scale the aggregator should reach in order to be able to control a significant power load. The challenge of the study derives from residential users' demand being much less predictable than that of industrial plants. For this reason they have resorted to queuing theory to study analytically the problem and quantify the trade-off between load control and tolerable service delays.

5.5. Content-Oriented Systems

Participants: Sara Alouf, Konstantin Avrachenkov, Nicaise Choungmo Fofack, Delia Ciullo, Alain Jean-Marie, Philippe Nain, Giovanni Neglia, Marina Sokol.

5.5.1. Performance evaluation of hierarchical TTL-based cache networks

N. Choungmo Fofack, P. Nain and G. Neglia, together with D. Towsley (Univ. of Massachusetts at Amherst, USA) have revisited and extended the work that has appeared in [82]. They consider caches that implement an expiration-based eviction policy to manage contents in their memory. These caches are called Time-To-Live (TTL)-based caches. These TTL-based caches can be used to model caches running classical replacement policies such as Least Recently Used (LRU) and Random Replacement (RND). The main characteristic of the latter TTL-based cache models is that they (re)initialize the TTL of a content at both cache hit and cache miss. In a paper that is currently under review, the case of a network of caches where requests for each content are routed as a polytree is analyzed and a framework to evaluate the performance of such general TTL-based cache networks is proposed.

5.5.2. Modeling modern DNS caches

Motivated by the recent behavior of Domain Name System (DNS) caches that do not respect the timeout marked (by Authoritative DNS servers) on resource records, N. Choungmo Fofack and S. Alouf propose in [44] a theoretical model based on renewal arguments to describe this modern behavior. The proposed model for a cache taken in isolation is validated with real traces collected by Inria's IT service at Sophia-Antipolis at one of the Inria's DNS caches. The model of a network of caches is validated by event-driven simulations. This study suggests that, when inter-request times have a concave cumulative distribution function, client caches (those caches that are fed directly by users requests) should keep each resource record for a constant duration (that may depend on its popularity). However, core caches should draw their timeout values for each record from a distribution which has as high coefficient of variation as possible.

5.5.3. An approximate analysis of general and heterogeneous cache networks

Jointly with M. Dehghan, D. L. Goeckel and D. Towsley (Univ. of Massachusetts at Amherst, USA), N. Choungmo Fofack proposes a simple, accurate, and computationally efficient framework to assess performance of network of caches with arbitrary topology, requests described by renewal processes, and caches running Least Recently Used (LRU), First-In First-Out (FIFO), or Random Replacement (RND) policies. Their framework is based on the characteristic time approximation of LRU, RND and FIFO caches that helps to model the latter as TTL-based caches. Classical results of the theory of (renewal) point processes (e.g. approximation of general point processes by renewal processes, thinning a renewal point process, aggregating/merging independent renewal processes) are used as well as theoretical results established in [82] and [44] on TTL-based caches (e.g. calculation of metrics of interest such hit and occupancy probabilities, characterization of miss streams).

5.5.4. Data placement

Jointly with J.-C. Bermond (Inria project-team COATI), D. Mazauric (Univ. Aix-Marseille) and J. Yu (UFV Vancouver), A. Jean-Marie has pursued the study of combinatorial designs that solve the problem of replicating optimally data over unreliable servers, with the objective of minimizing the variance of the availability of documents. In a forthcoming revision of [81], they use results from Design Theory, particularly the existence of "large triple systems" to solve multiple instances of the problem.

5.5.5. Semi-supervised learning with application to P2P systems

Semi-supervised learning methods constitute a category of machine learning methods which use labelled points together with unlabelled data to tune the classifier. The main idea of the semi-supervised methods is based on an assumption that the classification function should change smoothly over a similarity graph, which represents relations among data points. This idea can be expressed using kernels on graphs such as graph Laplacian. Different semi-supervised learning methods have different kernels which reflect how the underlying similarity graph influences the classification results. In [41] K. Avrachenkov, P. Gonçalves (Inria project-team DANTE) and M. Sokol analyze a general family of semi-supervised methods, provide insights about the differences among the methods and give recommendations for the choice of the kernel parameters and labelled points. In particular, it appears that it is preferable to choose a kernel based on the properties of the labelled points. They illustrate our general theoretical conclusions with an analytically tractable characteristic example, clustered preferential attachment model and classification of content in P2P networks.

5.6. Advances in Methodological Tools

Participants: Konstantin Avrachenkov, Alain Jean-Marie, Philippe Nain.

5.6.1. Perturbation analysis

In [21] K. Avrachenkov and J.-B. Lasserre (LAAS-CNRS) investigate the analytic perturbation of generalized inverses. Firstly the authors analyze the analytic perturbation of the Drazin generalized inverse (also known as reduced resolvent in operator theory). The approach is based on spectral theory of linear operators as well as on a new notion of group reduced resolvent. It allows one to treat regular and singular perturbations in a unified framework. The authors provide an algorithm for computing the coefficients of the Laurent series of the perturbed Drazin generalized inverse. In particular, the regular part coefficients can be efficiently calculated by recursive formulae. Finally, the authors apply the obtained results to the perturbation analysis of the Moore-Penrose generalized inverse in the real domain.

5.6.2. Markov processes

In [20] K. Avrachenkov, L. Cottatellucci (EURECOM), L. Maggi (CREATE-NET, Italy) and Y.-H. Mao (Beijing Normal Univ., China) consider both discrete-time irreducible Markov chains with circulant transition probability matrix P and continuous-time irreducible Markov processes with circulant transition rate matrix Q. In both cases they provide an expression of all the moments of the entropy mixing time. In the discrete case, they prove that all the moments of the mixing time associated with the transition probability matrix $\alpha P+(1-\alpha)P^*$ are maximum in the interval $0\le\alpha\le1$ when $\alpha=1/2$, where P^* is the transition probability matrix of the time-reversed chain. Similarly, in the continuous case, they show that all the moments of the mixing time associated with the transition rate matrix $\alpha Q+(1-\alpha)Q^*$ are also maximum in the interval $0\le\alpha\le1$ when $\alpha=1/2$, where Q^* is the time-reversed transition rate matrix.

In [23] K. Avrachenkov, in collaboration with A. Piunovskiy and Z. Yi (both from Univ. of Liverpool, UK), study a general homogeneous continuous-time Markov process with restarts. The process is forced to restart from a given distribution at time moments generated by an independent Poisson process. The motivation to study such processes comes from modeling human and animal mobility patterns, restart processes in communication protocols, and from application of restarting random walks in information retrieval. The authors provide a connection between the transition probability functions of the original Markov process and the modified process with restarts. Closed-form expressions for the invariant probability measure of the modified process are derived. When the process evolves on the Euclidean space there is also a closed-form expression for the moments of the modified process. The authors show that the modified process is always positive Harris recurrent and exponentially ergodic with the index equal to (or bigger than) the rate of restarts. Finally, the general results are illustrated by the standard and geometric Brownian motions.

5.6.3. Queueing theory

In [22] K. Avrachenkov, P. Nain and U. Yechiali (Tel Aviv Univ., Israel) consider two independent Poisson streams of jobs flowing into a single-server service system having a limited common buffer that can hold at most one job. If a type-i job ($i=1,2$) finds the server busy, it is blocked and routed to a separate type-i retrial (orbit) queue that attempts to re-dispatch its jobs at its specific Poisson rate. This creates a system with three dependent queues. Such a queueing system serves as a model for two competing job streams in a carrier sensing multiple access system. We study the queueing system using multi-dimensional probability generating functions, and derive its necessary and sufficient stability conditions while solving a Riemann-Hilbert boundary value problem. Various performance measures are calculated and numerical results are presented. In particular, numerical results demonstrate that the proposed multiple access system with two types of jobs and constant retrial rates provides incentives for the users to respect their contracts.

5.6.4. Control theory

In conjunction with E. Della Vecchia and S. Di Marco (both from National Univ. Rosario, Argentina), A. Jean-Marie has pursued the studies on the Rolling Horizon procedure and other approximations in stochastic control problems. Inspired by the work of A. Ruszczyński, they have considered Markov Decision problems where the metric to be optimized is a risk measure, a metric which generalizes the mathematical expectation and takes risk aversion of agents into account. For infinite-horizon, risk-averse discounted Markov Decision Processes, they have proved approximation bounds which imply the convergence of approximate rolling horizon procedures when the horizon length tends to infinity. They have also analyzed the effects of uncertainties on the transition probabilities, the cost functions and the discount factors [77].

In [17] K. Avrachenkov, U. Ayesta (LAAS-CNRS), J. Doncel (LAAS-CNRS) and P. Jacko (BCAM, Spain) address the problem of fast and fair transmission of flows in a router, which is a fundamental issue in networks like the Internet. They focus on the relaxed version of the problem obtained by relaxing the fixed buffer capacity constraint that must be satisfied at all time epoch. The relaxation allows one to reduce the multi-flow problem into a family of single-flow problems, for which one can analyze both theoretically and numerically the existence of optimal control policies of special structure. In particular, it is shown that the control can be represented by so-called index policies, but not always by threshold policies. The simulation and numerical results show that the index policy achieves a wide range of desirable properties with respect to fairness between different TCP versions, across users with different round-trip-time and minimum buffer required to achieve full utility of the queue.

5.6.5. Game theory

In [18] K. Avrachenkov, L. Cottatellucci (EURECOM) and L. Maggi (CREATE-NET, Italy) consider simple Markovian games, in which several states succeed each other over time, following an exogenous discrete-time Markov chain. In each state, a different simple static game is played by the same set of players. The authors investigate the approximation of the Shapley-Shubik power index in simple Markovian games (SSM). The authors prove that an exponential number of queries on coalition values is necessary for any deterministic algorithm even to approximate SSM with polynomial accuracy. Motivated by this, the authors propose and study three randomized approaches to compute a confidence interval for SSM. They rest upon two different assumptions, static and dynamic, about the process through which the estimator agent learns the coalition values. Such approaches can also be utilized to compute confidence intervals for the Shapley value in any Markovian game. The proposed methods require a number of queries, which is polynomial in the number of players in order to achieve a polynomial accuracy.

In [19] K. Avrachenkov, L. Cottatellucci (EURECOM) and L. Maggi (CREATE-NET, Italy) study multi-agent Markov decision processes (MDPs) in which cooperation among players is allowed. They find a cooperative payoff distribution procedure (MDP-CPDP) that distributes in the course of the game the payoff that players would earn in the long run game. They show under which conditions such a MDP-CPDP fulfills a time consistency property, contents greedy players, and strengthen the coalition cohesiveness throughout the game. Finally, the authors refine the concept of Core for Cooperative MDPs.

MAGIQUE-3D Project-Team

6. New Results

6.1. Inverse Problems

6.1.1. Reconstruction of an elastic scatterer immersed in a homogeneous fluid

Participants: Hélène Barucq, Rabia Djellouli, Élodie Estécahandy.

The determination of the shape of an obstacle from its effects on known acoustic or electromagnetic waves is an important problem in many technologies such as sonar, radar, geophysical exploration, medical imaging and nondestructive testing. This inverse obstacle problem (IOP) is difficult to solve, especially from a numerical viewpoint, because it is ill-posed and nonlinear. Its investigation requires as a prerequisite the fundamental understanding of the theory for the associated direct scattering problem, and the mastery of the corresponding numerical solution methods.

In this work, we are interested in retrieving the shape of an elastic obstacle from the knowledge of some scattered far-field patterns, and assuming certain characteristics of the surface of the obstacle. The corresponding direct elasto-acoustic scattering problem consists in the scattering of time-harmonic acoustic waves by an elastic obstacle Ω^s embedded in a homogeneous medium Ω^f, that can be formulated as follows:

$$\begin{aligned} \Delta p + (\omega^2/c_f^2)\, p = 0 & \qquad \text{in } \Omega^f \\ \nabla \cdot \sigma(u) + \omega^2 \rho_s\, u = 0 & \qquad \text{in } \Omega^s \\ \omega^2 \rho_f u \cdot n = \partial p/\partial n + \partial e^{i\,(\omega/c_f)\,x\cdot d}/\partial n & \qquad \text{on } \Gamma \\ \sigma(u) n = -pn - e^{i\,(\omega/c_f)\,x\cdot d} n & \qquad \text{on } \Gamma \\ \lim_{r\to+\infty} r\,(\partial p/\partial r - i\,(\omega/c_f)\,p) = 0 & \end{aligned} \tag{10}$$

where p is the fluid pressure in Ω^f whereas u is the displacement field in Ω^s, and $\sigma(u)$ represents the stress tensor of the elastic material.

This boundary value problem has been investigated mathematically and results pertaining to the existence, uniqueness and regularity can be found in [92] and the references therein, among others. We have obtained a new result proving the well-posedness of the problem when the fluid-solid interface is only lipschitzian. This has been published in the Journal of Mathematical Analysis and Applications [20]. We then propose a solution methodology based on a regularized Newton-type method for solving the IOP. The proposed method is an extension of the regularized Newton algorithm developed for solving the case where only the Helmholtz equation is involved, that is the acoustic case by impenetrable scatterers [86]. The direct elasto-acoustic scattering problem defines an operator $F : \Gamma \to p_\infty$ which maps the boundary Γ of the scatterer Ω^s onto the far-field pattern p_∞. Hence, given one or several measured far-field patterns $\widetilde{p}_\infty(\widehat{x})$, corresponding to one or several given directions d and wavenumbers k, one can formulate IOP as follows:

$$\text{Find a shape } \Gamma \ \text{ such that } \quad F(\Gamma)(\widehat{x}) = \widetilde{p}_\infty(\widehat{x}); \quad \widehat{x} \in S^1.$$

At each Newton iteration, we solve the forward problem using a finite element solver based on discontinuous Galerkin approximations, and equipped with high-order absorbing boundary conditions. We have first characterized the Fréchet derivatives of the scattered field and the characterization has been published in the Journal of Inverse and Ill-posed problems [18]. It is worth noting that they are solutions to the same boundary value problem as the direct problem with other transmission conditions. This work has been the object of several talks [63], [50], [36]. Elodie Estécahandy has defended her PhD thesis [14] in September 2013 and two papers will be submitted soon.

6.1.2. hp-adaptive inversion of magnetotelluric measurements

Participants: Hélène Barucq, Julen Alvarez Aramberri, David Pardo.

The magnetotelluric (MT) method is a passive electromagnetic exploration technique. It makes use of natural electric fields which propagate permanently into the Earth. Electric fields induce magnetic waves which can be detected at the surface to produce a map of the subsurface from the determination of the resistivity distribution. Magnetotelluric method is based on the mathematical relation between the magnetic and telluric variations which involve the electric resistivity of the subsurface. It is particularly relevant for the detection of metallic ores and for the study of geothermal sites. It is also used for oil and gas exploration because it provides information on sedimentary basins. It performs well on depth scales varying between a few tens of meters to hundred of kilometers, following the pioneering works of Tikhonov and Cagniard . Magnetotelluric measurements are governed by polarized Maxwell's equations in such a way that Helmholtz equations have to be solved. The geological mapping is constructed from the solution of an Inverse problem which requires computing the Impedance and/or the Resistivity distributions. In this work, we assimilate Earth with a horizontally layered model with possible 2D heterogeneities. Both the size of the direct problem and the required computational times may be excessively large. Indeed, on the one hand, the model of the source requires defining a horizontally sufficiently large thick plate to avoid undesirable effects that could take place around the edges. On the other hand, the inversion of MT measurements typically requires the computation of an accurate solution at the receivers located at different positions. Since traditional hp-goal oriented techniques [98], [97] provide an accurate solution in one single point, we use a multi-goal-oriented algorithm [99] to obtain accurate solutions at all receivers. To get accurate quantities at several positions, it is necessary to increase the size of the mesh. This induces high computational costs in particular because the solution of the inverse problem is based on reiterated solutions of the direct problem. To decrease the computational costs required to perform the inversion, we propose an adaptive multi-dimensional inversion algorithm, which consists in increasing step by step the dimension in which the direct problem and the inversion are solved. At first step, we compute the 1D primary field with a semi-analytical solution and we invert the 1D problem. After that, we introduce the 2D heterogeneities. Regarding the direct problem, we compute the secondary field, thereby, drastically reducing the size of the computational domain for this problem. Then, we perform the inversion using the solution to the 1D Inverse Problem as a regularization term, increasing the robustness of the inversion algorithm.

6.1.2.1. Reverse Time Migration with Elastic Wave Equations

Participants: Hélène Barucq, Henri Calandra, Julien Diaz, Jérôme Luquel.

Even if RTM has enjoyed the tremendous progresses of scientific computing, its performances can still be improved, in particular when applied to strong heterogeneous media. In this case, images have been mainly obtained by using direct arrivals of acoustic waves and the transition to elastic waves including multiples is not obvious, essentially because elastic waves equations are still more computationally consuming. We have thus chosen to consider high-order Discontinuous Galerkin Methods which are known to be well-adapted to provide accurate solutions based upon parallel computing. Now, one of the main drawback of RTM is the need of storing a huge quantity of information which is prohibitive when using elastic waves. For that purpose, we apply the Griewank algorithm [88] following Symes' ideas [101] for the acoustic RTM. The idea is to find a compromise between the number of wave equations to solve and the number of numerical waves that we have to store. This is the so-called Optimal Checkpointing. By reducing the occupancy of the memory, RTM should be efficient even when using elastic waves. The next step is the derivation of accurate imaging conditions, which could take advantage of all the information contained in the elastic wavefield. For acoustic media, Claerbout [83] proposed an imaging condition which is widely used and turns out to be sufficient to accurately reproduce interfaces. But Claerbout conditions do not take wave conversions into account and, since P-wave and S-wave interact with each other, it might be relevant to use an imaging condition including these interactions. This has been done successfully by J. Tromp and C. Morency [102] for seismology applications based upon the inversion of the global Earth. Their approach is based upon the adjoint state and it involves sensitivity kernels which are defined from the propagated and the back-propagated fields. Now, it has been shown in [93] that full wave form inversions using these sensitivity kernels may be polluted by numerical artifacts. One solution is to use a linear combination of the sensitivity kernels to delete artifacts. In this work, we propose then a new imaging condition which construction is inspired from [93] with some approximations required to keep admissible computational costs. We illustrate the properties of the new imaging condition on industrial benchmarks like the Marmousi model. In particular, we compare the new imaging condition with other imaging conditions by using as criteria the quality of the image and the computational costs required by the RTM. The results will be presented at the 2014 ECCOMAS conference in Barcelona.

6.2. Modeling

6.2.1. Implementation of a non-reflecting boundary condition on ellipsoidal boundary

Participants: Hélène Barucq, Anne-Gaëlle Saint-Guirons, Sébastien Tordeux.

The modeling of wave propagation problems using finite element methods usually requires the truncation of the computational domain around the scatterer of interest. Absorbing boundary condition are classically considered in order to avoid spurious reflections. This year we have proposed and tested a formulation which allows to take into account with no extra-cost a quasi-exact radiation condition based on a non local Dirichlet to Neumann operator.

6.2.2. Modeling of small heterogeneities in the context of the time domain wave equation

Participants: Vanessa Mattesi, Sébastien Tordeux.

We have proposed an approximate model to take into account small heterogeneities for the three dimensional time dependent wave equation. One of the most important result of this work is the generalization of the multipole theory (classically written for the Helmholtz equation) to the wave equation. This work has been presented at Waves 2013 and at JSA 2013 [58], [39]

6.2.3. A new modified equation approach for solving the wave equation

Participants: Hélène Barucq, Henri Calandra, Julien Diaz, Florent Ventimiglia.

In order to obtain high-order time-schemes, we are considering an alternative approach to the ADER schemes and to the modified equation technique described in section 3.2 . The two first steps of the construction of the schemes are similar to the previous schemes : we apply a Taylor expansion in time to the solution of the wave equation and we replace the high-order derivatives with respect to the time by high order space operators, using the wave equation. The difference is that we do not use auxiliary variables and we choose to discretize directly the high-order operators in space. These operators can not be discretized by classical finite elements. For the discretization of the biharmonic operator in an homogeneous acoustic medium, both C1 finite elements, such as the Hermite ones, and Discontinuous Galerkin Finite Elements (DGFE) can be used, while in a discontinuous medium, or for higher-order operators, DGFE should be preferred [80]. We have applied this method to the second order wave equation [15] and the numerical results showed that this technique induced less computational burden than the modified equation scheme or the ADER scheme.

In the framework of the PhD thesis of Florent Ventimiglia, we have extended the new method involving p-harmonic operator to the first order formulation of the acoustic wave equation, which is the formulation discretized in the DIVA platform of TOTAL. In this case, the high order operators in space are not are not powers of the Laplace operator but powers of the gradient. Hence, we also had to adapt the space discretization, and we have extended the DG formulation with centered fluxes proposed in [87] to higher order operators. A numerical analysis of performance in 2D indicates that, for a given accuracy, this method requires less computational costs and less storage than the High-Order ADER Scheme. These results have been presented to the SMAI conference [49], Waves 2013 [61], Numerico IV [61] and HF2013 [34]. A paper has been accepted in ESAIM Proceedings [49].

6.2.4. Constructing and using Absorbing Boundary Conditions

6.2.4.1. Higher Order On-Surface Radiation Conditions for elastic scatterers

Participants: Hélène Barucq, Chokri Bekkey, Juliette Chabassier, Julien Diaz.

The numerical simulation of wave propagation is generally performed by truncating the propagation medium and the team works on new Absorbing Boundary Conditions (ABCs), trying to improve the performance of existing conditions. As we explained at Section 3.2 , item **Boundary conditions**, we are developing ABCs for curved boundaries, based on the full factorization of the wave equation. These ABCs should take propagating, grazing and evanescent waves into account. In [17], we have considered the issue of constructing high-order ABCs taking into account both propagating and evanescent waves for the Helmholtz equation. In case of the simulation of acoustic waves diffracted by a solid immersed in a fluid, we investigate the performance of the new ABCs when used as On-Surface Radiation Conditions. The ABCs are set directly on the boundary of the solid. The unbounded problem is then replaced by a problem involving an acoustic pressure computed on the surface of the solid only. Preliminary results have been obtained by considering the toy problem where the scatterer is a disk. Analytic solutions are then available and we show that that taking into account evanescent waves in the ABC could improve the accuracy of classical ABCs by two orders of magnitude at mid-frequency range, for ka between 1 and 100, k being the frequency and a the typical size of the diffracting obstacle. These results have been presented to the Waves 2013 conference [47].

6.2.4.2. Radiation boundary condition at high frequency

Participants: Hélène Barucq, Elodie Estécahandy, Juliette Chabassier, Julien Diaz.

Regarding the solution of the Helmholtz equation at high frequency with finite element methods, it is current to refine the mesh in order to limit the effect of numerical pollution. It is then interesting to dispose of radiation boundary conditions which do not require to set the artificial boundary far from the scatterer. In this work, we have investigated the possibility of bringing closer the artificial boundary when standard conditions are enhanced by the modeling of grazing waves. The preliminary results we obtained show that this new ABC outperforms classical ABCs at high-frequency, for $ka > 50$, and that it is highly accurate, even when the boundary is very close to the scatterer. These results have been presented to the Waves 2013 conference [45]. Now, the next step is to consider an ABC taking the three types of waves into account.

6.2.4.3. Absorbing Boundary Conditions for Tilted Transverse Isotropic Elastic Media

Participants: Hélène Barucq, Lionel Boillot, Henri Calandra, Julien Diaz.

The simulation of wave propagation in geophysical media is often performed in domains which are huge compared to the wavelengths of the problem. It is then necessary to reduce the computational domain to a box. When considering acoustic or elastic isotropic media, this can be done by applying an Absorbing Boundary Condition (ABC) or by adding a Perfectly Matched Layer (PML). However, a realistic representation of the Earth subsurface must include anisotropy and, in particular, the so-called Tilted Transverse Isotropy. Perfectly Matched Layers are known to be unstable for this kind of media and, to the best of our knowledge, no ABC have been proposed yet. We have thus proposed a low-order ABC for TTI media. The construction is based on comparing and then connecting the slowness curves for isotropic and elliptic TTI waves. Numerical experiments illustrate the performance of the new ABC. They are performed by integrating the ABC in a DG formulation of Elastodynamics. When applied in a TTI medium, the new ABC performs well with the same level of accuracy than the standard isotropic ABC set in an isotropic medium. The condition demonstrates also a good robustness when applied for large times of simulation. These results have been presented to the Smai and to the Waves conferences [67], [68] and a paper has been submitted.

6.2.5. Modeling of the damping factor of Multiperforated plates

Participants: Estelle Piot, Vincent Popie, Sébastien Tordeux.

Multiperforated plates are classically used as a damping material. Melling has proposed in [94] a model to estimate the energy dissipated by these devises. However its well-known result should be corrected by a factor two to fit with experimental data. We have proposed a correct way to compute the energy dissipated by the multiperforated plates. This work has been presented at the Fifth International Scientific Conference and Young Scientists School "Theory and Computational Methods for Inverse and Ill-posed Problems.

6.2.6. Performance Assessment of IPDG for the solution of an elasto-acoustic scattering problem

Participants: Hélène Barucq, Rabia Djellouli, Élodie Estécahandy.

We present a solution methodology for the direct elasto-acoustic scattering problem that falls in the category of Discontinuous Galerkin methods. The method distinguishes itself from the existing methods by combining high-order Discontinuous Galerkin approximations, local stabilizations for the coupled problem and the use of curved element edges on the boundaries. We present some numerical results that illustrate the salient features and highlight the performance of the proposed solution methodology on the resonance phenomenon existing in the elastic scatterer for simple geometries such as circles. Moreover, the designed method ensures a convergence order with a gain of two order of magnitude compared to polygonal boundaries, and a potential to address both mid- and high-frequency regimes. These results have been accepted for publication in International Journal for Numerical Methods in Engineering [19].

6.2.7. On the influence of curvature on transmission conditions

Participants: Hélène Barucq, Martin Gander, Yingxiang Xu.

Domain decomposition methods are both highly successful parallel solvers and also important modeling tools, since problems in subdomains can be treated by adapted methods to the physics in each subdomain. Subdomain boundaries are therefore rarely straight lines. The focus of this paper is to study the influence of curvature on transmission conditions used in optimized Schwarz methods. For straight interfaces and simple geometries, optimized interface conditions are typically determined using Fourier analysis.Asymptotically, these optimized conditions are still valid for curved interfaces.Since however the curvature is the most important information for a smooth curve, we want to study in this paper if and how the interface curvature influences the constants in the optimized parameters.

We consider the model problem

$$(\Delta - \eta)u = f, \quad \text{on } \Omega = \mathbb{R}^2,\ \eta > 0, \tag{11}$$

and we require the solution to decay at infinity.We decompose Ω into two overlapping subdomains $\Omega_1 = (-\infty, a(y)) \times \mathbb{R}$ and $\Omega_2 = (b(y), \infty) \times \mathbb{R}$, where Γ_1 given by $a(y)$ and Γ_2 given by $b(y)$ are smooth curves satisfying $a(y) \geq b(y)$. A general parallel Schwarz algorithm is then given by

$$\begin{aligned} (\Delta - \eta)u_i^n &= f && \text{in } \Omega_i, \\ \mathcal{B}_i(u_i^n) &= \mathcal{B}_i(u_j^{n-1}) && \text{on } \Gamma_i,\ 1 \leq i \neq j \leq 2, \end{aligned} \tag{12}$$

where $\mathcal{B}_i, i = 1, 2$, are transmission conditions to be chosen. If $\mathcal{B}_i,\ i = 1, 2$ are chosen as $\partial_{n_i} + DtN_i$, with DtN_i the Dirichlet to Neumann operators, the iterates will converge in two steps. These operators are however non-local, and thus difficult to use in practice. Therefore, local approximations are used in optimized Schwarz methods. We presented two different approaches to take the curvature of interfaces into account in the transmission conditions of optimized Schwarz methods: micro-local analysis, and analysis using a circular model problem. In both cases, we obtained curvature dependent transmission conditions. A preliminary comparison shows that the transmission conditions based on optimization perform better on the model problem, and that it could be important to take the curvature into account in transmission conditions. In our opinion it is however essential to do a more thorough theoretical and numerical study on more general geometry, where micro-local analysis is still applicable, before we can definitely draw conclusions. This work has been published in the peer-reviewed proceedings of the conference Decomposition Methods in Science and Engineering XXI [51].

6.2.8. Operator Based Upscaling for Discontinuous Galerkin Methods

Participants: Hélène Barucq, Théophile Chaumont, Christian Gout.

Realistic numerical simulations of seismic wave propagation are difficult to handle because they must be performed in strongly heterogeneous media. Two different scales must then be taken into account. Indeed, the medium heterogeneities are currently very small compared to the characteristic dimensions of the propagation medium. To get accurate numerical solutions, engineers are then forced to use meshes that match the finest scale representing the heterogeneities. Meshing the whole domain with a fine grid leads then to huge linear systems and the computational cost of the numerical method is then too high to consider 3D realistic simulations. To dispose of a numerical method allowing to represent the heterogeneity of the medium accurately while computing on a coarse grid is thus relevant. This is the challenge of multiscale approaches like homogenization or upscaling. The ultimate objective of this work is to develop a software package for the Helmholtz equation set in heterogeneous domains. We focus on the operator-based upscaling method. Operator-based upscaling methods were first developed for elliptic flow problems (see [81]) and then extended to hyperbolic problems (see [90], [104], [103]). Operator-based upscaling method consists in splitting the solution into a coarse and a fine part. The coarse part is defined on a coarse mesh while the fine part is computed on a fine mesh. In order to speed up calculations, artificial boundary conditions (ABC) are imposed. By enforcing suitable ABCs on the boundary of every cells of the coarse mesh, calculations on the fine grid can be carried out locally. The coarse part is next computed globally on the coarse mesh. Operator-based upscaling methods were so far developed in joint with standard finite element discretization strategy. In this work, we investigate the idea of combining an operator based upscaling method with discontinuous Galerkin finite element methods(DGFEM). During this year, we have performed the mathematical analysis of the Helmholtz equation set in a domain represented by a discontinuous velocity. The analysis has been achieved both for the continuous and the discretized problem which is based on a quadrature scheme which allows to take the discontinuities of the velocity into account. We get new stability results for stratified-like domains and numerical experiments show that in case of industrial benchmarks, the quadrature scheme leads to lower computational costs with a very good level of accuracy. A paper is in preparation and the results will be presented at the 2014 ECCOMAS conference in Barcelona.

6.2.9. Efficient solution methodology based on a local wave tracking strategy for high-frequency Helmholtz problems.

Participants: Mohamed Amara, Sharang Chaudhry, Julien Diaz, Rabia Djellouli, Steven Fiedler.

We have proposed a procedure for selecting basis function orientation to improve the efficiency of solution methodologies that employ local plane-wave approximations. The proposed adaptive approach consists of a local wave tracking strategy. Each plane-wave basis set, within considered elements of the mesh partition, is individually or collectively rotated to best align one function of the set with the local propagation direction of the field. Systematic determination of the direction of the field inside the computational domain is formulated as a minimization problem. As the resultant system is nonlinear with respect to the directions of propagation, the Newton method is employed with exact characterization of the Jacobian and Hessian. To illustrate the salient features and evaluate the performance of the proposed wave tracking approach, we present error estimates as well as numerical results obtained by incorporating the procedure into a prototypical plane-wave based approach, the least-squares method (LSM) developed by Monk and Wang [95]. The numerical results obtained for the case of a two-dimensional rigid scattering problem indicate that (a) convergence was achievable to a prescribed level of accuracy, even upon initial application of the tracking wave strategy outside the pre-asymptotic convergence region, and (b) the proposed approach reduced the size of the resulting system by up to two orders of magnitude, depending on the frequency range, with respect to the size of the standard LSM system. These results has been presented to the Waves 2013 [43]and in a Research Report [71]. A paper has been submitted.

6.2.10. Mesh Free Frontier-based Formulation (MF3) for High Frequency Helmholtz Problems.

Participants: Mohamed Amara, Julien Diaz, Rabia Djellouli.

We have proposed a novel approach for solving efficiently Helmholtz problems. The proposed solution method employs a boundary-type formulation without however involving Green functions and/or incurring singular integrals. In addition, this approach does not necessitate the use of a mesh. For these reasons, the method is named Mesh Free Frontier-based Formulation (MF3). Furthermore, the sought-after field is locally approximated using a set of basis of functions that consists of a Bessel kind function computed at a prescribed finite set of points. The number of the functions determines mainly the size of the resulting system, the complexity, as well as the computational cost of the proposed method. Preliminary numerical results obtained in the case of 2D-Helmholtz problems in the high-frequency regime are presented to illustrate the computational efficiency of MF3 (the method delivers results with high accuracy level, about 10^-8 on theL^2 relative error, while requiring the solution of small linear systems). In addition, these results tend to suggest that MF3 is pollution free. These results has been presented to two conferences [32], [44]. A paper is in preparation.

6.2.11. Energy based simulation of a Timoshenko beam in non-forced rotation. Application to the flexible piano hammer shank.

Participants: Juliette Chabassier, Marc Duruflé.

A nonlinear model for a vibrating Timoshenko beam in non-forced unknown rotation is derived from the virtual work principle applied to a system of beam with mass at the end. The system represents a piano hammer shank coupled to a hammer head. An energy-based numerical scheme is then provided, obtained by non classical approaches. A major difficulty for time discretization comes from the nonlinear behavior of the kinetic energy of the system. Numerical illustrations are obtained by coupling this new numerical scheme to a global energy-preserving numerical solution for the whole piano. These numerical results show that the pianistic touch clearly influences the spectrum of the piano sound of equally loud isolated notes. These differences do not come from a possible shock excitation on the structure, nor from a changing impact point, nor a “longitudinal rubbing motion” on the string, since neither of these features are modeled in our study.

This work has been submitted for publication in Journal of Sound and Vibration [79]. discretization

6.2.12. Simulating the propagation of ultra short laser pulses in a dispersive non linear medium.

Participants: Juliette Chabassier, Marc Duruflé, Nayla Herran.

This collaboration with CEA-CESTA aimed at evaluating the limits of validity of the MIRÓ software provided by CEA. The MIRÓ software implements the propagation of laser pulses with a non-linear Schrodinger-like equation obtained from Maxwell’s equations in non-linear dispersive medium, assuming that the pulse spectrum is narrow and the non-linearity small enough. When considering intense ultra-short pulses, the spectrum bandwidth and the amplitude of the pulse may violate the constitutive assumptions of the model used by MIRÓ. This collaboration began with the internship of Nayla Herran (march 2013- august 2013). During this internship, different alternative models have been explored, and some of them were able to provide a solution much more accurate than MIRÓ’s models. The comparisons between alternative models and MIRÓ’s model have been performed in 1-D on small cases, for which the reference solution could be obtained from a direct numerical simulation of non-linear Maxwell’s equations. An efficient and accurate solver for non-linear Maxwell’s equations has been implemented in Montjoie software. This solver uses high order finite element in space, and high order Runge-Kutta-Nystrom scheme in time. The space grid moves with respect to the group velocity such that the computational domain stays relatively small and centered around the pulse. Thanks to this solver, we have been able to validate the different models and compare them. These 1-D promising results encourage us to continue this collaboration in order to obtain efficient and accurate numerical methods in 3-D. Our desire is also to be able to perform realistic cases involving physically relevant phenomena.

6.2.13. Asymptotic Modeling for Elasto-Acoustics

Participants: Julien Diaz, Victor Péron.

In the papers [31], [74], we derive equivalent conditions and asymptotic models for the diffraction problem of elastic and acoustic waves in a solid medium surrounded by a thin layer of fluid medium. Due to the thinness of the layer with respect to the wavelength, this problem is well suited for the notion of equivalent conditions and the effect of the fluid medium on the solid is as a first approximation local. We derive and validate equivalent conditions up to the fourth order for the elastic displacement. These conditions approximate the acoustic waves which propagate in the fluid region. This approach leads to solve only elastic equations. The construction of equivalent conditions is based on a multiscale expansion in power series of the thickness of the layer for the solution of the transmission problem.

Questions regarding the implementation of the conditions have been addressed carefully and the boundary conditions have been integrated without changing the structure of the code Hou10ni.

This work has been presented in two international conferences and Workshops : Workshop HPC-GA and WAVES'2013 [55].

A paper with numerical results for the elasto-acoustic problem with several configurations (a thin layer of variable thickness; coupling with an exterior "acoustic" medium) is in preparation.

6.2.14. Thin layer models for electromagnetics

Participants: Marc Duruflé, Victor Péron, Clair Poignard.

We have considered the transmission of electromagnetic waves through a thin layer. This thin layer can be replaced by transmission conditions. Media with thin inclusions appear in many domains: geophysical applications, microwave imaging, biomedical applications, cell phone radiations, radar applications, non-destructive testing. Our application concerns also media for which the conductivity drops inside the layer, such that the low-frequency regime has a different behavior from the mid-frequency regime. Different models are compared for these two regimes, drawbacks and disadvantages of each model are detailed. This work has been accepted for publication [27].

6.2.15. Corner Asymptotics of the Magnetic Potential in the Eddy-Current Model

Participants: Monique Dauge, Patrick Dular, Laurent Krähenbühl, Victor Péron, Ronan Perrussel, Clair Poignard.

The following results rely on a problematic developed in section 3.2 , item **Asymptotic modeling**.

In the paper [25], we describe the magnetic potential in the vicinity of a corner of a conducting body embedded in a dielectric medium in a bidimensional setting. We make explicit the corner asymptotic expansion for this potential as the distance to the corner goes to zero. This expansion involves singular functions and singular coefficients. We introduce a method for the calculation of the singular functions near the corner and we provide two methods to compute the singular coefficients: the method of moments and the method of quasi-dual singular functions. Estimates for the convergence of both approximate methods are proven. We eventually illustrate the theoretical results with finite element computations. The specific non-standard feature of this problem lies in the structure of its singular functions: They have the form of series whose first terms are harmonic polynomials and further terms are genuine non-smooth functions generated by the piecewise constant zeroth order term of the operator. This work has been presented in the international conference JSA 2013 [38].

6.2.16. Finite Element Subproblem Method

Participants: Patrick Dular, Christophe Geuzaine, Laurent Krähenbühl, Victor Péron, Ronan Perrussel.

In the paper [26], we develop a finite element subproblem method to correct the inaccuracies proper to perfect conductor and impedance boundary condition models, in particular near edges and corners of conductors, for a large range of conductivities and frequencies. Successive local corrections, supported by fine local meshes, can be obtained from each model to a more accurate one, allowing efficient extensions of their domains of validity. This work has been presented in the international conference Compumag 2013 [57].

We develop also a finite element subproblem method for progressive eddy current modeling. The modeling of eddy currents in conductors is split into a sequence of progressive finite element subproblems. The source fields generated by the inductors alone are calculated at first via either the Biot-Savart law or finite elements. The associated reaction fields for each added conductive region, and in return for the source regions themselves when massive, are then calculated with finite element models, possibly with initial perfect conductor and/or impedance boundary conditions to be further corrected. The resulting subproblem method allows efficient solving of parameterized analyses thanks to a proper mesh for each subproblem and the reuse of previous solutions to be locally corrected. This work has been presented in the international conference ISEF'2013 [56].

6.3. High Performance methods for solving wave equations

6.3.1. Coupling the DG code with task programming libraries

Participants: Emmanuel Agullo, Lionel Boillot, Georges Bosilca, Henri Calandra.

Last year we optimized the DG code implemented in the DIVA platform of Total by reducing the number of communications between each processors. This optimization, coupled with the use of Hybrid MPI and OpenMP parallel programming has allowed to prove the scalability of the code up to 512 cores. We are now planning to extend these tests up to 4000 cores. However, preliminary results emphasized the limitations due to low level issues such as threads placement, data communications and cache utilization. Therefore, we are now considering the implementation in DIVA of task programming libraries such as StarPU (http://runtime.bordeaux.inria.fr/StarPU/ or DAGuE (http://icl.cs.utk.edu/dague/index.html). These libraries handle the low level issues directly at the runtime level and allow the programmer to focus on the algorithm itself. They are also provide a valuable help to improve the portability of the code from one architecture to another, which will allow us to port DIVA on heterogeneous architectures such as CPU/GPU and Intel Xeon Phi. We have already coupled DIVA and DAGuE on the Symmetric Multiprocessor System (SMP) of Plafrim (https://plafrim.bordeaux.inria.fr) and compared the performance of the code with MPI and an OpenMP implementations [66], [64].

MAGNET Team

6. New Results

6.1. Probabilistic models for large graph

We have developped new approaches for the statistical analysis of large-scale undirected graphs. The main insight is to exploit the spectral decomposition of subgraph samples, and in particular their Fiedler eigenvalues, as basic features for density estimation and probabilistic inference. Our contributions are twofold. First, we develop a conditional random graph model for learning to predict links in information networks (such as scientific coauthorship and email communication). Second, we propose to apply the resulting model to graph generation and link prediction. This work is to published in the *Journal of Machine Learning Research*, the top journal in the field of machine learning.

6.2. Learning in hypergraphs

In this work, we focus on the problem of learning from several sources of heterogeneous data represented as input graphs that encode different relations over the same set of nodes. Our goal is to merge those input graphs by embedding them into an Euclidean space related to the commute time distance in the original graphs. Our algorithm output a combined kernel that can be used for different graph learning tasks. This work has been published in [5].

The approach designed in that paper has raised a new definition of undirected hypergraphs with bipartite hyperedges. A bipartite hyperedge is a pair of disjoint sets of nodes in which every node is associated with a weight. A bipartite hyperedge can be viewed as a relation between two teams of nodes in which every node has a weighted contribution to its team. Undirected hypergraphs generalize over undirected graphs. Consistently with the case of graphs, we have studied the hypergraph spectral framework. We have defined the notions of hypergraph gradient, hypergraph Laplacian, and hypergraph kernel as the Moore-Penrose pseudoinverse of a hypergraph Laplacian. Therefore, smooth labeling of (teams of) nodes and hypergraph regularization methods can be performed. Contrary to the graph case, we show that the class of hypergraph Laplacians is closed by the pseudoinverse operation (thus it is also the class of hypergraphs kernels), and is closed by convex linear combination. closure properties allow us to define (hyper)graph combinations and operations while keeping a hypergraph interpretation of the result. We exhibit a subclass of signed graphs that can be associated with hypergraphs in a constructive way. A hypergraph and its associated signed graph have the same Laplacian. This property allows us to define a distance between nodes in undirected hypergraphs as well as in the subclass of signed graphs. The distance coincides with the usual definition of commute-time distance when the equivalent signed graph turns out to be a graph. We claim that undirected hypergraphs open the way to solve new learning tasks and model new problems based on set similarity or dominance. We are currently exploring applications for modeling games between teams and for graph summarization. This work [8] has been submitted to *Journal of Machine Learning Research*.

6.3. Natural Language Processing

In [7] and [3], we develop a new algorithm for drastically improving a pairwise coreference classification system. Specifically, this algorithm works by learning the best partition over mention type pairs by training different pairwise coreference models for each pair type. In effect, our algorithm finds the optimal feature space (from a base feature set and set of types) for separating coreferential mention pairs, but it remains tractable by exploiting the structure of the hierarchies built from the pair types. In [6], we propose a new approach for the automatic identification of so-called implicit discourse relations. Our system combines hand labeled examples and automatically annotated examples (based on explicit relations) using different methods inspired by work on domain adapation. Our system is evaluated empirically and yields important performance gains compared to only using hand-labeled data. This paper has received the best paper award at the *TALN 2013* conference, the national NLP conference.

6.4. Query Induction

We have proposed a new algorithm for query learning that combines schema-guided pruning heuristics with the traditional learning algorithm for tree automata from positive and negative examples. We show that this algorithm is justified by a formal learning model, and that for stable queries it performs very well in practice of XML information extraction. This work [1] has alo been published in *JMLR*.

6.5. Learning Transducers

We have pursued the work on learning finite state tree-to-word transducers. Tree-to-word transformations are ubiquitous in computer science. They are the core of many computation paradigms from the evaluation of abstract syntactic trees to modern programming languages *XSLT*. We have extended the results obtained last year on the study of a class of sequential top-down tree-to-word transducers, called *STWs*. Transducers in *STWs* are capable of: concatenation in the output, producing arbitrary context-free languages, deleting inner nodes, and verifying that the input tree belongs to the domain even when deleting parts of it. These features are often missing in tree-to-tree transducers, and for instance, make *STWs* incomparable with the class of top-down tree-to-tree transducers. The class of *STWs* has several interesting properties, in particular we proposed in 2011 a normal for *STWs*.

In [4], we present a Myhill-Nerode characterization of the corresponding class of sequential tree-to-word transformations. Next, we investigate what learning of *STWs* means, identify fundamental obstacles, and propose a learning model with abstain. Finally, we present a polynomial learning algorithm.

MAGNOME Project-Team

6. New Results

6.1. Adopting new computing paradigms

Participants: David James Sherman [correspondant], Pascal Durrens, Natalia Golenetskaya, Florian Lajus, Xavier Calcas.

Analyses in comparative genomics are characteristically forms of datamining in high-dimension sets of relations between genes and gene products. For every linear increase in genomic data, these relations can grow at worst geometrically.

Natalia Golenetskaya's thesis[12] developed an integrated architecture that we call *Tsvetok*, which combines a novel NoSQL storage schema, domain-specific Map-Reduce algorithms, and existing resources to efficiently handle the fundamentally data-parallel analyses encountered in comparative genomics [48], [42], [51]. Tsvetok components are deployed in MAGNOME's private cloud and have been extensively tested using data and use cases derived from log analyses of the Génolevures web resource. We designed Map-Reduce solutions for the principal whole-genome analyses used by MAGNOME for comparative genomics, in particular new distributed algorithms for systematic identification of gene fusion and fission events in eukaryote genomes, and large-scale consensus clustering for protein families. These examples illustrate two strategies that can be used to scale algorithms in a Map-Reduce setting[12].

1. Converting classical graph-based algorithms with message propagation: instead of traversing a graph, which would incur high latency, information is sent forward in waves, and synchronized later. Some of the intermediate computations may be redundant, but overall running time is minimized.
2. Iterative sampling strategies, which run the standard algorithm on carefully chosen subsets, and later compute a consensus of the intermediate results. The iterations may take some time to converge, but the individual instances can be run within one machine.

Florian Lajus extended the Magus software platform to use the NoSQL storage components in Tsvetok, and has validated it on a large collection of fungal genomes. Xavier Calcas is currently integrating the Galaxy platform [8] with Magus.

6.2. Improving inference of metabolic models

Participants: David James Sherman [correspondant], Pascal Durrens, Razanne Issa, Anna Zhukova.

The Pantograph approach uses an annotated "scaffold" (reference) model and a collection of complementary predictions of homology between scaffold genes and target genes. The basis of the method is a weighing of the homology evidence to decide whether a reaction that is present in the scaffold ought be be present in the target.

We have improved on the method in two ways. First, we model the implicit knowledge represented in the boolean formula of each gene association, to derive hypotheses about the explicit role of individual genes; for example, a gene association $(S_1 \wedge S_2) \vee (S_1 \wedge S_3)$ may implicitly represent an enzyme complex formed from two subunits, the first encoded by gene S_1, and the second encoded by two paralogous genes S_2 and S_3 (figure 2). By using these hypotheses to rewrite gene associations, we improve the decision of whether a reaction is present in the target or not.

[8] http://usegalaxy.org

../../../../projets/magnome/IMG/tsvetok.png

Figure 1. General architecture of Tsvetok, showing the role of NoSQL (Apache Cassandra) and Map-Reduce (Apache Hadoop) paradigms

Second, we have adopted an abductive strategy for inferring reactions. In this strategy we consider that it is the reactions that explain the genes observed in the target genome. In the corresponding abductive logic program, the observations are the genes in the target, the integrity constraints are the rules that rewrite gene associations, and the hypotheses to be abduced are the reactions in the model. The scaffold model is compiled into a set of facts and predicates that express the reactions, their gene associations, and the integrity constraint rules; the abducibles generate assertions that specific reactions are in the target model. Combined with the facts of the genes observed in the target, this program generates, through abduction, the set of target reactions that explain the greatest number of genes.

The advantage of this approach is that it can invent, through specialization, reactions that are not present *per se* in the scaffold model.

../../../../projets/magnome/IMG/sbgn_col.png

Figure 2. An explicit model that is one possible explanation of the gene association $(S_1 \wedge S_2) \vee (S_1 \wedge S_3)$

6.3. Knowledge-based generalization of metabolic models

Participants: David James Sherman [correspondant], Pascal Durrens, Razanne Issa, Anna Zhukova.

There is an inherent tension between detail and understandability in large metabolic networks: detailed description of individual reactions is needed for simulation, but high-level views of reactions are needed for describing pathways in human terms. We defined knowledge-based methods that factor similar reactions into "generic" reactions in order to visualize a whole pathway or compartment, while maintaining the underlying model so that the user can later "drill down" to the specific reactions if need be[22], [23], [26] This method is available as a Python library from http://metamogen.gforge.inria.fr/.

Figures 3 and 4 illustrate model generation for *Yarrowia lypolitica* fatty acid oxidation in the peroxisome. Molecular species are represented as circular nodes, and the reactions as square ones, connected by edges to their reactants and products. Ubiquitous species (e.g. *oxygen*, *water*, *ATP*) are of smaller size and colored gray. Non-ubiquitous species are divided into fifteen equivalence classes and colored accordingly (red/blue for trivial species/reaction equivalence classes, different colors for non-trivial equivalence classes). The size of the model does not allow for readability of the species labels, thus we do not show them (figure 3).

The generalization algorithm identifies equivalent molecular species using an ontology, and groups together reactions that operate on the same abstract species. It finds the greatest generalization the preserves stoichiometry. The generalized model represents quotient species and reactions. For example, the violet *unsaturated FA-CoA* node is a quotient of *hexadec-2-enoyl-CoA*, *oleoyl-CoA*, *tetradecenoyl-CoA*, *trans-dec-2-enoyl-CoA*, *trans-dodec-2-enoyl-CoA*, *trans-hexacos-2-enoyl-CoA*, *trans-octadec-2-enoyl-CoA*, and *trans-tetradec-2-enoyl-CoA* (colored violet in figure 3). In a similar manner, the light-green *acCoA oxidase* quotient reaction, that converts *fatty acyl-CoA* (yellow) into *unsaturated FA-CoA* (violet), generalizes six corresponding light-green reactions of the initial model (figure 3).

The generalized model describes *β-oxidation* in a more generic way: as a transformation of *fatty acyl-CoA* (yellow) into *unsaturated FA-CoA* (violet), then into *hydroxy FA-CoA* (dark green), *3-oxo FA-CoA* (magenta), and back to *fatty acyl-CoA* (with a shorter carbon chain); while the specific model describes the same process in more details, specifying those reactions for each of the *fatty acyl-CoA* species presented in the organisms' cell (e.g. *decanoyl-CoA*, *dodecanoyl-CoA*, etc.). That is why the *beta-oxidation* chain of the reactions in the initial model, transforming step-by-step the *fatty-acyl-CoA* with the longest carbon chain into the one with the shortest chain, in the generalized model appears as a cycle (generalizing all the *fatty-acyl-CoA*s into one species, regardless the chain-length).

The specific model is appropriate for simulation, because it contains all of the precise reactions. The generalized model is suited for a human, because it reveals the main properties of the model and masks distracting details. For example, the generalized model highlights the fact that there is a particularity concerning *C24:0-CoA (stearoyl-CoA)* (red, inside the cycle): there exists a "shortcut" reaction (blue, inside the cycle), producing it directly from another *fatty acyl-CoA* (yellow), avoiding the usual four-reaction beta-oxidation chain, used for other *fatty acyl-CoA*s. This shortcut is not obvious in the specific model, because it is hidden among a plethora of similar-looking reactions.

6.4. Characterization of STAND protein families

Participants: David James Sherman, Pascal Durrens, Witold Dyrka [correspondant].

In collaboration with Sven Saupe and Mathieu Paoletti from IBGC Bordeaux (ANR Mykimun), we worked on characterization of the STAND protein family in the fungal phylum. We established an *in silico* screen based on state-of-the-art bioinformatic tools, which – starting from experimentally studied sequences from *Podospora anserina* – allowed us to determine the first systematic picture of fungal STAND protein repertoire (ms. in preparation). Most notably, we found evidence of extensive modularity of domain associations, and signs of concerted evolution within the recognition domain. Both results support the hypothesis that fungal STAND proteins, originally described in the context of vegetative incompatibility, are involved in a general fungal immune system. In addition, we investigated improved protein domain representations and elaborated a grammatical modelling method [15], which will be used to elucidate mechanisms of formation and operation of the STAND proteins.

6.5. Avoiding stiffness in BioRica

Participants: David James Sherman [correspondant], Joaquin Fernandez.

We previously formalized two strategies for integrating discrete control with continuous models, coefficient switches that control the parameters of the continuous model, and strong switches that choose different models [29], [27]. While these strategies have proved useful for modeling hybrid systems in biotechnology [31] and medicine [28], the resulting system model can be inefficient when the different subsystems evolve at very different time scales. In order to improve the efficiency of the resulting simulations, we investigated the use of Kofman's Quantized State Systems (QSS), and demonstrated that the QSS approach can be adapted to BioRica [13]. On the strength of this demonstration, we invited Joaquin Fernandez from Kofman's lab to Magnome. Joaquin had previously implemented an efficient library for QSS simulation, and during his stay succeeded in adapting it to our hybrid modeling framework. In his approach, SBML models with events are compiled into a hybrid model, using a variant of Modelica for surface syntax and using the QSS library for efficient simulation.

../../../../projets/magnome/IMG/pero_before_colored.png

Figure 3. Yarrowia lypolitica fatty acid oxidation model before generalization. Reactions of the specific model are divided into fifteen equivalence classes, represented by different colours

../../../../projets/magnome/IMG/pero_after.png

Figure 4. Generalization of the Yarrowia lypolitica fatty acid oxidation model, described as a transformation of fatty acyl-CoA (yellow) into unsaturated FA-CoA (violet), then into hydroxy FA-CoA (dark green), 3-oxo FA-CoA (magenta), and back to fatty acyl-CoA (with a shorter carbon chain)

6.6. Applications in biotechnology and health

Participants: David James Sherman, Pascal Durrens [correspondant], Florian Lajus, Xavier Calcas.

Using MAGNOME's Magus system and YAGA software, we have successfully realized a full annotation and analysis of several groups of related genomes:

- Seven new genomes, provided to the Génolevures Consortium by the CEA–Génoscope (Évry), including two distant genomes from the *Saccharomycetales* were annotated using previously published Génolevures genomes.
- Twelve wine starter yeasts linked to fermentation efficiency.
- Five pathogenic (to human) and non pathogenic Nakaseomycetes.
- Two oleaginous strains with applications in biofuels.

Winemaking yeasts. In collaboration with partners in the ISVV, Bordeaux, we have assembled and analyzed 12 wine starter yeasts, with the goal of understanding genetic determinants of performance in wine fermentation. Analysis included identification of strain-specific gains and losses of genes linked both to niche specificity and to performance in industrial applications (article in prep.). A further combined analysis with 50 natural and industrial strains showed a pattern of introgression concentrated in industrial strains (article in prep.).

Oleaginous yeasts. In collaboration with Prof Jean-Marc Nicaud's lab at the INRA Grignon, we developed the first functional genome-scale metabolic model of *Yarrowia lipolytica*, an oleaginous yeast studied experimentally for its role as a food contaminant and its use in bioremediation and cell factory applications.

Using MAGNOME's Pantograph method (see section 5.2) we produced an accurate functional model for *Y. lipolytica*, MODEL1111190000 in BioModels [9], that has been qualitatively validated against gene knockouts. This model has been enriched by Anna Zhukova with ontology terms from ChEBI and GO.

Pathogenic yeasts. A further group of five species, comprised of pathogenic and nonpathogenic species, was analyzed with the goal of identifying virulence determinants [39]. By choosing species that are highly related but which differ in the particular traits that are targeted, in this case pathogenicity, we are able to focus of the few hundred genes related to the trait [16]. The approximately 40,000 new genes from these studies were classified into existing Génolevures families as well as branch-specific families.

[9] http://biomodels.net/

MAGRIT Project-Team

6. New Results

6.1. Motion, Scene and Camera Reconstruction

Participants: Marie-Odile Berger, Srikrishna Bhat, Pierre Rolin, Gilles Simon, Frédéric Sur.

- **Metrological performance enhancement and resolution assessment for experimental solid mechanics**
  This work is motivated by image processing problems from experimental solid mechanics. One of the problem in this field is to measure heterogeneous strains on the surface of specimens subjected to mechanical tests, through an imaging device. Among full-field measurement techniques, the grid method consists in transferring a regular grid on the surface of the specimen and in taking images of the grid before and after deformation. Windowed Fourier analysis then gives an estimation of the surface displacement and strain components. In a collaboration with Institut Pascal (Université Blaise Pascal, Clermont Ferrand), we have shown that the estimations obtained by this technique are a first-order approximation of the convolution of the actual values with the analysis window. We have also characterized how the noise in the grid image impairs the displacement and strain maps [3]. This study has allowed us to improve the metrological performance of the grid method with deconvolution algorithms. A numerical and experimental study can be found in [10], [16], [21]. As any contactless measurement method, the resolution of the grid method is limited by the noise impairing the sensor. We have also characterized this resolution within a Poisson-Gaussian noise model, which is known to be realistic for CCD or CMOS sensors [4].
- **Matching in difficult conditions**
  Visual vocabularies are emerging as a new tool for building point correspondences for pose estimation. Within S. Bhat's PhD thesis [9] we have proposed several methods for visual word construction dedicated to point matching, with structure from motion and pose estimation applications in view. The three dimensional geometry of a scene is first extracted with bundle adjustment techniques based on keypoint correspondences. These correspondences are obtained by grouping the set of all SIFT descriptors from the training images into visual words. We obtain a more accurate 3D geometry than with classical image-to-image point matching. In a second on-line step, these visual words serve as 3D point descriptors that are robust to viewpoint change, and are used for building 2D-3D correspondences on-line during application, yielding the pose of the camera by solving the PnP problem. Several visual word formation techniques have been compared with respect to robustness to viewpoint change between learning and the test images.

  The PhD thesis of P. Rolin comes within the scope of camera pose estimation from an unstructured 3D point dataset, endowed with image descriptors. His work focuses on improving pose estimation with respect to strong viewpoint changes. 2D-3D correspondences are actually difficult to establish if there are too large viewpoint changes between the image whose pose is sought and the images that gave the 3D point dataset. P. Rolin currently assesses viewpoint simulation techniques in order to enhance the description of the 3D points with information from different viewpoints.
- **Acquisition of 3D calibrated data**

[3] F. Sur, M. Grédiac. Towards deconvolution to enhance the grid method for in-plane strain measurement. To appear in AIMS Inverse Problems and Imaging, American Institute of Mathematical Sciences, 2014.

[4] M. Grédiac, F. Sur. Effect of sensor noise on the resolution and spatial resolution of displacement and strain maps estimated with the grid method. To appear in Strain, Wiley, 2014.

In situ modeling is generating increasing interest in the community as it makes it possible to build AR applications in unprepared environments. In [19], we present a new method for interactive modeling of polygonal scenes, using a tablet PC, a laser rangefinder, an inertial measurement unit (IMU) and a camera. A well-founded calibration method is used to determine the orientation of the IMU and the origin and direction of the laser beam in the camera coordinate system. A new hybrid, driftless orientation tracking method is proposed, inspired by the tracking-by-synthesis algorithm adapted to 3-degree-of-freedom camera motions. Visual hints are provided during the tracking-and-modeling process in order to help the user get the best possible accuracy. These visual hints are based on a PCA analysis of the reconstructed laser point clouds and statistical measurements of the camera tracking accuracy.

6.2. Medical Imaging

Participants: René Anxionnat, Marie-Odile Berger, Nazim Haouchine, Erwan Kerrien, Matthieu Loosvelt, Pierre-Frédéric Villard, Brigitte Wrobel-Dautcourt, Ahmed Yureidini.

- **Interventional neuro radiology**
  Minimally invasive techniques impact surgery in such ways that, in particular, an imaging modality is required to maintain a visual feedback. Live X-ray imaging, called fluoroscopy, is used in interventional neuroradiology. Such images are very noisy, and cannot show but the vasculature and no other brain tissue. Most of all, and despite recent progress on the sensors, X-rays are bad for the patient's health and X-ray images are 2D projections deprived of any depth hint such as occlusions or shading. To quote a fellow physician: "it is rather uncanny to use 2D images to perform a gesture that is, by nature, 3D". Two of our long term aims in interventional neuroradiology are to reduce the operation time, and provide the interventional radiologists with a real-time visual feedback in 3D.

  All our research activity in this field is led in collaboration with the Department of Interventional Neuroradiology from Nancy University Hospital. This year was pivotal in this activity where some projects ended and other new projects started.

  We've been collaborating with Shacra Inria project-team (Lille-Nord Europe) in the context of the SOFA-InterMedS Inria Large-Scale Initiative for 4 years. Ahmed Yureidini is on the verge of defending his PhD thesis and the last step of his work consisted in validating the model he devised for the blood vasculature as a tree of local implicit surfaces [8]. Comparisons were made against simulations using triangular meshes against our implicit model and they showed a reduction by 2 orders of magnitude in computing time while numerical instabilities encountered with meshes (jaggy motions, unrealistic sticking of the catheter tip on the vessel surface, ...) were not observed with our implicit model. Publication of these results is under way.

  We also collaborate with Shacra team within the ANR IDeaS project. Computer simulations are very sensitive to inaccuracies in the various mechanical parameters or geometrical boundary conditions. Such inaccuracies are ubiquitous when dealing with patient-based data. We aim at developing Image-Driven Simulation to add the live X-ray images as new constraints to make the simulated surgical tool virtual visualization fit their position seen in the actual images. This year, a sensor was designed and tested to capture the motion of the line-shaped micro-tools (catheters, guidewires, etc...) and progress was made to design Kalman-like filters compliant with Sofa simulation platform.

  Our long-term collaboration with GE Healthcare took a new step this year with the arrival of Charlotte Delmas as a PhD student. She will work towards devising algorithms to reconstruct the micro-tools in 3D from fluoroscopy images.
- **Designing respiration models for patient based simulators**
  Respiration models are useful in many ways. They can be used in: 1) pulmonary radiotherapy, where the tumor displacement should be accurately known to be targeted by ionizing radiation, 2) thoracic surgery simulators, where breathing motion increases the realism of virtual patients, 3) interventional radiology, where augmented medical imaging that incorporates breathing motion can be used during treatment.

However building and parameterizing a fast and accurate respiration model is still an open problem. We continue this year to work on evolutionary methods to estimate the parameters of a complex 15-D respiration model on 5 patients [23]. A compound fitness function has been designed to take into account various quantities that have to be minimized.

The optimized parameters have been applied to an interventional radiology simulator that takes into account the respiration [14]. It also includes: segmentation, physically based modeling, haptics rendering, pseudo-ultrasound generation and the concept of a physical mannequin. It is the result of a close collaboration between different universities (Liverpool, Manchester, Imperial College, Banghor, Leeds, Hull) involving computer scientists, clinicians, clinical engineers and occupational psychologists.

- **Realistic simulation of organ dissection** Whilst laparoscopic surgical simulators are becoming increasingly realistic they can not, as yet, fully replicate the experience of live surgery. In particular tissue dissection is one task that is particularly challenging to replicate. Limitation of current attempts to simulate tissue dissection include: poor visual rendering; over simplification of the task and; unrealistic tissue properties. In an effort to generate a more realistic model of tissue dissection in laparoscopic surgery we worked on a novel method based on task analysis. Initially we have chosen to model only the basic geometrics of this task rather than a whole laparoscopic procedure. This year preliminary work has led to the development of a real time simulator performing organ dissection with a haptic thread at 1000Hz. 2D soft-tissue models replicate the process of tissue cutting.
- **Physics-based augmented reality**
  The development of AR systems for use in the medical field faces one major challenge: the correct superposition of pre-operative data onto intraoperative images. This task is especially difficult when laparospic surgery is considered since superposition must be achieved on deformable organs. Most existing AR systems only consider rigid registration between the pre and intraoperative data and the transformation is often computed interactively or from markers attached to the patient's body.

  In cooperation with the Shacra team, we have proposed in [17], [18] a framework for real-time augmentation of the vascular network and tumors during minimally invasive liver surgery. Internal structures computed from pre-operative CT scans can be overlaid onto the laparoscopic view for surgery guidance. Compared to state-of-the-art methods, our method uses a real-time biomechanical model to compute a volumetric displacement field from partial three-dimensional liver surface motion.

  The main contributions of this work are threefold: a) the use of a biomechanical model of liver deformation allows us to account for heterogeneity and anisotropy due to veins and arteries. In addition, the physical model is used as regularizer for the unreliable measurement of the visual tracking and as motion compensation in poorly textured areas; b) a real-time implementation of this virtual liver model has been proposed c) appropriate boundary conditions and external force have been defined which guide the biomechanical model using partial 3D motion estimated at the liver surface from a stereo video stream.

  Thanks to this framework, we are able to estimate, in real-time, relevant positions of internal structures of the liver (vessels and tumors) taking into account liver deformations and tissue heterogeneity.

MAIA Project-Team

6. New Results

6.1. Decision Making

6.1.1. Searching for Information with MDPs

Participants: Mauricio Araya, Olivier Buffet, Vincent Thomas, François Charpillet.

In the context of Mauricio Araya's PhD and PostDoc, we are working on how MDPs – or related models – can search for information. This has led to various research directions, such as extending POMDPs so as to optimize information-based rewards, or actively learning MDP models. This year has begun with the defense of Mauricio's PhD thesis in February. Since then, we have kept extending Mauricio's work and are preparing journal submissions.

While we have done some progress in this field, there are no concrete outcomes to present concerning optimistic approaches for model-based Bayesian Reinforcement Learning. Concerning POMDPs with information-based rewards, Mauricio's PhD thesis presents strong theoretical results that allow – in principle – deriving efficient algorithms from state-of-the-art "point-based" POMDP solvers. This year we have put this idea into practice, implementing variants of PBVI, PERSEUS and HSVI.

Preliminary results have been published (in French) in JFPDA'13 [32]. A journal paper with complete theoretical and empirical results is under preparation.

6.1.2. Adaptive Management with POMDPs

Participant: Olivier Buffet.

Samuel Nicol, Iadine Chadès (CSIRO), Takuya Iwamura (Stanford University) are external collaborators.

In the field of conservation biology, adaptive management is about managing a system, e.g., performing actions so as to protect some endangered species, while learning how it behaves. This is a typical reinforcement learning task that could for example be addressed through Bayesian Reinforcement Learning.

This year, we have worked in the context of bird migratory pathways, in particular the East Asian-Australasian (EAA) flyway, which is modeled as a network whose nodes are land areas where birds need to stay for some time. An issue is that these land areas are threatened due to sea level rise. The adaptive management problem at hand is that of deciding in the protection of which land areas to invest money so as to preserve the migratory pathways as efficiently as possible.

The outcome of this work is a data challenge paper published at IJCAI'13 [27], which presents the problem at hand, describes its POMDP model, gives empirical results obtained with state-of-the-art solvers, and challenges POMDP practitioners to find better solution techniques.

6.1.3. Solving decentralized stochastic control problems as continuous-state MDPs

Participants: Jilles Dibangoye, Olivier Buffet, François Charpillet.

External collaborators: Christopher Amato (MIT), Arnaud Doniec (EMD), Charles Bessonnet (Telecom Nancy), Joni Pajarinen (Aalto University).

Decentralized partially observable Markov decision processes (DEC-POMDPs) are rich models for cooperative decision-making under uncertainty, but are often intractable to solve optimally (NEXP-complete), even using efficient heuristic search algorithms. In this work, we present an efficient methodology to solving decentralized stochastic control problems formalized as a DEC-POMDP or its subclasses. This methodology is three-fold: (1) it converts the original decentralized problem into a centralized problem from the perspective of a solution method that can take advantage of the total data about the original problem that is available during the online execution phase; (2) it shows that the original and transformed problems are equivalent; (3) it solves the transformed problem using a centralized method and transfers the solution back to the original problem. We applied this methodology in various different decentralized stochastic control problems.

Our results include the application of this methodology over DEC-POMDPs [20], [33]. We recast them into deterministic continuous-state MDPs, where states — called *occupancy states* — are probability distributions over states and action-observation histories of the original DEC-POMDPs. We also demonstrate the occupancy state is a sufficient statistic for optimally solving DEC-POMDPs. We further show the optimal value function is a piecewise-linear and convex function of the occupancy states. With these results as a background, we prove for the first time that POMDP (and more generally continuous-state MDP) solution methods can, at least in principle, apply in DEC-POMDPs. This work has been presented at IJCAI'2013 [20] and (in French) at JFPDA'2013 [33], and an in-depth journal article is currently under preparation. We have already extended the results we obtained for general DEC-POMDPs in the case of transition- and observation-independent DEC-MDPs. Of particular interest, we demonstrated that the occupancy states can be further compressed into a probability distribution over the states — the first sufficient statistic in decentralized stochastic control problems that is invariant with time. This work has been presented at AAMAS'2013 [21], and an in-depth journal article is currently under preparation.

We believe our methodology lays the foundation for further work on optimal as well as approximate solution methods for decentralized stochastic control problems in particular, and stochastic control problems in general.

6.1.4. Abstraction Pathologies in Markov Decision Processes

Participants: Manel Tagorti, Bruno Scherrer, Olivier Buffet.

Jörg Hoffmann, former member of MAIA, is an external collaborator (from Saarland University).

Abstraction is a common method to compute lower bounds in classical planning, imposing an equivalence relation on the state space and deriving the lower bound from the quotient system. It is a trivial and well-known fact that refined abstractions can only improve the lower bound. Thus, when we embarked on applying the same technique in the probabilistic setting, our firm belief was to find the same behavior there. We were wrong. Indeed, there are cases where every direct refinement step (splitting one equivalence class into two) yields strictly worse bounds. We give a comprehensive account of the issues involved, for two wide-spread methods to define and use abstract MDPs.

This work has been presented and published in the ICAPS-13 workshop on Heuristics and Search for Domain-Independent Planning (HSDIP) [29] and (in French) in JFPDA-13 [37].

6.1.5. Evolutionary programming for Policies Space exploration

Participants: Amine Boumaza, Vincent Thomas.

Evolutionary Programming proposed by Fogel (initially introduced in 1966) is an approach to build an automaton optimizing a fitness function. Like other evolutionary algorithms, an initial population of automata is given, and the evolutionary programming algorithm will make this population evolve by progressively modifying automata (mutations) and keeping the most efficient ones in the next generation.

This process is close to the progressive construction by a policy iteration algorithm in a POMDP and we are currently investigating the links between these approaches.

This work has begun this year through an internship (Benjamin Bibler) and preliminary development has been made to solve the Santa Fe trail problem proposed by Koza (1992) which has become a benchmark to compare genetic and evolutionary programming approaches.

6.1.6. Evolutionary Learning of Tetris Policies

Participant: Amine Boumaza.

Learning Tetris controllers is an interesting and challenging problem due to the fact of the size of its search space where traditional machine learning methods do not work and the use of approximate methods is necessary (see 6.1.10). In this work we study the performance of a direct policy search algorithm namely the Covariance Matrix Adaptation Evolution Strategy (CMAES). We also proposed different techniques to reduce the learning time, one of which is racing. This approach concentrates the computation effort on promising policies and quickly disregards bad ones in order do reduce the computation time. This approach allowed to obtain policies of the same performance as those obtained without but at the fifth of the computation cost. The learned strategies are among the best performing players at this time scoring several millions of lines on average.

6.1.7. Evolutionary behavior learning

Participants: Amine Boumaza, François Charpillet, Iñaki Fernandèz.

Evolutionary Robotics (ER) deals with the design of agent behaviors using artificial evolution. Within this framework, the problem of learning optimal policies (or controllers) is treated as a policy search problem in the parameterized space of candidate policies. The search for the optimal policies in this context is driven by a fitness function that associates a value to the candidate policy by measuring its performance on the given task.

The work shown here describes the results of the master's thesis of Inãki Fernandèz which will be extended during a Ph.D. thesis started on october 2014.

- **Incremental policy learning with shaping**. Several methods have been proposed to accelerate the search for optimal policy in evolutionary robotics. In this work, we investigated the use of incremental learning and, more precisely, shaping, a well-known technique in behavioral psychology. The main idea is to learn to solve simple tasks and then exploit the learned behaviors to tackle increasingly harder tasks.

  Our preliminary results show that the best performances are obtained either in the setups with shaping or in the control experiment where the task difficulty is maximal. Nevertheless, a closer look at the results indicates that the best controllers for the shaping setups are not obtained at the end of the evolution, but rather at an earlier stage. This means that, for these shaping techniques, the best controllers have learned to solve the task when its difficulty was at an easy level and their performance is maintained later when the task difficulty increases. Although this was unforeseen, the results seem promising and deserve further investigation.

- **Online evolutionary learning**. As opposed to traditional evolutionary robotics which treat the learning problem as an off-line, centralized process, online onboard distributed evolutionary algorithms [67], [55] consider the learning process as executed at the agent level in a decentralized way. In this sense, each agent has its own controller or genome which is locally broadcasted from agent to agent and the best performing ones survive and spread. This gene-centered view of evolution is inspired from the theory introduced by Richard Dawkins: The selfish gene.

  The online aspect of the algorithms means that the agents are learning at the same time they are performing the task at hand. Another property that derives is that the agents are continuously learning which allows them to adapt to dynamically changing conditions and tasks. This is in opposition to the traditional view of evolutionary robotics (offline) where the outcome of evolution is tailored toward single task. Many challenging problems are raised in this framework and this thesis will address the problem of defining fitness functions that drive a swarm of agents to learn to solve a task. One other question is to study the dynamics of these algorithms both experimentally and theoretically using tools from distributed systems. Some promising work in this direction has been proposed [54].

6.1.8. Learning Bad Actions

Participant: Olivier Buffet.

Jörg Hoffmann, former member of MAIA, and Michal Krajňanský are external collaborators from Saarland University.

In classical planning, a key problem is to exploit heuristic knowledge to efficiently guide the search for a sequence of actions leading to a goal state.

In some settings, one may have the opportunity to solve multiple small instances of a problem before solving larger instances, e.g., trying to handle a logistics problem with small numbers of trucks, depots and items before moving to (much) larger numbers. Then, the small instances may allow to extract knowledge that could be reused when facing larger instances. Previous work shows that it is difficult to directly learn rules specifying which action to pick in a given situation. Instead, we look for rules telling which actions should not be considered, so as to reduce the search space. But this approach requires considering multiple questions: What are examples of bad (or non-bad) actions? How to obtain them? Which learning algorithm to use?

This research work is conducted as part of Michal Krajňanský's master of science (to be defended in early 2014). Early experiments show encouraging results, and we consider participating in the learning track of the international planning competition in 2014.

6.1.9. Complexity of the Policy Iteration algorithm

Participant: Bruno Scherrer.

We have this year improved the state-of-the-art upper bounds for the complexity of a standard algorithm for solving Markov Decision Processes: Policy Iteration.

Given a Markov Decision Process with n states and m actions per state, we study the number of iterations needed by Policy Iteration (PI) algorithms to converge to the optimal γ-discounted optimal policy. We consider two variations of PI: Howard's PI that changes the actions in all states with a positive advantage, and Simplex-PI that only changes the action in the state with maximal advantage. We show that Howard's PI terminates after at most $O\left(\frac{nm}{1-\gamma}\log\left(\frac{1}{1-\gamma}\right)\right)$ iterations, improving by a factor $O(\log n)$ a result by Hansen et al. (2013), while Simplex-PI terminates after at most $O\left(\frac{n^2m}{1-\gamma}\log\left(\frac{1}{1-\gamma}\right)\right)$ iterations, improving by a factor $O(\log n)$ a result by Ye (2011). Under some structural assumptions of the MDP, we then consider bounds that are independent of the discount factor γ: given a measure of the maximal transient time τ_t and the maximal time τ_r to revisit states in recurrent classes under all policies, we show that Simplex-PI terminates after at most $\widetilde{O}\left(n^3m^2\tau_t\tau_r\right)$ iterations. This generalizes a recent result for deterministic MDPs by Post & Ye (2012), in which $\tau_t \leq n$ and $\tau_r \leq n$. We explain why similar results seem hard to derive for Howard's PI. Finally, under the additional (restrictive) assumption that the state space is partitioned in two sets, respectively states that are transient and recurrent for all policies, we show that Simplex-PI and Howard's PI terminate after at most $\widetilde{O}(nm(\tau_t+\tau_r))$ iterations.

These results were presented at the JFPDA national workshop [36] and at the NIPS 2013 international conference [28].

6.1.10. Approximate Dynamic Programming and Application to the Game of Tetris

Participant: Bruno Scherrer.

Victor Gabillon and Mohammad Ghavamzadeh are external collaborators (from the Inria Sequel EPI). Matthieu Geist is an external collaborator (from Supélec Metz).

We present here three results: the first is a unified review of algorithms that are used to estimate a linear approximation of the value of some policy in a Markov Decision Process; the second concerns the analysis of a class of approximate dynamic algorithms for large scale Markov Decision Processes; the last is the successful application of similar dynamic programming algorithms on the Tetris domain.

In the framework of Markov Decision Processes, we have considered linear *off-policy* learning, that is the problem of learning a linear approximation of the value function of some fixed policy from one trajectory possibly generated by some other policy. We have made a review of *on-policy* learning algorithms of the literature (gradient-based and least-squares-based), adopting a unified algorithmic view. We have highlighted a systematic approach for adapting them to *off-policy* learning *with eligibility traces*. This lead to some known algorithms and suggested new extensions. This work has recently been accepted to JMLR and should be published at the beginning of 2014 [6].

We have revisited the work of Bertsekas and Ioffe (1996), that introduced λ policy iteration-a family of algorithms parametrized by a parameter λ-that generalizes the standard algorithms value and policy iteration, and has some deep connections with the temporal-difference algorithms described by Sutton and Barto (1998). We deepen the original theory developed by the authors by providing convergence rate bounds which generalize standard bounds for value iteration. We develop the theory of this algorithm when it is used in an approximate form. This work was published in JMLR [7].

Tetris is a video game that has been widely used as a benchmark for various optimization techniques including approximate dynamic programming (ADP) algorithms. A look at the literature of this game shows that while ADP algorithms that have been (almost) entirely based on approximating the value function (value function based) have performed poorly in Tetris, the methods that search directly in the space of policies by learning the policy parameters using an optimization black box, such as the cross entropy (CE) method, have achieved the best reported results. We have applied an algorithm we proposed in the past, called classification-based modified policy iteration (CBMPI), to the game of Tetris. Our experimental results show that for the first time an ADP algorithm, namely CBMPI, obtains the best results reported in the literature for Tetris in both small 10×10 and large 10×20 boards. Although the CBMPI's results are similar to those of the CE method in the large board, CBMPI uses considerably fewer (almost 1/6) samples (calls to the generative model) than CE. This work was presented at NIPS 2013 [26].

6.2. Ambiant Intelligence And Robotic Systems

6.2.1. Robotic systems : autonomy, cooperation, exploration, robustness, assistance

6.2.1.1. Local control based platooning

Participants: Jano Yazbeck, François Charpillet, Alexis Scheuer.

We consider decentralized control methods to operate autonomous vehicles at close spacings to form a platoon. We study models inspired by the flocking approach, where each vehicle computes its control from its local perceptions. We investigate different decentralized models in order to provide robust and callable solutions. Open questions concern collision avoidance, stability and multi-platoon navigation.

In order to reduce the tracking error (*i.e.* the distance between each follower's path and the path of its predecessor), we developed both an innovative approach [68] and a new lateral control law. This lateral control law reduces the tracking error faster than other existing control laws. An article, presenting this control law, its integration with a previously defined secure longitudinal control law [64] and the experimental results obtained with it, has been accepted to 2014 IEEE International Conference on Robotics and Automation.

6.2.1.2. Map Matching

Participants: François Charpillet, Maan El-Badaoui-El-Najjar.

We addressed an important issue for intelligent transportation system, namely the ability of vehicles to safely and reliably localize themselves within an a priori known road map network. For this purpose, we proposed an approach based on hybrid dynamic bayesian networks enabling to implement in a unified framework two of the most successful families of probabilistic model commonly used for localization: linear Kalman filters and Hidden Markov Models. The combination of these two models enables to manage and manipulate multi-hypotheses and multi-modality of observations characterizing Map Matching problems and it improves integrity approach. Another contribution is a chained-form state space representation of vehicle evolution which permits to deal with non-linearity of the used odometry model. Experimental results, using data from encoders' sensors, a DGPS receiver and an accurate digital roadmap, illustrate the performance of this approach, especially in ambiguous situations [8].

6.2.1.3. Adaptation of autonomous vehicle traffic to perturbations

Participants: Mohamed Tlig, Olivier Simonin, Olivier Buffet.

The aim of the European project InTraDE is to propose more efficient ways to handle containers in seaports through the use of IAVs (Intelligent Autonomous Vehicles).

In his PhD thesis, Mohamed Tlig considers the displacements of numerous such IAVs whose routes are a priori planned by a supervisor. However, in such a large and complex system, different unexpected events can arise and degrade the traffic: failure of a vehicle, human mistake while driving, obstacle on roads, local re-planning, and so on.

After working on a simple decentralized strategy to allow two queues of vehicles to share a single lane (presented in 2012, and this year in AATMO-13 [30]), we have started looking at improving vehicle flows in complete road networks. In particular, we have proposed an approach that allows multiple flows of vehicles to cross an intersection without stopping, allowing to reduce delays as well as energy consumption. Preliminary results have been presented (in French) at RJCIA-13 [38], and more advanced work is under submission.

The next step is to coordinate the controller agents located in each of the network's intersections so as to create "green waves" that would improve the flows not just locally, but globally.

6.2.1.4. Living assistant Robot

Participants: François Charpillet, Antoine Bautin, Abdallah Dib, Olivier Simonin.

With LAR (**living AssistanT Robot**), a PIA projet which started in March, Abdallah Dib joined our team for a PhD. His work is about the development of a low cost navigation system for a robot evolving in an indoor environment. The main issue of his work is to design a Simultaneous Localisation and Mapping algorithm working in a dynamic environment in which people are moving. This is very challenging if we restrict the sensing capabilities of the robot with low cost sensors such as RGB-D camera. An important service we expect the robot to achieve, is realizing similar services as the one we described below: fall detection, activity recognition.

6.2.1.5. Exploring an unknown environment with a team of mobile robots

Participants: François Charpillet, Olivier Simonin, Antoine Bautin, Nassim Kaldé.

This work has been realized during the ANR Cart-O-matic project. Antoine Bautin has been hired by the Maia team for this project for a PhD. The main objective of the project was to design and build a multi-robot system able to autonomously map an unknown building. This work has been done in the framework of a French robotics contest called Defi CAROTTE organized by the General Delegation for Armaments (DGA) and the French National Research Agency (ANR). The scientific issues of this project deal with Simultaneous Localization And Mapping (SLAM), multi-robot collaboration and object recognition. The Maia Team has been mainly involved in multi-robot collaboration and navigation [13], [1], [11].

Nassim Kaldé, a new PhD student started last year in order to carry out the work done by Antoine Bautin. The new directtion aims at addressing similar problems as the one we addressed in Cart-O-matic project but with dynamical environment, i.e. environment in which people are evolving with robots. An other point that Nassim Kaldé will address is social navigation, which is important for robot and human to coexist in a smart manner.

6.2.1.6. Features extraction for the control of redundant system with continuous sensori-motor space

Participants: Alain Dutech, Thomas Moinel.

Yann Boniface (CORTEX Team, Loria) is an external collaborator

In collaboration with the CORTEX team and supported by a M2R internship, many questions related to learning the control of a complex (mono)-agent system with a continuous sensori-motor space are explored. For several reasons, the classical framework of Reinforcement Learning is not easily used in that context:

- the value function to be learned has to be encoded using features that are not known at start,
- because of the richness of the sensori-motor space, a random exploration scheme is unlikely to find the rewarded states that are needed by the learning process,
- exploiting what is learned is difficult as one would need to find the maximum of the value function *while* it is learned.

../../../../projets/maia/IMG/bras6muscles_eng.png

Figure 1. Model of 2-joint human arm with 6 muscles.

Our work is focused on a planar model of the human arm with 2 joints and 6 muscles (see figure 1). Control signals are the activity of the motor-neurons that alter the length of the muscles, and thus the forces applied on the joints. This system is redundant but also highly non linear as many aspects of the model are described by non-linear differential equations (our model is a slight improvement over the one of Li [59]). The task to learn is to reach different positions from given starting points.

We have studied a developmental learning process with a simple muscle activation pattern. The idea is to start the learning process in an artificially reduced sensori-motor space (using rough perception and motor capacities) and slowly increase the size and complexity of this space when interesting behaviors are learned. Our approach [60] gives results comparable to other developmental techniques and raises several important research questions. Our work showed that we need an abstraction mechanism in order to define or refine the features used in actions but also in perceptions. This is a very difficult challenge that is one of the keys to the understanding (and design) of cognition. There is also a need for stronger generalization capabilities in the function approximation used in the process.

In parallel, we are taking inspiration from the field of neurosciences, and particularly on the coupling between the cortex and the cerebellum in motor control. Models based on the work of Kaladjian [58] should help us understand what control signals are used by the brain apparatus and how the learning of gestures is organized between these two regions. Our long term goal is to design mechanisms for learning features abstraction in the sensori-motor space while being guided by the improvement in behavior performances.

6.2.2. Ambiant intelligence

6.2.2.1. Personnaly Assisted Living

Participants: François Charpillet, Amandine Dubois, Olivier Simonin.

This action is supported by the Inria IPL Personally Assisted Living (PAL) which gathers 9 Inria teams associated with 6 research partners (technological, medical or social) which work together on three main issue guidelines: mobility assistance, assessing the degree of frailty of the persons, home activities analysis. The MAIA team is currently mainly involved in the 2 later topics, plus fall detection.

- **Evaluation of the degree of frailty of the elderly.** As argued in the famous paper of Fried *et al* [56] the estimation of frailty is highly significant to evaluate the risk of falls, disability, hospitalization and mortality. This issue is considered in Maia Team with different sensing devices: single RGB-D cameras [34], network of RGB-D cameras, sensing intelligent floor. One simple idea which is currently developed in the team is to determine either the center of mass of a person using one or several kinects, or the center of pressure and footsteps localization using an intelligent floor. The idea is to induce from these simple measures, the walking speed, the length of the steps and the position of the monitored persons.
- **People activity analysis.** The follow-up of the activity of elderly people over long period of time can be a good indicator of their well-being, but the evalution of the behavior of a person at home is an open challenge.

  To address this issue, we proposed this year a HMM based model capable of following simple activities such as sitting, walking, etc. An evaluation of this model has been conducted within a real smart environment with 26 subjects which were performing any of eight activities (sitting, walking, going up, squatting, lying on a couch, falling, bending and lying down). Seven out of these eight activities were correctly detected among which falling which was detected without false positives [24].
- **Fall detection.** Elderly fall is one of the major health issues affecting elderly people, especially at home. One of the objectives of the PhD work of Amandine Dubois is to design an automatic system to detect fall at home, which in its final version will be made up of a network of RGB-D sensors. A simple and robust method based on the identification and tracking of the center of mass of people evolving in an indoor environment has been developed. Using a simple Hidden Markov Model whose observations are the position of the center of mass, its velocity and the general shape of the body, we

can surprisingly monitor the activity of a person with high accuracy and thus detect falls with very good accuracy without false positives [22], [23]. An experimental study, that is reported here, has been driven in our smart apartment lab. 26 subjects were asked to perform a predefined scenario in which they realized a set of eight postures. 2 hours of video (216 000 frames) were recorded for the evaluation, half of it being used for the training of the model. The system detected the falls without false positives. This result encourages us to use this system in real situation for a better study of its efficiency.

6.2.2.2. Interconnected intelligent tiles

Participants: Mihai Andries, François Charpillet, Olivier Simonin.

We are also involved in the development of a new innovative sensing device: a Pressure-Sensing Floor with LED lighting making possible to provide a new way for people to interact with their environment. Sensitive or intelligent floors have attracted a lot of attention during the last two decades for different applications going from interaction capture in immersive virtual environments to robotics or human tracking, fall detection or activity recognition. Different technologies have been proposed so far either based on optical fiber sensing, pressure sensing or electrical near field. In the Maia Team, we have developed a more sophisticate approach in which both computation and sensing is distributed within the floor. This floor is made up of interconnected intelligent titles with can communicate with each other, have internal computation power, sense the environment activity (through four weight sensors, an accelerometer and a magnetometer) and can interact with users, robots or other sensor networks either by wireless/wire communication or through visual communication (each tile being equipped with 16 leds).

Several scientific challenges are open to us in the fields of decentralized spatial computing and in designing real application for assisting people suffering from loss of autonomy.

Some of these issues have been addressed this year. Mihai Andries, a PhD student, proposed two contributions demonstrating the relevancy of an intelligent floor such as the one we have developed. First contribution is about controlling a mobile robot through its interactions throughout the floor [10]. The second, least developed is about activity recognition of a person through its physical interaction on the floor. This approach has an important advantage compared to video based activity recognition: the privacy of people is without any doubt guaranteed. Let us mention too, the work of an internship student who developed a gait evaluation algorithm using the variation over time of the center of pressure that is sensed by the floor when one or several person walk over the floor.

6.2.2.3. Multi-Camera Tracking in Partially Observable Environment

Participants: Arsène Fansi Tchango, Olivier Buffet, Vincent Thomas, Alain Dutech.

Fabien Flacher (Thales THERESIS) is an external collaborator.

In collaboration with Thales ThereSIS - SE&SIM Team (Synthetic Environment & Simulation), we focus on the problem of following the trajectories of several persons with the help of several controllable cameras. This problem is difficult since the set of cameras cannot cover simultaneously the whole environment, since some persons can be hidden by obstacles or by other persons, and since the behavior of each person is governed by internal variables which can only be inferred (such as his motivation or his hunger).

The approach we are working on is based on (1) POMDP formalisms to represent the state of the system (person and their internal states) and possible actions for the cameras, (2) a simulator provided and developed by Thales ThereSIS and (3) particle filtering approaches based on this simulator.

From a theoretical point of view, we are currently investigating how to use a deterministic simulator and to generate new particles in order to keep a good approximation of the posterior distribution.

6.3. Understanding and mastering complex systems

6.3.1. Robustness of Cellular Automata and Reactive Multi-Agent Systems

Participants: Olivier Bouré, Vincent Chevrier, Nazim Fatès.

Our research on emergent collective behavior focuses on the analysis of the robustness of discrete models of complex systems. We ask to which extent systems may resist to various perturbations in their definitions. We progressed in the knowledge of how to tackle this issue in the case of cellular automata (CA) and multi-agent systems (MAS).

We proposed new definitions of asynchronism in lattice-gas cellular automata [3]. An experimental work was carried out and it was shown that the observation of an asynchronous version of a discrete model of swarm formation could help us gain insight on this well-studied model. The PhD thesis of O. Bouré [2] provides a detailed view of this work.

A study on the density classification problem, a well-studied problem of consensus in cellular automata, was carried out for infinite systems in 1D and 2D and for infinite trees [5], [4]. Positive results were provided and important conjectures were raised.

We proposed a survey on asynchronous cellular automata [25] and explained some of the difficulties in the classification of these objects [9].

In collaboration with colleagues from India, we proposed a complete characterisation of the reversibility of the set of the 256 Elementary Cellular Automata, which are known to be diffcult to study in all generality [53]. We also proposed a mathematical analysis of the second-order phase transitions that are observed in the most simple asynchronous cellular automata [48]. We also coordinated a special issue on asynchronous cellular automata in the *Natural Computing* journal [41].

6.3.2. Adaptive control of a complex system based on its multi-agent model

Participants: Vincent Chevrier, Tomas Navarrete.

Laurent Ciarletta (Madynes team, LORIA) is an external collaborator.

Complex systems are present everywhere in our environment: internet, electricity distribution networks, transport networks. These systems have as characteristics: a large number of autonomous entities, dynamic structures, different time and space scales and emergent phenomena. The thesis work of Tomas Navarrete is centered on the problem of control of such systems. The problem is defined as the need to determine, based on a partial perception of the system state, which actions to execute in order to avoid or favor certain global states of the system. This problem comprises several difficult questions: how to evaluate the impact at the global level of actions applied at a global level, how to model the dynamics of a heterogeneous system (different behaviors arise from different levels of interactions), how to evaluate the quality of the estimations obtained trhough the modeling of the system dynamics.

We propose a control architecture based on an "equation-free" approach. We use a multi-agent model to evaluate the global impact of local control actions before applying the most pertinent set of actions.

Our architecture has been prototypically implemented in order to confront the basic ideas of the architecture within the context of simulated "free-riding" phenomenon in peer to peer file exchange networks. We have demonstrated that our approach allows to drive the system to a state where most peers share files, even when the initial conditions are supposed to drive the system to a state where no peer shares. We have also performed experiments with different configurations of the architecture to identify the different means to improve the performance of the architecture.

This work helped us to better identify [16] the key questions that rise when using the multi-agent paradigm in the context of control of complex systems, concerning the relationship between the model entities and the target system entities.

6.3.3. Multi-Modeling and multi-simulation

Participants: Vincent Chevrier, Christine Bourjot, Benjamin Camus, Julien Vaubourg.

Laurent Ciarletta and Yannick Presse (Madynes team, LORIA) are external collaborators.

Laurent Ciarletta is the co-advisor of the thesis of Julien Vaubourg.

Models of Complex systems generally require different points of view (abstraction levels) at the same time in order to capture and to understand all the dynamics and the complexity. Consisting of different interacting parts, a model of a complex system also requires the joint and simultaneous use of modeling and simulation tools from different scientific fields.

We proposed the AA4MM meta-model [65] that solves the core challenges of multi-modelling and simulation coupling in an homogeneous perspective. In AA4MM, we chose a multi-agent point of view: a multi-model is a society of models; each model corresponds to an agent and coupling relationships correspond to interaction between agents.

This year we have made progress in the definition of multi-level modeling [15], [42]. We identified several facets of multi-level modeling and implemented them as different kinds of interactions in the AA4MM framework. A demonstration of these different multi-level couplings has been developed on a collective motion phenomenon.

In February started the MS4SG projet which involes MAIA, Madynes and EDF R&D on smart-grid simulation. A Phd thesis started on october 2013 by Julien Vaubourg in the MAIA team on the confrontation of the AA4MM principles against the specificities of smart-grid domain as a kind of complex system.

MANAO Team

5. New Results

5.1. Axis 1:Analysis and Simulation

5.1.1. Second Order Analysis of Variance in Multiple Importance Sampling

Participants: H. Lu, R. Pacanowski, X. Granier

Monte Carlo Techniques are widely used in Computer Graphics to generate realistic images. Multiple Importance Sampling reduces the impact of choosing a dedicated strategy by balancing the number of samples between different strategies. However, an automatic choice of the optimal balancing remains a difficult problem. Without any scene characteristics knowledge, the default choice is to select the same number of samples from different strategies and to use them with heuristic techniques (e.g., balance, power or maximum). We introduced [16] a second-order approximation of variance for balance heuristic. Based on this approximation, we automatically distribute samples for direct lighting without any prior knowledge of the scene characteristics. For all our test scenes (with different types of materials, light sources and visibility complexity), our method actually reduces variance in average (see Figure 9). This approach will help developing new balancing strategies.

Low to high glossy materials with five diffuse area light sources   Glossy materials with high-frequency environment map lighting

*Figure 9. Our per-pixel second-order approximation of the variance leads to a new and automatic approach for balancing the number of samples between two different sampling strategies. Except for light sources, the inset images show the sample distribution for each pixel. The yellow corresponds to the default balance heuristic strategy [101]. Compared to the balance heuristic, the variance is reduced by (**Left**) 26% and (**Right**) 20% in average (14% and 11% for the standard deviation).*

5.1.2. Rational BRDF

Participants: R. Pacanowski, L. Belcour, X. Granier

Over the last two decades, much effort has been devoted to accurately measuring Bidirectional Reflectance Distribution Functions (BRDFs) of real-world materials and to use efficiently the resulting data for rendering. Because of their large size, it is difficult to use directly measured BRDFs for real-time applications, and fitting the most sophisticated analytical BRDF models is still a complex task.

We have presented Rational BRDF [21], a general-purpose and efficient representation for arbitrary BRDFs, based on Rational Functions (RFs). Using an adapted parametrization, Rational BRDFs offer 1) a more compact and efficient representation using low-degree RFs, 2) an accurate fitting of measured materials with guaranteed control of the residual error, and 3) efficient importance sampling by applying the same fitting process to determine the inverse of the Cumulative Distribution Function (CDF) generated from the BRDF for use in Monte-Carlo rendering.

5.1.3. Decomposing intensity gradients into information about shape and material

Participants: P. Barla, G. Guennebaud, X. Granier

Recent work has shown that the perception of 3D shapes, material properties and illumination are interdependent, although for practical reasons, each set of experiments has probed these three causal factors independently. Most of these studies share a common observation though: that variations in image intensity (both their magnitude and direction) play a central role in estimating the physical properties of objects and illumination. Our aim is to separate retinal image intensity gradients into contributions of different shape and material properties, through a theoretical analysis of image formation [11].

We find that gradients can be understood as the sum of three terms: variations of surface depth conveyed through surface-varying reflectance and near-field illumination effects (shadows and inter-reflections); variations of surface orientation conveyed through reflections and far-field lighting effects; and variations of surface micro-structures conveyed through anisotropic reflections. We believe our image gradient decomposition constitutes a solid and novel basis for perceptual inquiry. We first illustrate each of these terms with synthetic 3D scenes rendered with global illumination. We then show that it is possible to mimic the visual appearance of shading and reflections directly in the image, by distorting patterns in 2D. Finally, we discuss the consistency of our mathematical relations with observations drawn by recent perceptual experiments, including the perception of shape from specular reflections and texture. In particular, we show that the analysis can correctly predict certain specific illusions of both shape and material.

5.2. Axis 2: From Acquisition to Display

5.2.1. Interactive Spatial Augmented Reality

Participants: B. Ridel, P. Reuter, X. Granier

We propose the revealing flashlight [26], a new interaction and visualization technique in spatial augmented reality that helps to reveal the details of cultural heritage artifacts (see Figure 3), since they often contain details that are difficult to distinguish due to aging effects such as erosion. We locally and interactively augment a physical artifact by projecting an expressive 3D visualization that highlights its features, based on an analysis of its previously acquired geometry at multiple scales.

Our novel interaction technique simulates and improves the behavior of a flashlight: according to 6-degree-of-freedom input, we adjust the numerous parameters involved in the expressive visualization - in addition to specifying the location to be augmented. This makes advanced 3D analysis accessible to the greater public with an everyday gesture, by naturally combining the inspection of the real object and the virtual object in a co-located interaction and visualization space. The revealing flashlight can be used by archeologists, for example, to help decipher inscriptions in eroded stones, or by museums to let visitors interactively discover the geometric details and meta-information of cultural artifacts. We confirm its effectiveness, ease-of-use and ease-of-learning in an initial preliminary user study and by the feedbacks of two public exhibitions.

5.2.2. High Dynamic Range, Multispectral, Polarization, and Light-Field Imaging

Participants: A. Manakov, J. Restrepo, R. Hegedüs, I. Ihrke

../../../../projets/manao/IMG/giana.png

Figure 10. (Top) Novel optical converter module that can be placed between a camera and its lens. This module can be configured flexibly to allow for multi-spectral, polarization, high-dynamic range, or light field imaging. It works by splitting the original image into a number of copies that can be optical filtered separately (Bottom, Left). Computational post-processing allows for unprecedented flexibility in image post-processing such as post-capture control of illumination, polarization state, exposure setting or focus. (Bottom,Right) The basis of the operational principle of the aforementioned prototype is imaging in mirror systems.

In [5] we propose a non-permanent add-on that enables plenoptic imaging with standard cameras (see also Figure 10 top and left). Our design is based on a physical copying mechanism that multiplies a sensor image into a number of identical copies that still carry the plenoptic information of interest. Via different optical filters, we can then recover the desired information. A minor modification of the design also allows for aperture sub-sampling and, hence, light-field imaging. As the filters in our design are exchangeable, a reconfiguration for different imaging purposes is possible. We show in a prototype setup that high dynamic range, multispectral, polarization, and light-field imaging can be achieved with our design.

5.2.3. Structure of a Planar Mirror System from Multiple Observations of a Single Point

Participants: I. Reshetouski, A. Manakov, I. Ihrke

We have investigated the problem of identifying the position of a viewer inside a room of planar mirrors with unknown geometry in conjunction with the room's shape parameters [25] (see also Figure 10 bottom right). We consider the observations to consist of angularly resolved depth measurements of a single scene point that is being observed via many multi-bounce interactions with the specular room geometry. Applications of this problem statement include areas such as calibration, acoustic echo cancelation and time-of-flight imaging. We theoretically analyze the problem and derive sufficient conditions for a combination of convex room geometry, observer, and scene point to be reconstructable. The resulting constructive algorithm is exponential in nature and, therefore, not directly applicable to practical scenarios.

To counter the situation, we propose theoretically devised geometric constraints that enable an efficient pruning of the solution space and develop a heuristic randomized search algorithm that uses these constraints to obtain an effective solution. We demonstrate the effectiveness of our algorithm on extensive simulations as well as in a challenging real-world calibration scenario.

5.2.4. Mirrors in Computer Graphics, Computer Vision and Time-of-Flight Imaging

Participants: I. Reshetouski, I. Ihrke

We have investigated the state of the art in dealing with the geometry of mirror systems [28].

Mirroring is one of the fundamental light/surface interactions occurring in the real world. Surfaces often cause specular reflection, making it necessary to design robust geometry recovery algorithms for many practical situations. In these applications the specular nature of the surface is a challenge. On the other side, mirrors, with their unique reflective properties, can be used to improve our sensing modalities, enabling applications such as surround, stereo and light field imaging. In these scenarios the specular interactions are highly desirable. Both of these aspects, the utilization and circumvention of mirrors are present in a significant amount of publications in different scientific areas. These publications are covering a large number of different problem statements as well as many different approaches to solutions. In this survey we focused on a collection and classification of the work in this area.

5.2.5. Computational Fabrication and Display of Material Appearance

Participant: I. Ihrke

We have investigated the state of the art in digital material fabrication and active display technology [22].

After decades of research on digital representations of material and object appearance, computer graphics has more recently turned to the problem of creating physical artifacts with controllable appearance characteristics.

While this work has mostly progressed in two parallel streams – display technologies as well as novel fabrication processes – we believe there is a large overlap and the potential for synergies between these two approaches. In this report, we summarize research efforts from the worlds of fabrication display, and categorize the different approaches into a common taxonomy. We believe that this report can serve as a basis for systematic exploration of the design space in future research.

5.3. Axis 3: Rendering, Visualization and Illustration

5.3.1. Real-Time Sampling from Captured Environment Map

Participants: H. Lu, R. Pacanowski, X. Granier

../../../../projets/manao/IMG/D0a-new.png ../../../../projets/manao/IMG/D1a-new.png

Figure 11. Time-varying light samples distribution for one pixel (cyan dot) on the dragon model when lit with a dynamic environment map [95]. This example runs in average at 145 fps using Multiple Importance Sampling with 50 samples for the Lafortune energy conserving Phong BRDF with a shininess exponent set to 150.

We have introduced [23] a simple and effective technique for light-based importance sampling of dynamic environment maps based on the formalism of Multiple Importance Sampling (MIS). The core idea is to balance per pixel the number of samples selected on each cube map face according to a quick and conservative evaluation of the lighting contribution: this increases the number of effective samples. In order to be suitable for dynamically generated or captured HDR environment maps, everything is computed on-line for each frame without any global preprocessing. Our MIS formalism can be easily extended to other strategies such as BRDF importance sampling.

5.3.2. Screen-Space Curvature for Production-Quality Rendering and Compositing

Participants: N. Mellado, P. Barla, G. Guennebaud, P. Reuter

Curvature is commonly employed for enhancing details in textured 3D models, or to modulate shading at the rendering or compositing stage. However, existing methods that compute curvature in object space rely on mesh-based surfaces and work at the vertex level. Consequently, they are not well adapted to production-quality models that rely on either subdivision surfaces with displacement and bump maps, or on implicit and procedural representations. In practice they would require a view-dependent scene discretization at each frame, to adapt geometry to visible details and avoid aliasing artifacts. Our approach [24] is independent of both scene complexity and the choice of surface representations since it computes mean curvature from scratch at each frame in screen-space. It works without any pre-process and provides a controllable screen-space scale parameter, which makes it ideal for production requirements, either during rendering or compositing.

5.3.3. Smooth Surface Contours with Accurate Topology

Participant: P. Bénard

Computing the visible contours of a smooth 3D surface is a surprisingly difficult problem, and previous methods are prone to topological errors, such as gaps in the outline. Our approach [13] is to generate, for each viewpoint, a new triangle mesh with contours that are topologically-equivalent and geometrically close to those of the original smooth surface. The contours of the mesh can then be rendered with exact visibility. The core of the approach is Contour-Consistency, a way to prove topological equivalence between the contours of two surfaces. Producing a surface tessellation that satisfies this property is itself challenging; to this end, we introduce a type of triangle that ensures consistency at the contour. We then introduce an iterative mesh generation procedure, based on these ideas. This procedure does not fully guarantee consistency, but errors are not noticeable in our experiments. Our algorithm can operate on any smooth input surface representation; we use Catmull-Clark subdivision surfaces in our implementation. We demonstrate results computing contours of complex 3D objects, on which our method eliminates the contour artifacts of other methods.

5.4. Axis 4: Editing and Modeling

5.4.1. Implicit Skinning and Modeling

Participant: G. Guennebaud

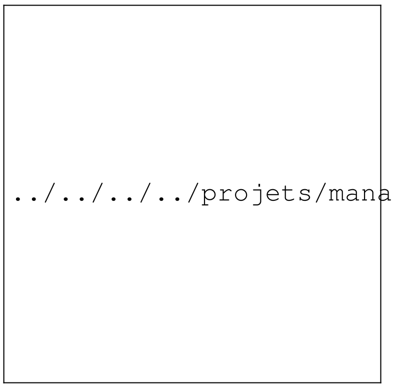
Geometric skinning techniques, such as smooth blending or dual-quaternions, are very popular in the industry for their high performances, but fail to mimic realistic deformations. Other methods make use of physical simulation or control volume to better capture the skin behavior, yet they cannot deliver real-time feedback. In collaboration with IRIT (Toulouse) and the Imagine team (Grenoble), we developed the first purely geometric method handling skin contact effects and muscular bulges in real-time. Our insight is to exploit the advanced composition mechanism of volumetric, implicit representations for correcting the results of geometric skinning techniques (cf. Figure 12 -a). The mesh is first approximated by a set of implicit surfaces. At each animation step, these surfaces are combined in real-time and used to adjust the position of mesh vertices, starting from their smooth skinning position. This deformation step is done without any loss of detail and seamlessly handles contacts between skin parts. As it acts as a post-process, our method fits well into the standard animation pipeline. Moreover, it requires no intensive computation step such as collision detection, and therefore provides real-time performances. This work has been published at Siggraph this year [20] and featured by the 3DVF website http://www.3dvf.com/actualite-6678-siggraph-2013-methode-skinning-implicite.html.

Still in collaboration with IRIT, we addressed the challenging problem of finding adequate bounds for implicit modeling with compact field functions. Recent advances in implicit surface modeling now provide highly controllable blending effects. These effects rely on the field functions of $\mathbb{R}^3 \to \mathbb{R}$ in which the implicit surfaces are defined. In these fields, there is an outside part in which blending is defined and an inside part. The implicit surface is the interface between these two parts. As recent operators often focus on blending, most efforts have been made on the outer part of field functions and little attention has been paid on the inner part. Yet, the inner fields are important as soon as difference and intersection operators are used. This makes its quality as crucial as the quality of the outside.

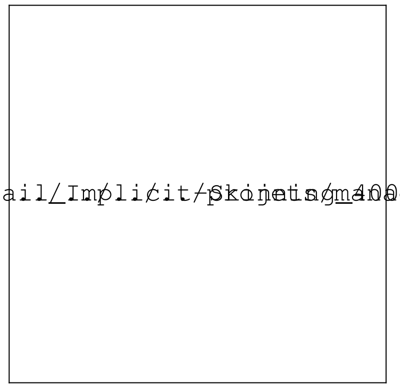
In this work we analyzed these shortcomings, and deduced new constraints on field functions such that differences and intersections can be seamlessly applied without introducing discontinuities or field distortions. In particular, we showed how to adapt state of the art gradient-based union and blending operators to our new constraints. Our approach enables a precise control of the shape of both the inner or outer field boundaries. We also developed a new set of asymmetric operators tailored for the modeling of fine details while preserving the integrity of the resulting fields. This work has been published at Shape Modeling International 2013 [14].

5.4.2. Surface reconstruction

Participants: J. Chen, G. Guennebaud, P. Barla, X. Granier



(a)



(b)

Figure 12. (a) Illustration of the implicit skinning technique. (b) Surface reconstruction from non-oriented normals

Reconstructing a smooth surface from a set of points is still a challenging problem. Most of the popular techniques assume correctly oriented points as inputs. However, in many situations, computing a consistent orientation of the normal field is as difficult as the reconstruction itself. In a recent work, we extended the Algebraic Point Set Surface method to support non oriented normals (cf. Figure 12 -b). By fitting algebraic spheres, our approach outperforms simple local methods based on non-oriented planar fit while still being fast since it involves only local computations. The core of this new technique also proved to be useful for image processing. This work as been published at Computer Graphics Forum [3].

5.4.3. Manipulation of Anisotropic Highlights

Participants: B. Raymond, P. Barla, G. Guennebaud, X. Granier

We have developed [19] a system for the direct editing of highlights produced by anisotropic BRDFs, which we call anisotropic highlights. We first provide a comprehensive analysis of the link between the direction of anisotropy and the shape of highlight curves for arbitrary object surfaces. The gained insights provide the required ingredients to infer BRDF orientations from a prescribed highlight tangent field. This amounts to a non-linear optimization problem, which is solved at interactive framerates during manipulation. Taking inspiration from sculpting software, we provide tools that give the impression of manipulating highlight curves while actually modifying their tangents. Our solver produces desired highlight shapes for a host of lighting environments and anisotropic BRDFs.

MARELLE Project-Team

6. New Results

6.1. Bourbaki, Sets and Ordinals

Participant: José Grimm [correspondant].

In previous years, we developped a formal library describing the part of the Bourbaki books on set theory, cardinals and ordinals, [18]. Here are ome additions to the library.

Since addition of ordinals is non-commutative, the sum of n ordinals x_1 to x_n depends on their ordering; the maximum number $f(n)$ is a priori bounded by $n!$, and we have shown that it satisfies a recurrence relation (R), Bourbaki asks, in an exercise, to show that $f(n) = 81f(n-5)$ for $n \geq 20$. This is an easy consequence of an explicit formula (F) for f. That (R) implies (F) can be expressed in pure Coq (with binary integers), but we have no idea how to prove it.

We proved some facts of the theory of models: the set V_ω of hereditarily finite sets satisfies ZF (but not the axiom of infinity); the von Neumann universe satisfies ZF and AF, there is a subset of the universe satisfiying ZF containing no inaccessible cardinal. We have also studied the set of formulas and show the theorem of Lövenheim-Skolem.

The main contribution this year is the study of some families of ordinals. If the family is internally closed and too big to be a set, then it is the image of a normal (continuous and strictly increasing) function, called the enumeration function of the family. The family of fix-points of a normal function satisfies this property, and the enumeration of this family is called the first derived function. There is a derivation at every order. For instance, the first derivation of $x \mapsto 1+x$ is $x \mapsto \omega x$, and the derivation of order n is $x \mapsto \phi(n,x)$. The least x such that $x = \omega^x$ is known as ϵ_0; the least x such that $x = \phi(x,0)$ is known as Γ_0.

We have shown that the inductive type T defined by zero and a constructor of type $T \to N \to T \to T$, without the terms that are not in "normal form" , is isomorphic to the set of ordinals less than ϵ_0; in the case of $T \to T \to N \to T \to T$, we get all ordinals less than Γ_0; we have also studied the case with one more T (the first two types were first implemented by Castéran, the last was suggested by Ackermann) [19]

6.2. Homotopy Type Theory

Participants: Yves Bertot [correspondant], Florent Bréhard.

Homotopy Type Theory is a domain born out of the conjuction of type theory, which serves as foundations for proof systems like Coq or Agda, and homotopy theory, and domain of mathematics which is concerned with equivalence classes of objects modulo continuous deformation. In particular, Homotopy Type Theory concentrates on paths (continuous substrate between various objects) and paths between paths: paths between points can be understood as lines, paths between lines can be understood as surfaces.

In particular, paths can be thought has having the same properties as the notion of equality that is usually defined inductively in type theory systems and homotopy type theory goes against the trend started in the 1990s where specialists thought an axiom should be added to express that all paths between paths should be equal. On the contrary, if all paths between paths are not equal, type theory can be used to model homotopy theory and that domain of mathematics because a new area of applications for type theory-based theorem provers.

V. Voevodsky organized a special year at Institute for Advanced Study in Princeton on this topic, and Yves Bertot participated to this special year, during which many experiments were performed, extensions to proof systems were designed, and a book was produced. In particular, Yves Bertot devised an extension of the Coq system with *private types* which makes it possible to simulate a new concept known as *higher inductive types*. On top of this extension, the members of the special year produced a collection of higher inductive types, describing circles, spheres, truncations.

During his internship in the Marelle project, Florent Bréhard studied the equivalence between several presentations of higher-dimension spheres using higher inductive types.

Work on higher inductive types was pursued more precisely by Bruno Barras from Saclay. We expect that the result of this work will supersede the experiments made possible by Yves Bertot's implementation of private types, but the concept of private type may retain applications in other domains.

6.3. Isolation of polynomial roots

Participants: Yves Bertot [correspondant], Julianna Zsidó.

Together with techniques to produce square-free polynomials (polynomials whose roots are all simple), Bernstein polynomials provide a way to decide whether a polynomial has roots in a given interval. Together with a dichotomy procedure, this makes it possible to isolate all the roots of a polynomial, or to show that no root of a given polynomial occur in a given interval. At the end of 2012, Julianna Zsidó started to study this procedure: she showed the properties of the procedure to obtain square-free polynomials and she then formalized a proof for a theorem known as the *theorem of three circles* which plays a rôle in proving that dichotomy will terminate. This work has been published as an article in the *Journal of Automated Reasoning*.

We expect to wrap up all this work by producing easy-to-use tactics to prove properties of polynomial formulas and generalizing it to polynomials in several variables.

During a summer internship, Konstantinos Lentzos worked on the representation of algebraic numbers (which can always be represented as roots of polynomials in a given interval) and the question of finding polynomials for algebraic numbers obtained through simple operations (like addition, multiplication, opposite, and inversion). However, this work was made extremely difficult by the problem of finding morphisms between various fields definable on top of a polynomial ring.

6.4. Properties of the π number

Participants: Yves Bertot [correspondant], Laurence Rideau, Laurent Théry.

As a testbed for the progress of formalized libraries in the domain of calculus, we studied an algorithm to compute π (the circle ratio) using arithmetic-geometric means. This study brought us to extend the libraries with improper integrals, studies of *arcsinh*, variable change in integrals, and error propagation proofs.

We also studied a formal proof of the spigot algorithm designed by Bailey, Borwein, and Plouffe, which is used to compute far digits in the hexadecimal representation of π as a fractional number. This relies on floating point computations and error control, for which we provided a formal proof.

6.5. Formal study of cryptography

Participants: Gilles Barthe [IMDEA Software Institute], François Dupressoir [IMDEA Software Institute], Benjamin Grégoire [correspondant], César Kunz [IMDEA Software Institute], Yassine Lakhnech [Univ. Grenoble 1], Benedikt Schmid [IMDEA Software Institute], Pierre-Yves Strub [IMDEA Software Institute], Santiago Zanella Béguelin [MSR].

The goal of this work is to provide a friendly tool easily usable by cryptographers without knowledge of formal proof assistants. The idea is to use the techniques formally proved in Certycrypt and to call SMT-provers. We provide two differents tools:

- Easycrypt (see http://www.easycrypt.info/) is a toolset for reasoning about relational properties of probabilistic computations with adversarial code. Its main application is the construction and verification of game-based cryptographic proofs. This year, Easycrypt has been fully reimplemented, allowing more modularity in proofs and an interactive prover has been integrated.
- ZooCrypt (see http://www.easycrypt.info/zoocrypt/) is an automated tool for analyzing the security of padding-based public-key encryption schemes (i.e. schemes built from trapdoor permutations and hash functions). ZooCrypt includes an experimental mechanism to generate EasyCrypt proofs of security of analyzed schemes.

This year we published papers concerning formal proofs for properties of elliptic curves, differential privacy, padding-based encryption, and probabilistic relational verification.

6.6. Approximation of Mathematical functions

Participants: Guillaume Hanrot, Érik Martin-Dorel, Micaela Mayero [Université de Paris 13], Ioana Paşca [Université de Nimes], Laurence Rideau, Laurent Théry [correspondant].

In a collaboration supported by ANR project Tamadi, we study the approximation of mathematical functions (exponential and trigonometric functions) using polynomial functions.

This year, we completed the formal verification of our library that computes Taylor Models for the usual mathematical functions of one variable within Coq. A presentation of this work has been done at SYNASC'2013.

The SLZ algorithm checks that there is no hard-to-round floating numbers for a given range in a given floating-point format. It usually consists of a very long computation returning a yes/no answer. Formally proving the implementation of the algorithm is current outside reach since it requires very sophisticated numerical libraries that are currently impossible to verify formally. We have defined a notion of certificate for these computations based on Hensel's lemma and derived an executable checker within Coq that is capable to verify such computations. A publication has been submitted.

6.7. Formal verification in Geometry

Participants: Laurent Fuchs, Laurent Théry [correspondant].

Grassmann-Cayley Algebras are a convenient algebraic way of talking about geometrical concepts. We have further improved our certified Grassmann-Cayley Algebra library to accommodate unbalanced binary trees. A publication has been accepted and will be published in 2014.

6.8. SMT automation for Ssreflect

Participants: Antoine Grospellier, Laurent Théry [correspondant].

The proof of the Feit-Thompson theorem (also known as the odd-order theorem) has been carried on with little use of automation. We have customised the existing connection between Coq and SMT solvers using Why to accomodate Ssreflect specificities. The preliminary results are encouraging.

MASAIE Project-Team

5. New Results

5.1. Release of Wolbachia as a preventive action against dengue

We have designed a model of infection by *Wolbachia* of a *Aedes aegypti* population, to take into account the biology of this infection and also the data that can be obtained. The objective is to use this model for predicting the sustainable introduction of this bacteria. We provide a complete mathematical analysis of the model proposed and give the basic reproduction ratio $\mathcal{R}_0$ for *Wolbachia* . We observe a bistability phenomenon. Two equilibria are asymptotically stable : an equilibrium where all the population is uninfected and an equilibria where all the population is infected. A third unstable equilibrium exists. We are in a backward bifurcation situation. The bistable situations occurs with natural biological values for the parameters. Our model is an example of an epidemiological model with only vertical transmission.

This infection model is then connected with a classical dengue model. We prove that for the complete model the equilibrium with *Wolbachia* for the mosquitoes and without dengue for the human is asymptotically stable. We prove that, if a sufficiently great population of infected mosquitoes is introduced, dengue will disappear.

We use the data of a real trial of releases of infected mosquitoes in Cairns (Australia) to calibrate our model. Our model behave remarkably well versus the observed data. We use then the calibrated model to simulate different scenarii of appearance of dengue. We use a pessimistic situation where the basic reproduction ratio $\mathcal{R}_0$ of dengue is 24.5. The simulations confirm our findings, dengue epidemics does not occur, and show that the introduction of *Wolbachia* is a promising way of control dengue.

../../../../projets/masaie/IMG/GordonYorkpercent-vert

Figure 1. Frequencies observed and predicted. The red squares are the frequencies of infection given by the model. The blue circles are the frequencies observed

../../../../projets/masaie/IMG/dengueUetW30.png

Figure 2. In green the infected human with Wolbachia present. In blue when mosquitoes have not been infected by Wolbachia.

5.2. Arboviruses on urban environments

We investigate the influence of *human movement* for the onset of an arboviral (mosquito-borne) epidemics (such as Dengue, Chikungunya, West Nile or Yellow fever) on an urban environment. The metapopulation model has a standard SIR (human)/SI (mosquito) model as the basic dynamics on the patches. The nodes consist of notification districts used by public health authorities. The subsystems are coupled by human movement. Our main result provides quantitative relations between three reproduction numbers: *local* - at each isolated subsystem, *uniform* or *mixing* - aggregating the data of the whole region, and the *network* reproduction number - for the coupled dynamics. We observe that the epidemics can spread among the patches as a consequence solely of human movement: while all nodes may have, if isolated, local reproduction ratio less than one and, moreover, the uniform reproduction number being also less than one, however, the network reproduction number can be greater than one. An estimate is provided on the overall effect of vector control on a chosen patch [13].

5.3. Analysis of the dynamics of some models for vector-borne diseases with host circulation

In this work we study the dynamics of a vector borne disease on a metapopulation model that accounts for host circulation. For such models, the movement network topology gives rise to a contact network topology, corresponding to a bipartite graph. Under the assumption that the contact network is strongly connected, we can define the basic reproductive number R_0 and show that this system has only two equilibria: the so called disease free equilibrium (DFE); and a unique interior equilibrium that exists if, and only if, the basic reproduction number, R_0, is greater that unity. We are also able to show that the DFE is globally asymptotically stable, if $R_0 \leq 1$. If $R_0 > 1$, the dynamics is uniformly persistent and, with further assumptions on the contact network structure, we also show that the endemic equilibrium (EE) is globally asymptotically stable [17].

5.4. Analysis and observer design for a schistosomiasis model

Human schistosomiasis is a behavioral and occupational disease associated with poor human hygiene, insanitary animal husbandry and economic activities. Among human parasitic diseases, schistosomiasis ranks second behind malaria as far as the socio-economic and public health importance in tropical and subtropical areas are concerned. The spread and persistence of schistosomiasis have made of it one of the most complex host-parasite process to model mathematically because of the different steps of growth of larvae assumed by the parasite and the requirement of two host elements (definitive human host and intermediate snail hosts) during their life cycle.

An efficient method to control the schistosomiasis infection that may require relatively little funding is a biological control. Particularly, trematode parasites or competitive snails of the intermediate snail hosts have been proved to be effective in controlling schistosomiasis in the Caribbean area.

We have studied a schistosomiasis infection model that involves human and intermediate snail hosts as well as an additional mammalian host and a competitor snail species. This mathematical analysis of the model gives insight about the epidemiological consequences of the introduction of a competitor resistant snail species [15].

We have also proposed a solution to the state estimation problem for a schistosomiasis infection dynamical model described by a continuous non linear system when only the infected human population is measured. We have constructed an estimator that is able to give dynamical estimates of the variables that can not be measured [14].

MATHRISK Project-Team

6. New Results

6.1. Credit risk

Participants: Aurélien Alfonsi, Céline Labart, Jérôme Lelong.

We have ended our study on stochastic local intensity model. We have shown by the mean of a particles system that this model is well defined and have obtained an efficient way to perform Monte-Carlo algorithms for this model.

6.2. Liquidity risk

Participants: Aurélien Alfonsi, A. Schied.

A. Alfonsi and A. Schied (Mannheim University) are working on price impact models that describe how the price is modified by large trades. The paper together with J. Acevedo on a time-dependent price impact is now accepted for publication. With A. Schied and F. Klöck [45], we have studied the cross price impact between different assets and identified conditions on the resilience of this impact that avoid manipulations strategies. With P. Blanc, we are working on the optimal execution problem when there are many large traders that modify the price.

6.3. Systemic Risk

Participants: Andreea Minca, Agnès Sulem.

We are working on the theory of the stochastic control of financial networks.

In two related articles, we find the optimal strategy of a government who seeks to make equity infusions in a banking system prone to insolvency and to bank runs. The first article combines stochastic control and the random graph representation of the financial system developed in Andreea's thesis. The second article combines the network representation of a financial system and the solvency-based mechanism of contagion with another potent source of distress, which is funding illiquidity [31] and [60].

6.4. Estimation of the parameters of a Wishart process

Participants: Aurélien Alfonsi, Ahmed Kebaier, Clément Rey.

This research has started this year together with the thesis of Clément Rey. We are studying the Maximum Likelihood Estimator for the Wishart processes and in particular its convergence in the ergodic and the non ergodic case.

6.5. An Affine term structure model for interest rates that involve Wishart diffusions

Participants: Aurélien Alfonsi, E. Palidda.

Affine term structure models (Dai and Singleton, Duffie, ...) consider vector affine diffusions. Here, we would like to extend this model by including some Wishart dynamics, and to get a model that could better fit the market. We also develop some numerical pricing methods for this model to make its implementation possible.

6.6. Applications of optimal transport

Participants: Aurélien Alfonsi, Benjamin Jourdain, Arturo Kohatsu-Higa.

A. Alfonsi and B. Jourdain study the Wasserstein distance between two probability measures in dimension n sharing the same copula C. The image of the probability measure dC by the vectors of pseudo-inverses of marginal distributions is a natural generalization of the coupling known to be optimal in dimension $n = 1$. In dimension $n > 1$, it turns out that for cost functions equal to the p-th power of the L^q norm, this coupling is optimal only when $p = q$ i.e. when the cost function may be decomposed as the sum of coordinate-wise costs.

As another application of optimal transport, they are working with A. Kohatsu-Higa on the uniform in time estimation of the Wasserstein distance between the time-marginals of an elliptic diffusion and its Euler scheme. To generalize in higher dimension the result that they obtained previously in dimension one using the optimality of the explicit inverse transform, they compute the derivative of the Wasserstein distance with respect to the time variable thanks to the theory developped by Ambrosio Gigli and Savare. The abstract properties of the optimal coupling between the time marginals then enable them to estimate this time derivative.

6.7. Capital distribution and portfolio performance in the mean-field Atlas model

Participants: Benjamin Jourdain, J. Reygner.

B. Jourdain and J. Reygner study a mean-field version of rank-based models of equity markets, introduced by Fernholz in the framework of stochastic portfolio theory. They first obtain an asymptotic description of the market when the number of companies grows to infinity. They then discuss the long-term capital distribution in this asymptotic model, as well as the performance of simple portfolio rules. In particular, they highlight the influence of the volatility structure of the model on the growth rates of portfolios.

6.8. Public Private Partnerships

Participants: Gilles Edouard Espinosa, Caroline Hillairet, Benjamin Jourdain, Monique Pontier.

With Gilles Edouard Espinosa, Caroline Hillairet and Monique Pontier, Benjamin Jourdain is interested in the problem of outsourcing the debt for a big investment, according two situations: either the firm outsources both the investment (and the associated debt) and the exploitation to a private consortium, or the firm supports the debt and the investment but outsources the exploitation. They prove the existence of Stackelberg and Nash equilibria between the firm and the private consortium, in both situations. They compare the benefits of these contracts.They conclude with a study of what happens in case of incomplete information, in the sense that the risk aversion coefficient of each partner may be unknown by the other partner [51].

6.9. Backward Stochastic (Partial) Differential equations with jumps and stochastic control

Participants: Roxana-Larisa Dumitrescu, Marie-Claire Quenez, Agnès Sulem.

We have studied optimization problems for BSDEs with jumps, optimal stopping for dynamic risk measures induced by BSDEs with jumps and associated reflected BSDEs, and generalized Dynkin games associated to double barriers reflected BSDEs with jumps [32], [38], [42]. A. Sulem, with B. Øksendal and T. Zhang has also studied optimal stopping for Stochastic Partial Differential equations and associated reflected SPDEs [34], and optimal control of Forward-Backward SDEs [54].

6.10. Utility maximization and Arbitrage Theory

Participants: Claudio Fontana, Bernt Øksendal, Agnès Sulem.

B. Øksendal and A. Sulem have contributed to the issue of robust utility maximization in jump diffusion markets via a stochastic maximum approach and the links with robust duality [53].

In the period January - October 2013, the main subject of investigation of C. Fontana has been arbitrage theory, with a special emphasis on no-arbitrage conditions weaker than the classical notion of No Free Lunch with Vanishing Risk (NFLVR). In particular, in the context of financial market models based on diffusion processes (see [35]), we have provided a characterization of several no-arbitrage conditions as well as a generalization of the second fundamental theorem of asset pricing. In the context of jump-diffusion models under partial information (see [25]), we have studied the relation between market viability (in the sense of solvability of portfolio optimization problems) and the existence of a martingale measure given by the marginal utility of terminal wealth, without a-priori assuming no-arbitrage restrictions on the model. Finally, in the paper [41], we have provided a critical analysis of the paper Arbitrage, Approximate Arbitrage and the Fundamental Theorem of Asset Pricing (Wong & Heyde, 2010), where the authors aim at proposing an original and simple proof of the fundamental theorem of asset pricing in the context of incomplete diffusion-based models. We have shown that the method of Wong & Heyde (2010) can only work in the well-known case of complete markets, exhibiting an explicit counterexample.

6.11. Regularity of probability laws using an interpolation method

Participant: Vlad Bally.

This work was motivated by previous papers of Nicolas Fournier, J. Printemps, E. Clément, A. Debusche and of myself, concerning the regularity of the law of the solutions of some equations with coefficients with little regularity - for example diffusion processes with Hölder coefficients (but also many other examples including jump type equations, Bolzmann equation or Stochastic PDE's). Since we do not have sufficient regularity the usual approach by Malliavin calculus fails in this framework. Then one may use an alternative idea which roughly speaking is the following: We approximate the law of the random variable X (the solution of the equation at hand) by a sequence $X(n)$ of random variables which are smooth and consequently we are able to establish integration by parts formulas for $X(n)$ and we are able to obtain the absolutely continuity of the law of $X(n)$ and to establish estimates for the density of the law of $X(n)$ and for its derivatives. Notice that the derivatives of the densities of $X(n)$ generally blow up - so we can not derive directly results concerning the density of the law of X. But, if the speed of convergence of $X(n)$ to X is stronger than the blow up, then we may obtain results concerning the density of the law of X. It turns out that this approach fits in the framework of interpolation spaces and that the criterion of regularity for the law of X amounts to the characterization of an interpolation space between a space of distributions and a space of smooth functions. Although the theory of interpolation spaces is very well developed and one already know to characterize the interpolation spaces for Sobolev spaces of positive and negative indices, we have not found in the (huge) literature a result which covers the problem we are concerned with. So, although our result may be viewed as an interpolation result, it is a new one. The above work is treated in the paper [62] (in collaboration with Lucia Caramellino). As an application we discussed in [48] the regularity of the law of a Wiener functional under a Hörmander type non degeneracy condition.

6.12. A stochastic parametric representation for the density of a Markov process

Participant: Vlad Bally.

Classical results in the PDE theory (due to A. Friedmann) assert that, under uniform ellipticity conditions, the law of a diffusion process has a continuous density (the approach of A. Friedmann is analytical and concerns PDE's instead of the corresponding diffusion process). The method developed by A. Friedmann becomes well known as the "parametric method". In collaboration with A. Kohatzu Higa [49] we gave a probabilistic approach which represents the probabilistic counterpart of the parametric method. We obtained a probabilistic representation for the density of the law of the solution of a SDE and more generaly, for a class of Markov processes including solutions of jump type SDE's. This representation may be considered as a perfect simulation scheme and so represents a starting point for Monte Carlo simulation. However the random variable which appears in the stochastic representation has infinite variance, so direct simulation gives unstable results (as some preliminary tests have proved). In order to obtain an efficient simulation scheme some more work on the reduction of variance has to be done.

6.13. Regularity of probability laws using an interpolation method

Participant: Vlad Bally.

The distance between two density functions and convergence in total variation. In collaboration with Lucia Caramellino we obtained estimates of the distance between the densities of the law of two random variables using an abstract variant of Malliavin calculus. We used these estimates in order to study the convergence in total variation of a sequence of random variables. This has been done in [47]. We are now working on more specific examples concerning the Central Limit Theorem. In the last years the convergence in entropy distance and in total variation distance for several variants of the CLT has been considered in papers of S. Bobkov, F. Gotze, G. Peccati, Y. Nourdin, D. Nualart and G. Polly. So this seems to be a very active research area. Moreover, in an working paper in collaboration with my Phd student R. Clement, we use the same methods in order to study the total variation distance between two Markov semigroups and in particular for approximation schemes. A special interest is devoted to higher order schemes - as for example the Victoire Nyomia scheme.

MAVERICK Project-Team

6. New Results

6.1. Visual perception

6.1.1. Decomposing intensity gradients into information about shape and material

Participants: Pascal Barla, Romain Vergne, Roland W. Fleming.

Recent work has shown that the perception of 3D shapes, material properties and illumination are interdependent, although for practical reasons, each set of experiments has probed these three causal factors independently. Most of these studies share a common observation though: that variations in image intensity (both their magnitude and direction) play a central role in estimating the physical properties of objects and illumination. Our aim is to separate retinal image intensity gradients into contributions of different shape and material properties, through a theoretical analysis of image formation. We find that gradients can be understood as the sum of three terms: variations of surface depth conveyed through surface-varying reflectance and near-field illumination effects (shadows and inter-reflections); variations of surface orientation conveyed through reflections and far-field lighting effects; and variations of surface micro-structures conveyed through anisotropic reflections. We believe our image gradient decomposition constitutes a solid and novel basis for perceptual inquiry. We first illustrate each of these terms with synthetic 3D scenes rendered with global illumination. We then show that it is possible to mimic the visual appearance of shading and reflections directly in the image, by distorting patterns in 2D. Finally, we discuss the consistency of our mathematical relations with observations drawn by recent perceptual experiments, including the perception of shape from specular reflections and texture. In particular, we show that the analysis can correctly predict certain specific illusions of both shape and material.

6.1.2. Predicting the effects of illumination in shape from shading

Participants: Roland W. Fleming, Romain Vergne, Steven Zucker.

Shading depends on different interactions between surface geometry and lighting. Under collimated illumination, shading is dominated by the 'direct' term, in which image intensities vary with the angle between surface normals and light sources. Diffuse illumination, by contrast, is dominated by 'vignetting effects' in which image intensities vary with the degree of self-occlusion (the proportion of incoming direction that each surface point 'sees'). These two types of shading thus lead to very different intensity patterns, which raises the question of whether shading inferences are based directly on image intensities. We show here that the visual system uses 2D orientation signals ('orientation fields') to estimate shape, rather than raw image intensities and an estimate of the illuminant. We rendered objects under varying illumination directions designed to maximize the effects of illumination on the image. We then passed these images through monotonic, non-linear intensity transfer functions to decouple luminance information from orientation information, thereby placing the two signals in conflict (Figure 6). In Task 1 subjects adjusted the 3D shape of match objects to report the illusory effects of changes of illumination direction on perceived shape. In Task 2 subjects reported which of a pair of points on the surface appeared nearer in depth. They also reported perceived illumination directions for all stimuli. We find that the substantial misperceptions of shape are well predicted by orientation fields, and poorly predicted by luminance-based shape from shading. For the untransformed images illumination could be estimated accurately, but not for the transformed images. Thus shape perception was, for these examples, independent of the ability to estimate the lighting. Together these findings support neurophysiological estimates of shape from the responses of orientation selective cell populations, irrespective of the illumination conditions.

6.1.3. Evaluation of Depth of Field for Depth Perception in DVR

Participants: Pascal Grosset, Charles Hansen, Georges-Pierre Bonneau.

../../../../projets/maverick/IMG/orientation-fields.jpg

Figure 6. Top: one single shape is shaded using multiple light source directions. Bottom: one single light source is used to shade multiple shapes. Each of them has approximatively the same percept as for the corresponding light direction. Similarity between orientation fields can be seen in the insets.

We study the use of Depth of Field for depth perception in Direct Volume Rendering (Figure 7). Direct Volume Rendering with Phong shading and perspective projection is used as the baseline. Depth of Field is then added to see its impact on the correct perception of ordinal depth. Accuracy and response time are used as the metrics to evaluate the usefulness of Depth of Field. The on site user study has two parts: static and dynamic. Eye tracking is used to monitor the gaze of the subjects. From our results we see that though Depth of Field does not act as a proper depth cue in all conditions, it can be used to reinforce the perception of which feature is in front of the other. The best results (high accuracy & fast response time) for correct perception of ordinal depth is when the front feature (out of the users were to choose from) is in focus and perspective projection is used. Our work has been published in the proceedings of the Pacific Graphics conference in 2013 [16].

../../../../projets/maverick/IMG/DepthOfFieldExperiment.jpg

Figure 7. A user study is conducted to evaluate the use of Depth of Field for depth perception in Direct Volume Rendering. The user has to say which of the two highlighted features is in the front. Statistical analysis of the results is performed to assess the effect of Depth of Field in the visualization. Results are detailed in the Pacific Graphics paper [16].

6.2. Visualization

6.2.1. Morse-Smale complexes

Participants: Léo Allemand-Giorgis, Georges-Pierre Bonneau, Stefanie Hahmann.

Preserving meaningful local extrema of scalar data in a visualization while removing nearby extrema with similar values is a powerful way for enhancing the appearance of significant features. For the special case of monotonic data, e.g. data with no local extrema in the interior of the domain, the visualization should not introduce spurious local extrema. We study a new piecewise polynomial interpolant that preserves the monotonicity of scalar data defined on a 2D uniform grid. Based on this interpolant, we also plan to introduce a new method for visualizing data that has been simplified according to its Morse-Smale complex, a combinatorial structure connecting the critical points and partitioning the domain into a set of monotonic regions. In contrast with previous analogous works, our approach uses piecewise polynomial functions defined in each monotonic region instead of optimizing values on the original mesh vertices. We have presented our first results in a workshop and have submitted a paper for a book chapter about our new monotonic interpolant.

6.2.2. *Computation of components' interfaces in highly complex assemblies*

Participants: François Jourdes, Georges-Pierre Bonneau, Stefanie Hahmann, François Faure.

The preparation of CAD models from complex assemblies for simulation purposes is a very time-consuming and tedious process, since many tasks such as meshing and idealization are still completed manually. Herein, the detection and extraction of geometric interfaces between components of the assembly is of central importance not only for the simulation objectives but also for all necessary shape transformations such as idealizations or detail removals. It is a repetitive task in particular when complex assemblies have to be dealt with. This paper proposes a method to rapidly and fully automatically generate a precise geometric description of interfaces in generic B-Rep CAD models. The approach combines an efficient GPU ray-casting technique commonly used in computer graphics with a graph-based curve extraction algorithm. Not only is it able to detect a large number of interfaces efficiently, but it also provides an accurate Nurbs geometry of the interfaces, that can be stored in a plain STEP file for further downstream treatment. We demonstrate our approach on examples from aeronautics and automotive industry, see Figure 8 . Our results have been funded in by the ANR Project ROMMA. They have been published as a journal paper in [13], and presented at the Solid and Physical Modeling conference in 2013.

6.3. Image creation and editing

6.3.1. *Accurate Binary Image Selection from Inaccurate User Input*

Participants: Subr Kartic, Paris Sylvain, Soler Cyril, Kautz Jan.

Selections are central to image editing, since they are the starting point of common operations such as copy-pasting and local edits. Creating them by hand is particularly tedious and scribble-based techniques have been introduced to assist the process. By interpolating a few strokes specified by users, these methods generate precise selections. However, most of the algorithms assume a 100% accurate input, and even small inaccuracies in the scribbles often degrade the selection quality, which imposes an additional burden on users. In this work, we propose a selection technique tolerant to input inaccuracies (See example on Figure 9). We use a dense conditional random field (CRF) to robustly infer a selection from possibly inaccurate input. Further, we show that patch-based pixel similarity functions yield more precise selection than simple point-wise metrics. However, efficiently solving a dense CRF is only possible in low-dimensional Euclidean spaces, and the metrics that we use are high-dimensional and often non-Euclidean. We address this challenge by embedding pixels in a low-dimensional Euclidean space with a metric that approximates the desired similarity function. The results show that our approach performs better than previous techniques and that two options are sufficient to cover a variety of images depending on whether the objects are textured. This work has been published to the Eurographics Conference [15].

6.3.2. *Discrete Texture Design Using a Programmable Approach*

Participants: Hugo Loi, Thomas Hurtut, Romain Vergne, Thollot Joëlle.

Many rendering methods use discrete textures (planar arrangements of vector elements) instead of classic bitmaps. Discrete textures are resolution-insensitive and easily allow to modify the elements' geometry or spatial distribution. However, manually drawing such textures is a time-consuming task. Automating this production is a long-time studied subject. The methods designed for this purpose deal with a difficult tradeoff between the reachable variety of textures and the usability for a community of users. In this work, we show that considering discrete textures as programs allow for a larger variety of textures than relying on a given model. This work has been as a Siggraph 2013 talk [18].

(a) (b) (c)

../../../../projets/maverick/IMG/JonctionPercageRayonVue2.png

(d) (e) (f)

Figure 8. Computing Interfaces of an Aircraft part assembling the wings with the body of an aircraft (model courtesy of EADS). (a,b) two views of the components, (c) exploded view, (d) ray casting for detecting proximities between parts, (e) boundary reconstruction, (f) final interfaces

../../../../projets/maverick/IMG/scribbles.png

Figure 9. Accurate selection from inacurrate input. Although our algorithm is fed with inaccurate selection, it succeeds in separating the background from the foreground components in the image.

../../../../projets/maverick/IMG/programable_textures.jpg

Figure 10. Discrete element texture generated with our programmable approach.

6.4. Complex scenes

6.4.1. Appearance pre-filtering

Participants: Heitz Eric, Neyret Fabrice.

Last year work and HPG'12 paper "Representing Appearance and Pre-filtering Subpixel Data in Sparse Voxel Octrees" was dealing with the light and view dependant aspect of complex surfaces due to sub-pixels details. This was done by replacing sub-voxel height fields by gaussian slope distribution and height-correlated colors by its gradient, feeding a Cook-Torrance-like microfacet brdf.

In continuation of this and in the same spirit of replacing sub-pixel values by gaussian distributions to be shaded using the frame of microfacets brdf, this year we addressed the filtering of color maps (on surfaces and per se), displacement map, and reflectance maps, thus obtaining a complete model of the local rendering integral (see Figure 11).

Note that Eric did his work partly during his 6-monthes stay of University of Montreal in the scope of Exploradoc regional founding. He also colaborated with nVIDIA for an on-going work related to animation of GigaVoxels, and we were invited for a stay of several weeks at Weta Digital, NZ to help them applying our techniques.

6.4.2. Filtering Color Mapped Textures and Surfaces

Participants: Heitz Eric, Neyret Fabrice, Nowrouzezahrai Derek, Poulin Pierre.

Indeed, several ubiquitus CG operations like filtering non-linear functions of the data are still mostly unsolved despite their important flaws. Typically, density, noise data, normals or height are filtered before feeding a color look-up texture, despite the strong non-linearity of the transform forbids factoring it out of the integral. This result on very visible flaws such as thin blue bones+air foams appearing as red muscle at distance in volume visualisation, silhouettes and horizon getting the middle tint instead of the integral of tints, procedural noise bump maps and height fields appearing as smooth instead of rough.

Figure 11. The resulting pixel color is the integral of the local rendering on the surface, which combines 4 fields which are: the color texture parameter, the visibility from eye, the visibility from light, and the shading. As these are varying and non-linear, the ubicuitous simplification of averaging each term separately is not valid.

Assuming Gaussian distribution of colors within a pixel or voxel, the filtered colors values can be represented as color lobes (i.e. histogramms) instead of scalars. In all the cases where the subpixel/voxel raw data can also be represented as gaussian distribution (e.g. Perlin noise), the filtering is just the inner product of the two lobes. It can easily be tabuled as a 1D LUT MIP-map which LOD corresponds to the standard deviation and thus the scale. Since microfacet brdf models allow to estimate the *visible* slope statistics accounting for light and view visibility, this allows for emerging light-view dependant color effect both acurately and very efficiently. Note that the same scheme applies for colors corelated to orientation rather than heights (see Figure 12). This provides a multiscale representation where subpixel/subvoxel data is represented through lobes which can be precalculated or calculated on demand from the thiner level.

This work was published at ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games (I3D) 2013 [17]. An extended version "Filtering Non-Linear Transfer Functions on Surfaces" was published at IEEE Transactions on Visualization and Computer Graphics 2013 [11].

6.4.3. Linear Efficient Antialiased Displacement and Reflectance Mapping

Participants: Dupuy Jonathan, Heitz Eric, Neyret Fabrice.

Here, the last term of the local rendering integration is addressed: the filtering of subpixel/subvoxel geometry and brdf as an appearant brdf applied on a macro-geometry. By re-derivating accurately the brdf of a displacement map assumed to have sub-pixel gaussian distribution (with an exact masking term, more accurate cross-correlated light-view, and offseted appearant lobe), and by noting that the reflectance of the environment can be pre-filtered like the textures of the previous paper, we finally obtain a complete model of pre-fitltered appearance of surfaces (see Figure 13). This work, co-first-authored with Jonathan Dupuy, was published at ACM Transactions on Graphics and presented at Siggraph Asia [8].

6.5. Realistic rendering

6.5.1. Interactive Rendering of Acquired Materials on Dynamic Geometry Using Frequency Analysis

Participants: Bagher M. Mahdi, Soler Cyril, Subr Kartic, Belcour Laurent, Holzschuch Nicolas.

Figure 12. (Left:) Correct pre-filtering of non-linear functions (e.g., LUT) of a noise texture. (Right:) Correct pre-filtering of the appearance of details, with light and view macroscopic dependency and color correlated with depth.

Shading acquired materials with high-frequency illumination is computationally expensive. Estimating the shading integral requires multiple samples of the incident illumination. The number of samples required may vary across the image, and the image itself may have high- and low-frequency variations, depending on a combination of several factors. Adaptively distributing computational budget across the pixels for shading is a challenging problem. In this work, we depict complex materials such as acquired reflectances, interactively, without any precomputation based on geometry. In each frame, we first estimate the frequencies in the local light field arriving at each pixel, as well as the variance of the shading integrand. Our frequency analysis [1] accounts for combinations of a variety of factors: the reflectance of the object projecting to the pixel, the nature of the illumination, the local geometry and the camera position relative to the geometry and lighting. We then exploit this frequency information (bandwidth and variance) to adaptively sample for reconstruction and integration. For example, fewer pixels per unit area are shaded for pixels projecting onto diffuse objects, and fewer samples are used for integrating illumination incident on specular objects (See Figure 14). This work has been published in IEEE Transactions on Visualization and Computer Graphics [3], as a follow up of a previous paper published at the I3D conference.

6.5.2. 5D Covariance Tracing for Efficient Defocus and Motion Blur

Participants: Belcour Laurent, Soler Cyril, Subr Kartic, Holzschuch Nicolas, Durand Frédo.

The rendering of effects such as motion blur and depth-of-field requires costly 5D integrals. We dramatically accelerate their computation through adaptive sampling and reconstruction based on the prediction of the anisotropy and bandwidth of the integrand. For this, we develop a new frequency analysis of the 5D temporal light-field, and show that first-order motion can be handled through simple changes of coordinates in 5D. We further introduce a compact representation of the spectrum using the co- variance matrix and Gaussian approximations. We derive update equations for the 5×5 covariance matrices for each atomic light transport event, such as transport, occlusion, BRDF, texture, lens, and motion. The focus on atomic operations makes our work general, and removes the need for special-case formulas. We present a new rendering algorithm that computes 5D covariance matrices on the image plane by tracing paths through the scene, focusing on the single-bounce case. This allows us to reduce sampling rates when appropriate and perform reconstruction of images with complex depth-of-field and motion blur effects (See Figure 15). This work was published at ACM Transactions on Graphics [5] and presented at Siggraph'2013.

../../../../projets/maverick/IMG/LEAD-R.png

Figure 13. (Left:) Correct pre-filtering of non-linear functions (e.g., LUT) of a noise texture. (Right:) Correct pre-filtering of the appearance of details, with light and view macroscopic dependency and color correlated with depth.

../../../../projets/maverick/IMG/bandwidth_realtime.png

Figure 14. Our simplified bandwidth prediction technique is suitable for real-time rendering. It allows us to only compute a subset of the image pixels while concentrating integration cost to the pixels that have the largest variance of the integrand.

../../../../projets/maverick/IMG/covariance.png

Figure 15. Our covariance analysis of the power spectrum of local illumination allows us to accurately predict the sampling rates an reconstruction filters to significantly increase the convergence of path tracing.

6.5.3. Accurate and Efficient Filtering using Anisotropic Filter Decomposition

Participants: Soler Cyril, Bagher Mahdi, Nowrouzezahrai Derek.

Efficient filtering remains an important challenge in computer graphics, particularly when filters are spatially-varying, have large extent, and/or exhibit complex anisotropic profiles. We explored an efficient filtering approach for these difficult cases based on anisotropic filter decomposition (IFD). By decomposing complex filters into linear combinations of simpler, displaced isotropic kernels, and precomputing a compact prefiltered dataset, we are able to interactively apply any number of—potentially transformed—filters to a signal (See Figure 16). Our performance scales linearly with the size of the decomposition, not the size nor the dimensionality of the filter, and our prefiltered data requires reasonnable storage, comparing favorably to the state-of-the-art. We apply IFD to interesting problems in image processing and realistic rendering. This work is currently under submission and a technical report is already available [21].

`../../../../projets/maverick/IMG/isofilter.png`

Figure 16. Isotropic filter decomposition allows us to efficiently perform anisotropic filtering at the cost of minimal pre-computation.

6.5.4. Double- and Multiple-scattering Effects in Translucent Materials

Participants: Holzschuch Nicolas, Gascuel Jean-Dominique.

Some materials, such as coffee, milk or marble, have a soft translucent aspect because of sub-surface scattering: light enters them, is scattered several times inside before leaving in a different place. A full representation of sub-surface scattering effects in illumination simulation is computationally expensive. The main difficult comes from multiple scattering events: the high number of events increases the uncertainty on the result, forcing us to allocate more time for the computations. Recently, we showed that there is a strong correlation between the surface effects of multiple scattering inside the material and the effects after just two scatter events. This knowledge will help in accelerating multiple scattering effects (see figure 17 . We exploited this knowledge to provide a model and implementation for fast computation of double-scattering events, using a precomputed density function stored in a compact way. This work has been published in IEEE Computer Graphics and Applications [12].

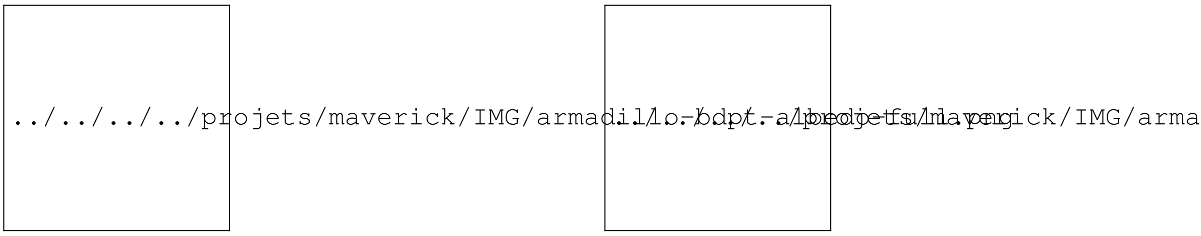


Figure 17. Our approximation achieves results very close to the reference solution: (left) the BDPT reference, in 32min; (right) our solution, in 1.7 min. .

6.6. Inverse problems

6.6.1. Sparse BRDF Approximation using Compressive Sensing

Participants: Zupancic Benoit, Soler Cyril.

BRDF acquisition is a tedious operation, since it requires measuring 4D data. On one side of the spectrum lie explicit methods, which perform many measurements to potentially produce very accurate reflectance data after interpolation. These methods are generic but practically difficult to setup and produce high volume data. On the other side, acquisition methods based on parametric models implicitly reduce the infinite dimensionality of the BRDF space to the number of parameters, allowing acquisition with few samples. However, parametric methods require non linear optimization. They become unstable when the number of parameters is large, with no guaranty that a given parametric model can ever fit particular measurements.

We experiment a new acquisition method where the measurement of the BRDF is performed from a single image, knowing the normals and illumination. To tackle such a severely underconstrained problem, we express the BRDF in a high dimensional basis, and perform the reconstruction using *compressive sensing*, looking for the most sparse solution to the linear problem of fitting the measurement image. Doing so, we leverage the coherency between the measured pixels, while keeping the high dimension of the space the BRDF is searched into.

This work is a very first attempt at reconstructing BRDFs using compressive sensing. In Fig 18 we used a synthetic input image, for the sake of checking the feasibility of the recovery algorithm, in the particular case of an isotropic spatially constant BRDFs. The possibility to extend our theory to non spatially varying and anisotropic BRDFs is currently under investigation. We would like to orient our work toward BRDF acquisition with consumer hardware. In particular, our preliminary results indicate that compressive sensing could achieve an very accurate acquisition with additional input, such as a video of a static object under probed lighting.

This word has been published as a poster to the Siggraph Asia'2013 conference [24].

../../../../projets/maverick/IMG/ssbrdf.png

Figure 18. Reconstruction of a BRDF from a single image. This work stands up as a preliminary experiment to prove that BRDF acquisition can be performed with a much lower amount of data than what the Nyquist rate predicts for data reconstruction, if using compressive sensing methods.

6.6.2. Floating tangents for approximating spatial curves with G1 piecewise helices

Participants: Derouet-Jourdan Alexandre, Bertails-Descoubes Florence, Thollot Joëlle.

Curves are widely used in computer science to describe real-life objects such as slender deformable structures. Using only 3 parameters per element, piecewise helices offer an interesting and compact way of representing digital curves. In our work [7], we present a robust and fast algorithm to approximate Bezier curves with G1 piecewise helices. Our approximation algorithm takes a Bezier spline as input along with an integer N and returns a piecewise helix with N elements that closely approximates the input curve. The key idea of our method is to take N+1 evenly distributed points along the curve, together with their tangents, and interpolate these tangents with helices by slightly relaxing the points. Building on previous work, we generalize the proof for Ghosh's co-helicity condition, which serves us to guarantee the correctness of our algorithm in the general case. Finally, we demonstrate both the efficiency and robustness of our method by successfully applying it to various datasets of increasing complexity, ranging from synthetic curves created by an artist to automatic image-based reconstructions of real data such as hair, heart muscular fibers or magnetic field lines of a star.

6.6.3. Inverse Dynamic Hair Modeling with Frictional Contact

Participants: Derouet-Jourdan Alexandre, Daviet Gilles, Bertails-Descoubes Florence, Thollot Joëlle.

In the latest years, considerable progress has been achieved for accurately acquiring the geometry of human hair, thus largely improving the realism of virtual characters. In parallel, rich physics-based simulators have been successfully designed to capture the intricate dynamics of hair due to contact and friction. However, at the moment there exists no consistent pipeline for converting a given hair geometry into a realistic physics-based hair model. Current approaches simply initialize the hair simulator with the input geometry in the absence of external forces. This results in an undesired sagging effect when the dynamic simulation is started, which basically ruins all the efforts put into the accurate design and/or capture of the input hairstyle. In this work [6] we propose the first method which consistently and robustly accounts for surrounding forces — gravity and frictional contacts, including hair self-contacts — when converting a geometric hairstyle into a physics-based hair model. Taking an arbitrary hair geometry as input together with a corresponding body mesh, we interpret the hair shape as a static equilibrium configuration of a hair simulator, in the presence of gravity as well as hair-body and hair-hair frictional contacts. Assuming that hair parameters are homogeneous and lie in a plausible range of physical values, we show that this large underdetermined inverse problem can be formulated as a well-posed constrained optimization problem, which can be solved robustly and efficiently by leveraging the frictional contact solver of the direct hair simulator. Our method was successfully applied to the animation of various hair geometries, ranging from synthetic hairstyles manually designed by an artist to the most recent human hair data automatically reconstructed from capture.

Maxplus Project-Team

6. New Results

6.1. Théorie spectrale max-plus et géométrie métrique/Max-plus spectral theory and metric geometry

6.1.1. Introduction

Participants: Marianne Akian, Stéphane Gaubert, Cormac Walsh.

Étant donné un noyau $a : S \times S \to \mathbb{R} \cup \{-\infty\}$, on peut lui associer le problème spectral max-plus

$$\sup_{y \in S} a(x,y) + u(y) = \lambda + u(x), \quad \forall x \in S, \tag{13}$$

dans lequel on cherche le vecteur propre $u : S \to \mathbb{R} \cup \{-\infty\}$ et la valeur propre correspondante $\lambda \in \mathbb{R} \cup \{-\infty\}$. Comme nous l'avons rappelé dans les §3.2 et 3.3 , le problème spectral (9) intervient en contrôle ergodique: l'ensemble S est l'espace des états, et l'application $a(x,y)$ fournit le gain associé à la transition $x \to y$. Le cas où S est fini est classique, l'on a alors un résultat précis de représentation de l'espace propre, à l'aide d'un certain graphe, dit graphe critique. Des résultats existent également lorsque S est compact et que le noyau vérifie certaines propriétés de régularité.

Dans [61], nous avons considéré le cas où S est non compact. Lorsque $\lambda = 0$, l'espace propre est analogue à l'espace des fonctions harmoniques défini en théorie (classique ou probabiliste) du potentiel. En introduisant l'analogue max-plus de la frontière de Martin, nous avons obtenu un analogue de la formule de représentation de Poisson des fonctions harmoniques : toute solution u de (9) peut être représentée sous la forme :

$$u = \sup_{w \in \mathcal{M}_m} w + \mu_u(w) \ , \tag{14}$$

où $\mathcal{M}_m \subset (\mathbb{R} \cup \{-\infty\})^S$ est l'analogue max-plus de la frontière de Martin minimale (l'ensemble des fonctions harmoniques extrémales normalisées), et où μ_u joue le rôle de la mesure spectrale. Nous avons montré aussi que les éléments de l'espace de Martin minimal peuvent être caractérisés comme les limites de "quasi-géodésiques". La frontière de Martin max-plus généralise dans une certaine mesure la frontière d'un espace métrique construite à partir des horo-fonctions (fonctions de Busemann généralisées), ou horo-frontière. Ces résultats inspirent les travaux des sections suivantes, qui portent sur des cas remarquables d'espaces métriques (§6.1.3) ou sur des applications en théorie des jeux (§6.1.2).

English version

Let the kernel $a : S \times S \to \mathbb{R} \cup \{-\infty\}$ be given. One may associate the max-plus spectral equation (9), where the eigenvector $u : S \to \mathbb{R} \cup \{-\infty\}$ and the eigenvalue $\lambda \in \mathbb{R} \cup \{-\infty\}$ are unknown. As we recalled in §3.2 and refmonotone, this spectral problem arises in ergodic optimal control: the set S is the *state space*, and the map $a(x,y)$ is the *transition reward*. The case when S is finite is classical, a precise spectral theorem is known, with a characterisation of the eigenspace in terms of a critical graph. Some results have been shown when S is compact, assuming that the kernel a satisfies some regularity properties.

In [61], we considered the case where S is non-compact. When $\lambda = 0$, the eigenspace is analoguous to the set of harmonic functions defined in classical or probabilistic potential theory. By introducing a max-plus analogue of the classical Martin boundary, we obtained an analogue of the Poisson representation of harmonic functions, showing that any solution u of (9) may be represented as in (10) where $\mathcal{M}_m \subset (\mathbb{R} \cup \{-\infty\})^S$ is a max-plus analogue of the minimal Martin boundary (the set of normalised extremal harmonic functions), and μ_u plays the role of the spectral measure. We also showed that the elements of the minimal Martin boundary can be characterised as limits of certain "almost-geodesics". The max-plus Martin boundary generalises to some extent the boundary of metric spaces defined in terms of horofunctions (generalised Busemann functions), or horoboundary. These results have inspired the work of the next sections, which deal either with interesting examples of metric spaces (§6.1.3) or with applications to zero-sum games (§6.1.2).

6.1.2. *Asymptotiques d'itérées d'applications contractantes au sens large et jeux à somme nulle en horizon long/Asymptotics of iterates of nonexpansive mappings and zero-sum games*

Participants: Jérôme Bolte, Stéphane Gaubert, Guillaume Vigeral.

On s'intéresse ici à l'existence du paiement moyen pour les jeux répétés, et plus généralement, à l'existence du vecteur de "taux de fuite" $\lim_k f^k(x)/k$ où f est une application de $\mathbb{R}^n$ dans lui même, nonexpansive pour une norme quelconque. Dans le cas particulier des jeux, f est un opérateur de Shapley, qui est nonexpansif pour la norme sup. On montre dans [45] que la limite existe si l'application f est définissable dans une structure o-minimale. Ceci généralise des résultats de Bewley, Kohlberg, et Neyman, qui montraient que la limite existe si f est semi-algébrique. L'extension au cas o-minimal permet notamment de traiter des opérateurs de type "log-exp" apparaissant en contrôle sensible au risque. Ce travail traite aussi de la question de savoir si un jeu dont les fonctions de paiement et de transition sont définissables dans une structure o-minimale admet un opérateur de Shapley f définissable. Un contre exemple montre que f n'est pas forcément définissable dans la même structure, mais l'on montre qu'il en est ainsi dès que les probabilités de transition ont une structure séparable.

English version

We study the question of the existence of the mean payoff for repeated games, and more generally, the existence of a vector of "escape rates", $\lim_k f^k(x)/k$, where f is a self-map of $\mathbb{R}^n$, non-expansive in some norm. In the special case of zero-sum games, f is a Shapley operator, and it is sup-norm nonexpansive. We showed in [45] that this limit does exist as soon as the map f is definable in an o-minimal structure. This generalizes results of Bewley, Kohlberg, and Neyman, who showed that this limit exists if f is semi-algebraic. The extension to the case of o-minimal structures allows one in particular to deal with log-exp type operators arising in risk sensitive control. This work also adresses the question of knowing whether a game with definable payment and transition functions has a Shapley operator that is definable in the same structure. We gave a counter example showing that this may not be the case, but showed that the Shapley operator is definable as soon as the transition probabilities have a separable structure.

6.1.3. *Isométries de la géométrie de Hilbert/Isometries of the Hilbert geometry*

Participants: Cormac Walsh, Bas Lemmens [Kent University, UK].

L'un des intérêts de l'horo-frontière est de renseigner sur le groupe des isométries d'un espace métrique. En effet, ce groupe agit naturellement sur l'horo-frontière, et cette action peut parfois être mieux comprise que l'action du groupe sur l'espace d'origine.

Nous avons utilisé ces idées pour étudier le groupe des isométries pour la métrique de Hilbert. De La Harpe [179] a donné plusieurs conjectures relatives à ce groupe. Nous avons montré dans [51], en utilisant l'horo-frontière, que le groupe des isométries est exactement le groupe des transformations linéaires projectives à moins que le domaine ne soit une coupe d'un cône symmétrique non-Lorentzien. Dans ce dernier cas, le groupe linéaire projectif est d'index 2 dans le groupe des isométries. Le cas particulier où le domaine est un polytope a été traité précédemment dans [136].

Dans [51] nous déterminons aussi le groupe des isométries pour une métrique fortement reliée à la métrique de Hilbert, à savoir la métrique de Thompson sur un cone.

English version

One use for the horofunction boundary is to study the group of isometries of a metric space. This is because this group has a well defined action on the horoboundary and it is likely that in many cases this action will be easier to understand than the action on the space itself.

We have been applying these ideas to investigate the isometries of the Hilbert geometry. De La Harpe [179] has previously made several conjectures about the isometry group of this space. We have shown [51] using the horofunction boundary that the isometry group is exactly the group of projective linear transformations unless the domain on which the geometry is defined is a cross section of a non-Lorentzian symmetric cone, in which case the projective linear group is of index two in the isometry group.

The special case when the domain is a polytope was previously considered in [136].

In the paper [51], we also determine the isometry group of closely related metric, the Thompson geometry on a cone.

6.1.4. Consensus non-commutatif et contraction d'opérateurs de Kraus/Noncommutative consensus and contraction of Kraus maps

Participants: Stéphane Gaubert, Zheng Qu.

Dans le travail [47], on s'est intéressé à la vitesse de convergence vers l'équilibre d'une itération de la forme $x^{k+1} = T(x^k)$, $x^k \in X$, où T est une application linéaire préservant un cône dans un espace de Banach X, telle que $T(e) = e$, pour un certain vecteur e dans l'interieur du cône. On s'intéresse aussi à l'itération dans l'espace dual, $y^{k+1} = T^*(y^k)$, $y^k \in X^*$, lorsque $\langle y^0, e\rangle = 1$.

Le cas classique est celui où $T(x) = Px$ est un opérateur de Markov. L'itération primale traduit alors la convergence vers le "consensus", et l'itération duale traduit la convergence de la distribution de probabilité en temps k vers l'état stationnaire. Dans ce cas, le taux de contraction (en un coup) $\kappa(P)$ d'une itération primale, pour la semi-norme de Hilbert $\|z\|_H := \max_i z_i - \min_j z_j$, ainsi que le taux de contraction d'une itération duale, pour la métrique en variation totale, coïncident et sont caractérisés par une formule dûe à Doeblin et Dobrushin (coefficient d'ergodicité),

$$\kappa(P) := 1 - \min_{i,j} \sum_{s=1}^{n} \min(P_{is}, P_{js}).$$

On a donné ici une généralisation de cette formule au cas d'opérateurs abstraits, qui s'applique en particulier aux opérateurs de Kraus qui interviennent en information quantique. Ces derniers opérent sur l'espace des matrices symmétriques, et sont de la forme

$$T(x) = \sum_k a_k x a_k^* \qquad \text{avec} \qquad \sum_k a_k a_k^* = I \ .$$

Dans [34], nous avons étendu ces résultats aux flots non-linéaires sur les cones.

English version

In a recent work [27], we studied the speed of convergence to equilibrium of an iteration of the form $x^{k+1} = T(x^k)$, $x^k \in X$, where T is a linear map preserving a cone in a Banach space X, such that $T(e) = e$, for some vector e in the interior of the cone. We also considered the iteration in the dual space X^*, $y^{k+1} = T^*(y^k)$, $y^k \in X^*$, where $\langle y^0, e\rangle = 1$.

The classical application arises when $T(x) = Px$ is a Markov operator. Then, the primal iteration represents the dynamics of consensus, whereas the dual iteration represents the evolution of the probability distribution as a function of time. Then, the (one-shot) contraction rate $\kappa(P)$ of the primal iteration, with respect to Hilbert's seminorm $\|z\|_H := \max_i z_i - \min_j z_j$, and the contraction rate of the dual iteration, with respect to the total variation metric, coincide, and are characterized by a formula of Doeblin and Dobrushin (ergodicity coefficient),

$$\kappa(P) := 1 - \min_{i,j} \sum_{s=1}^{n} \min(P_{is}, P_{js}).$$

We gave here a generalization of this formula to an abstract operators on a cone. This covers in particular the Kraus maps arising in quantum information theory. The latter maps act on the space of symmetric matrices. They can be written as

$$T(x) = \sum_k a_k x a_k^* \qquad \text{with} \qquad \sum_k a_k a_k^* = I \ .$$

In [34], we generalized these results to non-linear flows over cones.

6.2. Algèbre linéaire max-plus et convexité abstraite/Max-plus linear algebra and abstract convex analysis

6.2.1. Convexité max-plus ou tropicale/Max-plus or tropical convexity

Participants: Xavier Allamigeon, Stéphane Gaubert, Eric Goubault [CEA], Ricardo Katz [Conicet, Argentine].

On étudie les analogues max-plus ou tropicaux des ensembles convexes. Ceux-ci sont utiles en particulier pour représenter de manière effective les ensembles d'états accessibles de systèmes à événements discrets [9], ils sont aussi apparus récemment en géométrie tropicale, dans toute une série de travaux à la suite de Sturmfels et Develin [96]. Les polyèdres max-plus peuvent aussi être vus comme des limites de déformations de polyèdres classiques, sur lesquels ils donnent un éclairage de nature combinatoire. Toutes ces motivations ont inspiré la recherche d'analogues des résultats fondamentaux d'analyse convexe classique: séparation, projection, points extrémaux, à la suite en particulier de [8].

Dans un travail de X. Allamigeon, S. Gaubert, et E. Goubault [64], [16], on a mis en évidence un critère combinatoire pour la caractérisation des sommets des polyèdres tropicalement convexes. Celui-ci s'exprime à l'aide d'hypergraphes orientés, et de leurs composantes fortement connexes. Ce critère possède la propriété d'être vérifiable en un temps presque linéaire en la taille de l'hypergraphe.

On en déduit un analogue tropical de la méthode de la double description [16] (méthode très utilisée sur les polyèdres classiques, et dûe à Motzkin *et al.* [148]). Cet algorithme permet de calculer les sommets d'un polyèdre défini de façon externe (intersection de demi-espaces ou d'hyperplans tropicaux). Grâce au critère combinatoire précédent, l'algorithme améliore de plusieurs ordres de grandeur les techniques connues jusqu'alors. Ceci est confirmé par de nombreuses expérimentations. Ce travail est motivé par des applications à l'analyse statique [63] et aux systèmes à événements discrets [99], dans lesquelles la manipulation de tels polyèdres est le goulot d'étranglement.

Il est connu qu'un polyèdre tropical peut être représenté comme l'enveloppe convexe d'un ensemble minimal de points et rayons, donnés par ses sommets et ses rayons extrêmes [111]. Dans un travail réalisé par X. Allamigeon et R. Katz [17], et effectué en partie lors de visites de R. Katz à Inria, on étudie la question duale de la caractérisation des représentations minimales par demi-espaces. On montre qu'un polyèdre tropical possède *essentiellement* une unique représentation minimale par demi-espaces, lorsque leurs apex appartiennent au polyèdre. On montre que les apex de ces demi-espaces non-redondants correspondent à certains sommets du complexe tropical introduit par Develin et Sturmfels [96]. On introduit également un critère combinatoire pour l'élimination de demi-espaces redondants à l'aide d'hypergraphes orientés.

Dans un travail en cours de X. Allamigeon, P. Benchimol, S. Gaubert et R. Katz, nous étudions la tropicalisation des représentations par demi-espaces des polyèdres convexes sur le corps des séries de Puiseux. Nous démontrons ainsi une conjecture de Develin et Yu [97]. Celle-ci assure qu'étant donné un polytope tropical pur, il existe un polytope *relevé* sur les séries de Puiseux, dont les demi-espaces associés aux faces se "tropicalisent" en une représentation par demi-espaces du polytope tropical initial.

Des applications de ces travaux à l'algorithmique, concernant en particulier les jeux répétés, sont discutées dans la Section 6.4.2 . Une application aux systèmes temps réel est discutée dans la Section 6.5.4 .

English version

We study the max-plus or tropical analogues of convex sets. These have been used in particular to represent effectively the accessible sets of certain discrete event systems [9]. They also appeared in tropical geometry, following the work of Sturmfels and Develin [96]. Max-plus polyhedra can be thought of as limits of deformations of classical polyhedra, on which they give a combinatorial insight. These motivations have inspired the investigation of analogues of basic results of classical convex analysis: separation, projection, representation by extreme points, following [8].

In a work of X. Allamigeon, S. Gaubert, and E. Goubault [16], we introduce a combinatorial criterion for the characterization of the vertices of tropically convex polyhedra. It is expressed in terms of directed hypergraphs and their strongly connected components. This criterion can be verified in almost linear time in the size of the hypergraph.

This allows to develop a tropical analogue of the double description method [16] (this method is widely used for classical convex polyhedra, and is due to Motzkin *et al.* [148]). This algorithm is able to determine all the vertices of a polyhedron defined externally (intersection of tropical half-spaces of hyperplanes). Thanks to the combinatorial criterion mentioned above, the algorithm improves the existing methods by several orders of magnitude. This is confirmed by several experiments. This is motivated by applications to static analysis [63] and discrete event systems [99], in which computing such polyhedra turns out to be the bottleneck.

It is well-known that a tropical polyhedron can be represented as the convex hull of a minimal set of points and rays, provided by its vertices and extreme rays [111]. In a work of X. Allamigeon and R. Katz [17], partly done during visits of R. Katz at Inria, the dual problem of characterizing the minimal representations by half-spaces is studied. We show that a tropical polyhedron admits *essentially* a unique minimal external representation by half-spaces, provided that their apices belong to the polyhedron. We prove that the apices of these half-spaces correspond to certain vertices of the tropical complex introduced by Develin and Sturmfels [96]. We also establish a combinatorial criterion allowing to eliminate redundant half-spaces using directed hypergraphs.

In an ongoing work of X. Allamigeon, P. Benchimol, S. Gaubert and R. Katz, we study the tropicalization of the representation by half-spaces of convex polyhedra over the field of Puiseux series. In particular, we prove a conjecture of Develin and Yu [97]. It states that, given a pure tropical polytope, there exists a lifting polytope over Puiseux series, such that the facet-defining half-spaces are "tropicalized" into a representation by half-spaces of the initial polytope.

Some algorithmic applications of this work concerning in particular mean payoff games, will be discussed in Section 6.4.2 . Applications to real time systems will be discussed in Section 6.5.4 .

6.2.2. *Systèmes linéaires max-plus/Max-plus linear systems*

Participants: Marianne Akian, Stéphane Gaubert, Alexander Guterman [Moscow State University].

Dans [37], on montre des formules de Cramer pour des systèmes linéaires sur diverses extensions du semi-anneau max-plus. Les éléments de ces extensions sont des nombres tropicaux enrichis d'une information de multiplicité, de signe ou d'angle par exemple. On obtient ainsi des résultats d'existence et d'unicité qui généralisent plusieurs résultats de [121], [153], [107], [161], [127]. De plus, pour certaines extensions du semi-anneau max-plus, les preuves fournissent des algorithmes de type Jacobi ou Gauss-Seidel pour résoudre les systèmes linéaires.

English version

In [37], we prove general Cramer type theorems for linear systems over various extensions of the tropical semiring, in which tropical numbers are enriched with an information of multiplicity, sign, or argument. We obtain existence or uniqueness results, which extend or refine earlier results in [121], [153], [107], [161], [127]. Moreover, some of our proofs lead to Jacobi and Gauss-Seidel type algorithms to solve linear systems in suitably extended tropical semirings.

6.3. Algèbre max-plus, déformations et asymptotiques /Max-plus algebra, deformations and asymptotic analysis

6.3.1. Introduction

Comme indiqué dans le §3.7 , l'algèbre max-plus est la limite d'une déformation de l'algèbre classique, ou plutôt du semi-corps des réels positifs. Elle peut aussi fournir des estimations de ces déformations, puisque

$$\max(a,b) \leq \epsilon \log(e^{a/\epsilon} + e^{b/\epsilon}) \leq \epsilon \log(2) + \max(a,b) \ . \tag{15}$$

L'utilisation de ces propriétés a déjà conduit dans le passé aux travaux sur les perturbations de valeurs propres [55], [54], [53], ou sur les grandes déviations [1], [59]. Dans les travaux qui suivent, nous exploitons ces propriétés dans des contextes reliés ou similaires à ceux de nos travaux précédents.

English version

As detailled in §3.7 , max-plus algebra is the limit of a deformation of classical algebra, or more precisely of the semi-field of usual real positive numbers. It can also give estimations for these deformations using for instance (11). By using these properties, we already obtained some works on singular perturbations of matrix eigenvalues [55], [54], [53], or on large deviations [1], [59]. In the works described below, we are exploiting again these properties in contexts that are related or similar to those of our earlier works.

6.3.2. Aspects tropicaux des algorithmes de scaling matriciel/Tropical aspects of matrix scaling problems

Participants: Marianne Akian, Stéphane Gaubert, Meisam Sharify Najafabadi [Univ. Manchester].

Une partie du travail de thèse de M. Sharify [167] portait sur les méthodes de mise à l'échelle pour améliorer la précision du calcul de valeurs propres. En appliquant les techniques de [53], [54], on montrait notamment que l'ordre de grandeur des valeurs propres d'un faisceau matriciel est donné (sous des conditions de non-dégénerescence) par les valeurs propres tropicales, qui peuvent être calculées de manière robuste, et fournissent ainsi une mise à l'échelle pour calculer les valeurs propres classiques.

Nous avons poursuivi ce travail dans [41]. On calcule cette fois l'ordre de grandeur des valeurs propres d'un polynôme matriciel au moyen des racines tropicales du polynôme obtenu en appliquant une norme donnée aux coefficients. Les racines dépendent de la norme choisie, et la norme de Frobenius est optimale en un certain sens. On obtient des bornes générales pour les ratios entre modules des valeurs propres et racines tropicales qui généralisent les bornes obtenues par Polya et Ostrowski dans le cas de polynômes scalaires. On raffine aussi ces bornes, en particulier lorsque les racines tropicales sont bien séparées les unes des autres.

English version

A part of the PhD work of M. Sharify [167] dealt with scaling methods to improve the accuracy of eigenvalue numerical computions. Applying the techniques of [53], [54], we showed in particular that the order of magnitude of the eigenvalues of a matrix pencil can be determined (under nondegeracy conditions) by computing tropical eigenvalues. The latter can always be computed accurately and provide a scaling which can be combined with standard numerical methods for matrix pencils.

We have pursued this work in [41]. Now, we compute the order of magnitude of the eigenvalues of a matrix polynomial by using the tropical roots of a polynomial obtained by applying a norm to the coefficients of the original matrix polynomial. The tropical roots depend on the chosen norm, and the Frobenius turns out to be optimal in a certain sense. We obtain indeed general bounds on the ratios between the modulus of the eigenvalues of the matrix polynomial and the tropical roots which generalize the bounds of Polya and Ostrowski available for scalar polynomials. We also improve these bounds, in particular when the tropical roots are well separated.

6.3.3. Méthodes tropicales de localisation de valeurs propres de matrices/Tropical methods for the localisation of matrix eigenvalues

Participants: Marianne Akian, Stéphane Gaubert, Andrea Marchesini.

Le travail de stage de M2 d'Andrea Marchesini a conduit à la publication [14] dans laquelle on montre des inégalités de type majorisation entre les valeurs propres d'une matrice et les valeurs propres tropicales de la matrice de ses modules. En particulier, la majoration est une généralisation de l'inégalité de Friedland [106] concernant le rayon spectral.

La thèse d'Andrea Marchesini s'inscrit dans le prolongement de son stage de M2 dans l'équipe et certains des travaux de la thèse de Meisam Sharify [167]. Le but est d'obtenir des inégalités de type majorisation permettant d'estimer a priori les valeurs propres de matrices ou de faisceaux de matrices, en faisant éventuellement intervenir des hypothèses de bon conditionnements. En particulier on recherche la localisation de ces valeurs propres en fonction de valeurs propres de matrices agrégées ou simplifiées. On cherchera aussi à obtenir le même type de localisation ou d'estimation dans le cas des vecteurs propres associés, par exemple en utilisant les techniques de compléments de Schur de [54] ou les idées de Murota [149].

L'idée est ensuite d'utiliser ces résultats de localisation pour améliorer la précision des algorithmes de calcul numérique de valeurs propres de matrices, en particulier en construisant des changements d'échelle exploitant les calculs tropicaux, à effectuer préalablement à l'appel d'algorithmes classiques comme QZ. Les travaux de Stéphane Gaubert et Meisam Sharify [115] ont montré l'intérêt de cette approche, notamment pour les problèmes de faisceaux quadratiques de valeurs propres issus de systèmes mécaniques pour lesquels on dispose de nombreux exemples pathologiques pour les algorithmes existants. Dans un travail en collaboration avec Françoise Tisseur et James Hook de l'Université de Manchester, on montre l'intérêt des changements d'échelle en termes de le conditionnement des valeurs propres.

English version

The M2 internship of Andrea Marchesini led to the publication [14], in which we show majorization type inequalities between the eigenvalues of a matrix and the tropical eigenvalues of the matrix obtained by applying the modulus entrywise. In particular, the bound is a generalization of the inequality of Friedland [106] concerning the spectral radius.

The PhD thesis follows his M2 internship and some of the works of Meisam Sharify's PhD thesis [167]. The aim is to obtain majorization type inequalities allowing one to estimate the eigenvalues of matrices or matrix polynomials, using possibly assumptions on condition numbers. In particular, one may look for estimates of these eigenvalues using the eigenvalues of aggregated or simplified matrices. One may also try to find the same type of estimates for the associated eigenvectors, for instance by using techniques of Schur complements from [54] or ideas of Murota [149].

One would like to use these estimation results to improve the accuracy of eigenvalue numerical computions, in particular by constructing scaling methods using tropical techniques, which may be used before calling usual algorithms as QZ. The works of Stéphane Gaubert and Meisam Sharify [115] showed the interest of this approach, in particular for quadratic matrix polynomials issued from mechanical systems for which there exists several pathological examples for existing algorithms. In a work with Françoise Tisseur and James Hook from Manchester University, we show the interest of these scaling methods on the eigenvalue conditioning.

6.3.4. Mesures et applications maxitives/Maxitive measures and maps

Participants: Marianne Akian, Stéphane Gaubert, Paul Poncet.

La thèse de Paul Poncet [154] concernait essentiellement ce que l'on appelle l'analyse idempotente, c'est-à dire l'étude des espaces fonctionnels ou linéaires de dimension infinie sur l'algèbre tropicale, ou tout autre semi-anneau idempotent. Paul Poncet a développé pour cela un point de vue treillis continus comme dans [1], ou plus généralement domaines. Depuis la soutenance, plusieurs articles issus du manuscrit de thèse sont en cours de publication ou de soumission, et d'autres travaux pousuivant ceux de la thèse sont en cours avec les membres de l'équipe.

La première partie de la thèse traitait des mesures maxitives, en particulier de l'existence d'une densité cardinale ou d'une densité d'une mesure par rapport à une autre (théorème de Radon-Nikodym), et de la régularité d'une mesure maxitive. Ces travaux sont publiés ou en cours de publication dans [49] et [23] respectivement.

La deuxième partie concernait les convexes dans les semi-treillis ou l'algèbre max-plus, pour lesquels Paul Poncet a pu établir des théorèmes de type Krein-Milman, réciproque de Milman, et représentation de Choquet. [48] traite du cas des semi-treillis.

On sait que les résultats sur les convexes tropicaux de dimension infinie de [154] permettent de retrouver partiellement les résultats sur la frontière de Martin max-plus décrits dans la section 6.1.1 . Dans un travail commun nous essayons d'obtenir d'autres applications et extensions du théorème de représentation de Choquet tropical. En particulier on considère le cas d'ensembles ordonnés qui ne sont pas forcément des treillis tels que le cône des matrices symmetriques positives muni de l'ordre de Loewner.

English version

The PhD thesis work of Paul Poncet [154] concerned essentially what is called idempotent analysis, that is the study of infinite dimensional functional or linear spaces over tropical algebra, or any other idempotent semiring. For this aim, Paul Poncet developped the point of view of continuous lattices, as in [1], or more generally of domains. Since the defense of his thesis, several papers derived from the thesis manuscript have been submitted and some are published or up to be published. Some other works pursuing the thesis work are done with team members.

The first part of the Paul Poncet's thesis concerned maxitive measures, in particular the existence of a cardinal density of a measure, or that of a density of a measure with respect to another (Radon-Nikodym theorem), and the regularity of a maxitive measure. These works are now published or accepted for publication in [49] and [23] respectively.

A second part concerned convex sets in lattices or max-plus algebra, for which Paul Poncet showed results such as a Krein-Milman type theorem, a Milman converse type theorem, and a Choquet representation type theorem. [48] concerns the case of semilattices.

We know that the results on infinite dimensional tropical convex sets of [154] allow one to recover at least partially the results on max-plus Martin boundaries described in Section 6.1.1 . In a joint work, we try to obtain other applications and extensions of the max-plus Choquet representation theorem. In particular, we consider the case of ordered sets that are not necessarily semilattices, such as the cone of nonnegative symmetric matrices endowed with the Loewner order.

6.4. Algorithmes/Algorithms

6.4.1. Itération sur les politiques pour le contrôle stochastique et les jeux répétés à somme nulle/Policy iterations for stochastic control and repeated zero sum games

Participants: Marianne Akian, Jean Cochet-Terrasson [CGA], Sylvie Detournay, Stéphane Gaubert.

L'algorithme d'itération sur les politiques est bien connu pour résoudre efficacement les équations de la programmation dynamique associées à des problèmes de contrôle stochastique avec critère à horizon infini (Howard) ou ergodique (Howard, et Denardo et Fox). Récemment, il a été généralisé au cas de problèmes de jeux à deux joueurs et somme nulle dégénérés (avec paiements ergodiques et de type "multi-chaîne"), au moyen de techniques d'algèbre max-plus et de théorie du potentiel non linéaire [84]. Chaque itération de base de cet algorithme utilise la résolution d'un système d'équations linéaires dont l'opérateur est monotone, mais dont la taille peut être grande, soit parce qu'il provient d'une discrétisation fine d'une équation aux dérivées partielles, soit parce qu'il est associé à un problème discret de grande taille comme le graphe du Web.

La thèse de Sylvie Detournay [95] a permis de développer et d'étudier un algorithme associant une méthode d'itération sur les politiques du type de celle introduite par Cochet-Terrasson et Gaubert dans [84] et une méthode multigrille algèbrique, afin de résoudre des problèmes de jeux à somme nulle dégénérés, éventuellement posés directement sous forme discrète. L'ensemble des codes nouveaux associés, écrits en C, est déposé sur le projet "pigames" de la gforge et sera disponible librement.

Sylvie Detournay a en particulier implémenté et raffiné l'algorithme proposé dans [84], en l'associant soit à des méthodes de résolution exacte de systèmes linéaires, soit à des méthodes multigrilles algébriques, en utilisant aussi des méthodes multigrilles multiplicatives pour le calcul de la mesure invariante de chaînes de Markov irréductibles, comme celles introduites par De Sterck. Ceci a permis l'obtention de résultats numériques dans le cas de discrétisations d'équations d'Isaacs associées à des jeux de poursuite déterministes ou aléatoires. Cela a aussi permis de tester de manière systématique l'algorithme sur des instances aléatoires de jeux de type Richman. Certains de ces résultats, ainsi que la présentation de l'algorithme (de manière plus concrète que dans [84], et avec les détails d'implémentation) sont présentés dans [24]. Des details supplémentaires ainsi que la preuve de convergence de l'algorithme peuvent être trouvés dans [56].

Des résultats récents de Ye ainsi que Hansen, Miltersen et Zwick montrent que l'algorithme d'itération sur les politiques, restreint à la classe des jeux à somme nulle (à 1 ou 2 joueurs) actualisés de facteur d'actualisation donné, est fortement polynomial. Dans [40], [29], on montre que ceci est le cas aussi pour l'algorithme d'itération sur les politiques pour les jeux à somme nulle et paiement moyen, restreint à la classe des jeux qui ont temps moyen de retour ou d'arrivée à un état donné borné. La preuve utilise des techniques de théorie de Perron-Frobenius non-linéaire, permettant de ramener le problème à paiement moyen à un problème actualisé (de facteur d'actualisation dépendant de l'état et des actions). La même technique permet aussi de traiter le cas de jeux à somme nulle actualisés dont le facteur d'actualisation peut dépendre de l'état et des actions et prendre éventuellement des valeurs supérieures à 1.

English version

Policy iteration is a powerful and well known algorithm to solve the dynamic programming equation associated to stochatic control (one player game) problems with infinite horizon criterion (Howard) or ergodic criterion (Howard and Denardo and Fox). It has recently been extended to degenerate two players problems (with ergodic payoff and in "multichain" cases) using ideas from max-plus algebra and nonlinear potential theory [84]. One basic iteration of the algorithm consists in solving a linear system the operator of which is monotone, but with a size which may be large since it comes from the discretization of a partial differential equation or since it is associated to a large size discrete problem arising from instance from the Web graph.

The PhD thesis of Sylvie Detournay [95] developped and studied an algorithm for degenerate two player games (that may come from a discrete time and finite state space model) combining a policy iteration such as the one introduced in [84] by Cochet-Terrasson et Gaubert, and an algebraic multigrid method (AMG). All new corresponding algorithms, coded in C, belong to the gforge project "pigames" and will be distributed openly.

In particular, Sylvie Detournay has implemented and refined the algorithm proposed in [84], while associating it either to direct linear solvers, or to the AMG methods already used in the nondegenerate case, and using also multiplicative AMG methods for computing invariant measures of Markov chains, such as the one introduced by De Sterck. This allowed her to obtain numerical results in the case of discretisations of Isaacs equations associated to deterministic or stochastic pursuit games. This also allowed her to test systematically the algorithm on random instances of Richman type games.

Some of these results, together with the presentation of the algorithm (in a more practical manner than in [84], with implementation details) are gathered in [24]. Additional details and the convergence proof of the algorithm can be found in [56].

Recent results of Ye and Hansen, Miltersen and Zwick show that policy iteration for one or two player (perfect information) zero-sum stochastic games, restricted to instances with a fixed discount rate, is strongly polynomial. In [40], [29], we show that policy iteration for mean-payoff zero-sum stochastic games is also strongly polynomial when restricted to instances with bounded first mean return time to a given state. The proof is based on methods of nonlinear Perron-Frobenius theory, allowing us to reduce the mean-payoff problem to a discounted problem with state dependent discount rate. Our analysis also shows that policy iteration remains strongly polynomial for discounted problems in which the discount rate can be state dependent (and even negative) at certain states, provided that the spectral radii of the nonnegative matrices associated to all strategies are bounded from above by a fixed constant strictly less than 1.

6.4.2. Algorithmique des polyèdres tropicaux/Algorithmics of tropical polyhedra

Participants: Xavier Allamigeon, Pascal Benchimol, Stéphane Gaubert, Eric Goubault [CEA], Michael Joswig [TU Darmstadt].

X. Allamigeon, S. Gaubert, et E. Goubault, ont développé dans [63], [16] plusieurs algorithmes permettant de manipuler des polyèdres tropicaux. Ceux-ci correspondent aux travaux décrits dans §6.2.1 . Ils permettent notamment de déterminer les sommets et rayons extrêmes d'un polyèdre tropical défini comme intersection de demi-espaces, ou inversement, de calculer une représentation externe à partir d'un ensemble de générateurs. Ces algorithmes sont implémentés la bibliothèque TPLib (voir §5.3).

Dans un travail en cours de X. Allamigeon, P. Benchimol, S. Gaubert et M. Joswig, nous avons défini un analogue tropical de l'algorithme du simplexe qui permet de résoudre les problèmes de *programmation linéaire tropicale*, *i.e.*

$$\begin{array}{rl} \text{minimiser} & \max\limits_{1\leq j\leq n} c_j + x_j \\ \text{sous les contraintes} & \max\big(\max\limits_{1\leq j\leq n}(a^+_{ij}+x_j), b^+_i\big) \geq \max\big(\max\limits_{1\leq j\leq n}(a^-_{ij}+x_j), b^-_i\big), \quad i=1,\cdots,m \\ & x\in(\mathbb{R}\cup\{-\infty\})^n \end{array} \tag{16}$$

où les entrées du programme $a^{\pm}_{ij}$, $b^{\pm}_i$, c_j sont à valeur dans $\mathbb{R}\cup\{-\infty\}$. Ces problèmes sont intimement liés à la résolution de jeux répétés à somme nulle, puisque résoudre un jeux à paiement moyen déterministe est équivalent à déterminer si un problème de programmation linéaire admet un point réalisable [57].

Comme son homologue usuel, le simplexe tropical pivote entre des points de base (tropicaux), jusqu'à atteindre l'optimum du programme linéaire. La différence fondamentale avec l'algorithme du simplexe classique est que le pivotage est réalisé de manière purement combinatoire, en s'appuyant sur des descriptions locales du polyèdre tropical défini par les contraintes à l'aide d'(hyper)graphes orientés. Ceci nous a permis de prouver que *l'étape de pivotage (incluant le calcul des coûts réduits) a la même complexité en temps que dans l'algorithme classique, i.e.* $O(n(m+n))$. Ceci est d'autant plus inattendu que la structure des arêtes tropicales entre deux points de base sont géométriquement plus complexes (elles sont constituées de plusieurs segments de droite, jusqu'à n).

Le simplexe tropical a la propriété d'être fortement corrélé avec l'algorithme du simplexe classique. Grâce au principe de Tarski, le simplexe usuel peut être transposé tel quel sur des programmes linéaires dont les coefficients en entrée sont non plus des réels, mais sur le corps $\mathbb{R}\{\{t\}\}$ des séries de Puiseux généralisées en une certaine indéterminée t, *i.e.* des objets de la forme :

$$c_{\alpha_1}t^{\alpha_1} + c_{\alpha_2}t^{\alpha_2} + \cdots \tag{17}$$

où les α_i sont des réels, les coefficients c_{α_i} sont des réels non-nuls, et où la séquence des $\alpha_1, \alpha_2, \cdots$ est strictement croissante et soit finie, soit non-bornée. L'opposé du plus petit exposant de la série, $-\alpha_1$, est appelé *valuation* de la série. Un programme linéaire tropical est dit *relevé* en un problème linéaire sur $\mathbb{R}\{\{t\}\}$, si la valuation des coefficients en entrée de ce dernier sont égaux aux coefficients du problème tropical. Dans nos travaux, nous avons établi la correspondance suivante entre le simplexe usuel et le simplexe tropical : *pour tout programme linéaire tropical générique, l'algorithme du simplexe tropical trace l'image par la valuation du chemin sur l'algorithme du simplexe usuel sur n'importe quel relèvement du programme tropical dans* $\mathbb{R}\{\{t\}\}$.

Les résultats présentés ci-dessus sont rassemblés dans le preprint [43]. Ils ont fait l'objet de plusieurs présentations en conférence [32], [33].

Ces résultats ouvrent la possibilité de relier la complexité du l'algorithme du simplexe usuel avec celles des jeux déterministes. Pour ces derniers, on sait seulement que leur résolution est dans la classe de complexité $\mathsf{NP} \cap \mathsf{coNP}$, et on ignore s'il existe un algorithme de complexité polynomiale. De façon similaire, on ne sait pas caractériser de façon précise la complexité de l'algorithme du simplexe usuel. Celle-ci dépend fortement de la règle de pivotage utilisée, et il existe des problèmes sur lesquelles de nombreuses règles de pivotage ont une complexité exponentielle. L'existence d'une règle de pivotage qui permettrait au simplexe de terminer en temps polynomial sur n'importe quelle instance est encore aujourd'hui une question ouverte.

Dans un deuxième travail, nous avons relié les deux problèmes ouverts précédents, grâce à l'algorithme du simplexe tropical. Nous avons en effet exhibé une classe de règles de pivotage, dites *combinatoires*, et avons montré qu'elles satisfont la propriété suivante : *s'il existe une règle de pivotage combinatoire qui permet de résoudre tout problème de programmation linéaire usuel en temps polynomial, alors on peut résoudre les jeux à paiement moyen en temps (fortement) polynomial.* Le terme *combinatoire* fait référence au fait que la règle est définie en fonction du signe des mineurs de la matrice des coefficients du problème linéaire.

Ce dernier résultat est décrit dans le preprint [42].

English version

X. Allamigeon, S. Gaubert, and E. Goubault, have developed in [63], [16] algorithms allowing one to manipulate tropical polyhedra. They correspond to the contributions described in §6.2.1 . In particular, they can be used to determine the vertices and extreme rays of a tropical polyhedron defined as the intersection of half-spaces, or inversely, to compute an external description from a set of generators. These algorithms are implemented in the library TPLib (see §5.3).

In an ongoing work of X. Allamigeon, P. Benchimol, S. Gaubert and M. Joswig, we introduced a tropical analogue of the simplex algorithm, allowing one to solve problems of *tropical linear programming*, which are of the form (12), where the coefficients of the program, $a_{ij}^{\pm}$, $b_i^{\pm}$, c_j take their values in the max-plus semiring $\mathbb{R} \cup \{-\infty\}$. These problems are closely related to mean payoff games, as solving a game of this kind is equivalent to determine whether a tropical linear program admits a feasible point [57].

Like the classical simplex algorithm, the tropical simplex algorithm performs pivoting operations between basis points, until it reaches the optimum. The main discrepancy with the classical algorithm is that the pivoting is now a purely combinatorial operation, which is performed by using a local description of the polyhedron by a directed hypergraph. This allowed us to show that *a tropical pivoting step (including computing reduced costs) has the same complexity as in the classical simplex algorithm, i.e.* $O(n(m+n))$. This is all the more surprising as the tropical edge between two given points has a geometrically more complex structure in the tropical case (it is constituted of up to n ordinary line segments).

The tropical simplex algorithm turns out to be closely related to the classical one. Thanks to Tarski's principle, the latter is also valid for linear programs over the field $\mathbb{R}\{\{t\}\}$ of generalized Puiseux series in an indeterminate t. These series are of the form (13), where the α_i are real numbers, the coefficients c_{α_i} are non-zero reals, and the sequence $\alpha_1, \alpha_2, \cdots$ is strictly increasing and either finite or unbounded. The opposite of the smallest exponent of the series, $-\alpha_1$, is called *valuation*. A tropical linear program is said to be *lifted* to a linear program over $\mathbb{R}\{\{t\}\}$ if the valuation of the coefficients of the latter are sent to the coefficients of the former by the valuation. We showed the following relation between the classical simplex algorithm and its tropical analogue: *for all generic tropical linear program, the tropical simplex algorithm computes the image by the valuation of the path of the classical simplex algorithm, applied to any lift in* $\mathbb{R}\{\{t\}\}$ *of the original program*.

These results are gathered in the preprint [43]. They have been presented in several conferences [32], [33].

They allow one to relate the complexity of the classical simplex algorithm with the complexity of mean payoff games. The latter is unsettled, these games are known to be in the class $\mathsf{NP} \cap \mathsf{coNP}$ but it is not known whether they can be solved in polynomial time. Basic complexity issues regarding the classical simplex algorithm are also unsettled: its execution time depends on the pivoting rule, and many pivoting rules have been shown to have exponential worst case behaviors. The existence of a pivoting rule leading the simplex to terminate in polynomial time is still an open question.

In a second work, we related these two open questions, via the tropical simplex algorithm. We identified a class of pivoting rules, which are said to be *combinatorial*, and show that they have the following property: *if there is a combinatorial pivoting rule allowing one to solve every classical linear programming problem in polynomial time, then, mean payoff games can be solved in (strongly) polynomial time.* By *combinatorial*, we mean that the rule depends only of the coefficients of the system through the signs of minors of the coefficients matrix.

This result is given in the preprint [42].

6.4.3. Problèmes d'accessibilité dans les hypergraphes orientés et leur complexité/Reachability problems in directed hypergraphs and their complexity

Participant: Xavier Allamigeon.

Les hypergraphes orientés sont une généralisation des graphes orientés, dans lesquelles chaque arc relie un ensemble de sommets à un autre. Ils jouent un rôle important dans les travaux récents sur la convexité tropicale (voir §6.2.1), puisqu'ils offrent une représentation naturelle des cônes définis sur le sous-semi-anneau booléen $\mathbb{B} = \{-\infty, 0\}$.

Dans un travail de X. Allamigeon [15], on étudie la complexité de problèmes d'accessibilité sur les hypergraphes orientés. Nous introduisons un algorithme de complexité presque linéaire permettant de déterminer les composantes fortement connexes terminales (qui n'accèdent à aucune autre composante si ce n'est elles-mêmes) d'un hypergraphe.

Nous établissons également une borne inférieure sur-linéaire sur la taille de la réduction transitive de la relation d'accessibilité dans les hypergraphes. Cela indique que la relation d'accessibilité dans les hypergraphes orientés est combinatoirement plus complexe que celle des graphes orientés. Cela suggère aussi que des problèmes comme le calcul des composantes fortement connexes est plus difficile sur les hypergraphes que sur les graphes. Nous mettons d'ailleurs en évidence une réduction en temps linéaire du problème du calcul des ensembles minimaux dans une famille d'ensembles donnée, vers le problème du calcul de toutes les composantes fortement connexes d'un hypergraphe. Le problème du calcul des ensembles minimaux a été largement étudié dans la littérature [155], [175], [174], [156], [157], [158], [101], [69], et aucune algorithme en temps linéaire n'est connu à ce jour.

English version

Directed hypergraphs are a generalization of directed graphs, in which the tail and the head of the arcs are sets of vertices. It appears that they play an important role in the recent works on tropical convexity (see §6.2.1), since they offer a natural representation of cones defined over the boolean sub-semiring $\mathbb{B} = \{-\infty, 0\}$.

In a work of X. Allamigeon [15], we study the complexity of reachability problems on directed hypergraphs. We introduce an almost linear-time algorithm allowing to determine the terminal strongly connected components (a component is said to be *terminal* when no other component is reachable from it).

We also establish a super-linear lower bound over the size of the transitive reduction of the reachability relation in directed hypergraphs. This indicates that the reachability relation is combinatorially more complex in directed hypergraphs than in directed graphs. This also suggests that reachability problems such as computing all strongly connected components are likely to be harder in hypergraphs than in graphs. Besides, we show that the minimal set problem can be reduced in linear time to the problem of computing all strongly connected components in hypergraphs. The former problem consists in finding all minimal sets among a given family of sets. It has been well studied in the literature [155], [175], [174], [156], [157], [158], [101], [69], and no linear time algorithm is known.

6.4.4. Approximation max-plus de fonctions valeurs et équations de Riccati généralisées/Max-plus approximation of value functions and generalized Riccati equations

Participants: Stéphane Gaubert, Zheng Qu, Shanjian Tang [Fudan University, Shanghai].

La thèse de Zheng Qu, supervisée par S. Gaubert et S. Tang, a porté sur le développement de méthodes tropicales en programmation dynamique approchée [12].

Les méthodes d'approximation max-plus conduisent à approcher la fonction valeur d'un problème de contrôle ou de jeux par un supremum d'un nombre fini de formes quadratiques, voir notamment [113]. On s'intéresse ici à l'analyse théorique (complexité) ainsi qu'à l'amélioration de ces méthodes. Dans certains cas, ces formes quadratiques sont propagées par des flots d'équations de Riccati généralisées. Afin d'effectuer des analyses d'erreur, on exploite les propriétés de contraction du flot de Riccati pour certaines métriques connues sur le cône des matrices positives, et en particulier pour la métrique de Thompson. Celle-ci n'est rien d'autre que $d_T(A,B) = \|\log \mathrm{spec}\,(A^{-1}B)\|_\infty$, où spec désigne la suite des valeurs propres d'une matrice, et log s'entend composante par composante.

Ceci nous a amené à étudier le problème général du calcul du taux de contraction d'un flot monotone sur un cône, pour la métrique de Thompson. En effet, les propriétés de contraction de l'équation de Riccati standard sont connues (résultats de Bougerol pour la métrique Riemanienne invariante, et de Wojtowski pour la métrique de Thompson), mais les techniques de preuve employées dans ce cadre (semigroupes de matrices symplectiques) ne s'étendent pas aux équations généralisées.

On donne dans [114], [28] une formule explicite générale pour le taux de contraction pour la métrique de Thompson d'un flot monotone, faisant seulement intervenir le générateur du flot et sa dérivée. On a notamment appliqué ce résultat à une équation de Riccati généralisée associé à des problèmes de contrôle stochastique avec critère quadratique, dans lesquels la dynamique comporte un terme bilinéaire en le contrôle et le bruit. On a montré dans ce cas que la métrique de Thompson est la seule métrique de Finsler invariante pour laquelle le flot est nonexpansif, et l'on a caractérisé la constante de contraction locale.

Une application de ces résultats à l'analyse d'une méthode de réduction de la malédiction de la dimension, dûe à McEneaney, a été donnée dans [28], [50].

English version

The PhD work of Zheng Qu, supervised by S. Gaubert and S. Tang, dealt with the development of tropical methods in approximate dynamic programming [12].

The max-plus methods lead to approach the value function of an optimal control or zero-sum game problem by a supremum of a finite number of quadratic forms, see in particular [113]. We are interested here in the theoretical analysis (complexity) of this class of methods, as well as of their improvement. In certain cases, the quadratic forms are propagated by the flows of generalized Riccati equations. In order to perform an error analysis, we need to use some contraction properties of the Riccati flow, for certain known metrics on the space of positive matrices, like Thompson's metric. The latter is nothing but $d_T(A,B) = \|\log \mathrm{spec}\,(A^{-1}B)\|_\infty$, where spec denotes the sequence of eigenvalues of a matrix, and log is understood entrywise.

This led us to study the general problem of computing the contraction rate of an order-preserving flow on a cone, with respect to Thompson's metric. Indeed, the contraction properties of the standard Riccati flow are known (theorem of Bougerol for the invariant Riemanian metric, of Wojtowski for the Thompson's metric), but the proof of these properties (based on symplectic semigroups) does not carry over to generalized Riccati equations.

We gave in [114],[28] a general explicit formula for the contraction rate with respect to Thompson's metric of an order-preserving flow, involving only the generator of the flow and its derivative. We applied in particular this result to a generalized Riccati equation, associated to stochastic optimal control problems with a quadratic cost and a bilinear dynamics (presence of a bilinear term between the control and the noise). We showed that in this case, the Thompson's metric is the only invariant Finsler metric in which the generalized Riccati flow is nonexpansive, and we characterized the local contraction rate of this flow.

Z. Qu has applied these results in [28], [50] to the analysis of a method of reduction of the curse of dimensionality, introduced by McEneaney.

6.4.5. Points fixes d'applications monotones homogènes et jeux à somme nulle/Fixed points of order preserving homogeneous maps and zero-sum games

Participants: Marianne Akian, Stéphane Gaubert, Antoine Hochart.

Les opérateurs de Shapley sont les opérateur de programmation dynamique pour des jeux à somme nulle, ce sont précisément les opérateurs qui préservent l'ordre et commutent avec l'addition d'une constante. Le travail de M2 d'Antoine Hochart a traité d'une sous-classe d'opérateurs de Shapley, qui commutent en outre avec la multiplication par une constante positive. Nous les appellerons ici *sans-paiement*, car ils apparaissent dans des classes de jeux où les paiements instantanés sont nuls - le paiement a lieu seulement le dernier jour (*recursive games*). Ils apparaissent aussi dans l'étude structurelle de familles paramétriques de jeux répétés avec espace d'état fini et information parfaite, si l'on suppose par exemple que les probabilités de transitions sont fixées, mais que les paiements sont des paramètres. À toute famille paramétrique de jeux est associée un opérateur sans paiements et les points fixes de ce dernier sont précisément les vecteurs de paiement moyen réalisables. Un problème de base consiste à vérifier si un opérateur sans paiement n'a que des points fixes triviaux (réduits à des multiples du vecteur unité), et si possible, de déterminer des caractéristiques plus précises de l'ensemble des points-fixes, par exemple, savoir s'il existe un point fixe d'argmin donné. Le premier problème est connu être co-NP-complet, même pour des jeux déterministes. Nous montrons cependant que le second problème (point fixe d'argming prescrit) peut être résolu en temps polynomial. La preuve repose sur la construction d'une correspondance de Galois entre les faces d'un hypercube qui sont invariantes par l'opérateur, ainsi que sur une réduction à un problème d'accessibilité dans un hypergraphe orienté.

English version

Shapley operators are the dynamic programming operators of zero-sum stochastic games, they can be characterized as order preserving maps commuting with the addition of a constant. The M2 work of Antoine Hochart has dealt with a subclass of Shapley operators which are characterized by the property of commuting with the multiplication by a positive constant. We call them *payment-free*, as they arise in the study of *recursive games*, in which the payment only occurs when the game stops. They also arise in the study of structural properties of parametric mean payoff games (the transition probabilities are fixed, not the transition payoffs) with finite action spaces and perfect information: their fixed point set can be shown to give all the possible mean payoff vectors of such games. A basic problem is to check whether the fixed point set of such an operator is trivial (reduced to the multiples of the unit vector), and more precisely to determine its characteristics, for instance decide whether there is a fixed point with a prescribed argmin. The former problem is already known to be co-NP-complete, even for deterministic games. We showed however that the latter can be solved in polynomial time. The proof relies on the construction of a Galois connection between faces of the hypercube that are invariant by the operator, and on a reduction to a reachability problem in a directed hypergraph.

6.5. Applications

6.5.1. Introduction

Nous présentons maintenant plusieurs travaux de nature appliquée, touchant à des domaines variés, dans lesquels nous exploitons certaines des techniques mathématiques présentées précédemment, et particulièrement celles qui relèvent de la théorie de Perron-Frobenius non-linéaire et de la convexité tropicale. Ces applications utilisent aussi des techniques d'algèbre linéaire ou d'optimisation convexe.

English version

In this section, we describe several applied works in which we use some of the theoretical tools developed by the team, including non-linear Perron-Frobenius theory and tropical convexity. Some of these applications also make an intensive use of linear algebraic and convex programming methods.

6.5.2. Propriétés des valeurs propres de Perron et de Floquet, et application en chronothérapeutique/Properties of Perron and Floquet eigenvalue, with an application to chronotherapeutics

Participants: Frédérique Billy [Projet BANG, Inria], Jean Clairambault [Projet BANG, Inria], Olivier Fercoq, Stéphane Gaubert, Thomas Lepoutre [Projet BANG puis DRACULA, Inria].

On s'intéresse à des modèles de systèmes dynamiques monotones structurés en âge représentant la croissance de populations de cellules (saines ou tumorales), à la suite de travaux de Clairambault et Perthame. Il s'agit de comprendre l'influence du contrôle circardien sur la croissance des cellules. Dans le cas stationnaire, le taux de croissance est représenté par une valeur propre de Perron. Dans le cas périodique, il s'agit d'une valeur propre de Floquet. Les travaux [39], [73], [72] portent sur l'identification de ces modèles ainsi que sur un problème de contrôle thérapeutique, consistant à minimiser le taux de croissance des cellules tumorales sous une contrainte de non-toxicité du traitement (maintien d'une population de cellules saines). Ce travail s'appuie en particulier sur un algorithme d'optimisation de la valeur propre de Perron d'une matrice développé par Fercoq dans un autre contexte [104].

Un développement récent de ce travail peut être trouvé dans [39]. Un travail théorique sur ce type de modèles est présenté dans [46].

English version

We study monotone dynamical systems representing the growth of cells (healthy or tumoral), following a work of Clairambault and Perthame. The goal is to understand how the circadian control influences the growth of cells. In the case of stationnary monotone systems, this growth is measured by the Perron root. In the time periodic case, this Perron root is replaced by a Floquet multiplier.

The works [39], [73], [72] deal with the identification of these models, together with a therapeutic control problem, consisting in minimizing the growth rate of tumoral cells, under a non-toxicity constraint (preserving the population of healthy cells). This works relies in particular on a fast algorithm to optimize the Perron eigenvalue of a matrix, developed by Fercoq in a different context [104].

A recent development of this work can be found in [39]. A theoretical work on this kind of models has been presented in [46].

6.5.3. Preuve formelle d'inégalités non-linéaires/Formal proofs of non-linear inequalities

Participants: Xavier Allamigeon, Stéphane Gaubert, Victor Magron, Benjamin Werner [LIX].

La thèse de Victor Magron [11], dirigée par Benjamin Werner, codirigée par Stéphane Gaubert et Xavier Allamigeon, a porté sur la certification de bornes inférieures de fonctions multivariées à valeurs réelles, définies par des expressions semi-algébriques ou transcendantes, et sur la preuve de validité de celles-ci au moyen de certificats dans l'assistant de preuves Coq.

De nombreuses inégalités de cette nature apparaissent notamment dans la preuve par Thomas Hales de la conjecture de Kepler. Voici un exemple typique d'inégalité à prouver.

LEMME 9922699028 FLYSPECK. *Soit K, $\Delta\mathbf{x}$, l, t et f définis comme suit:*

$$
\begin{aligned}
K &:= [4, 6.3504]^3 \times [6.3504, 8] \times [4, 6.3504]^2 \ , \\
\Delta\mathbf{x} &:= x_1 x_4(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6) \\
&\quad + x_2 x_5(x_1 - x_2 + x_3 + x_4 - x_5 + x_6) \\
&\quad + x_3 x_6(x_1 + x_2 - x_3 + x_4 + x_5 - x_6) \\
&\quad - x_2 x_3 x_4 - x_1 x_3 x_5 - x_1 x_2 x_6 - x_4 x_5 x_6 \ , \\
l(\mathbf{x}) &:= -\pi/2 + 1.6294 - 0.2213(\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0) \\
&\quad + 0.913(\sqrt{x_4} - 2.52) + 0.728(\sqrt{x_1} - 2.0) \ , \\
t(\mathbf{x}) &:= \arctan \frac{\partial_4 \Delta\mathbf{x}}{\sqrt{4 x_1 \Delta\mathbf{x}}} \ , \\
f(\mathbf{x}) &:= l(\mathbf{x}) + t(\mathbf{x}) \ .
\end{aligned}
$$

Alors, $\forall \mathbf{x} \in K, f(\mathbf{x}) \geq 0$.

On s'est donc intéressé à des fonctions non-linéaires, faisant intervenir des opérations semi-algébriques ainsi que des fonctions transcendantes univariées ($\cos$, $\arctan$, $\exp$, etc).

De manière classique, on peut approcher les fonctions transcendantes qui interviennent de la sorte par des polynômes, ce qui permet de se ramener à des problèmes d'optimisation semi-algébriques, que l'on peut résoudre par des techniques de sommes de carrés creuses conduisant à des problèmes SDP. Cependant, en pratique, cette approche est limitée par la taille des SDP à résoudre, qui croît rapidement avec le degré des approximations polynomiales.

Dans ce travail de thèse, on a développé une méthode alternative, qui consiste a borner certains des constituants de la fonction non-linéaire par des suprema de formes quadratiques dont les Hessiens sont judicieusement choisis. On reprend donc ici l'idée des approximations "max-plus" initialement introduites en contrôle optimal, en s'appuyant sur des techniques d'interprétation abstraite (généralisation non-linéaire de la méthode des gabarits de Manna et al.). Ainsi, on obtient une nouvelle technique d'optimisation globale, basée sur les gabarits, qui exploite à la fois la precision des sommes de carrés et la capacité de passage à l'échelle des méthodes d'abstraction.

L'implémentation de ces méthodes d'approximation a abouti à un outil logiciel : `NLCertify`. Cet outil génère des certificats à partir d'approximations semi-algébriques et de sommes de carrés. Son interface avec Coq permet de bénéficier de l'arithmétique certifiée disponible dans l'assistant de preuves, et ainsi d'obtenir des estimateurs et des bornes valides pour chaque approximation.

Les performances de cet outil de certification ont été démontrées sur divers problèmes d'optimisation globale ainsi que sur des inégalités essentiellement serrées qui interviennent dans la preuve de Hales (projet Flyspeck).

Ce travail est exposé dans [25], [26].

English version

The PhD work of Victor Magron [11], supervised by Benjamin Werner, and cosupervised by Stéphane Gaubert and Xavier Allamigeon, dealt with the certification of lower bounds for multivariate functions, defined by semi-algebraic or transcendental expressions, and their correctness proof through certificates checked in the Coq proof assistant.

Many inequalities of this kind appear in particular in the proof by Thomas Hales of Kepler's conjecture. Here is a typical example of inequality.

LEMMA 9922699028 FLYSPECK. *Let K, $\Delta\mathbf{x}$, l, t and f be defined as follows:*

$$\begin{aligned}
K &:= [4, 6.3504]^3 \times [6.3504, 8] \times [4, 6.3504]^2 \ , \\
\Delta\mathbf{x} &:= x_1x_4(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6) \\
&\quad +x_2x_5(x_1 - x_2 + x_3 + x_4 - x_5 + x_6) \\
&\quad +x_3x_6(x_1 + x_2 - x_3 + x_4 + x_5 - x_6) \\
&\quad -x_2x_3x_4 - x_1x_3x_5 - x_1x_2x_6 - x_4x_5x_6 \ , \\
l(\mathbf{x}) &:= -\pi/2 + 1.6294 - 0.2213(\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0) \\
&\quad +0.913(\sqrt{x_4} - 2.52) + 0.728(\sqrt{x_1} - 2.0) \ , \\
t(\mathbf{x}) &:= \arctan \frac{\partial_4 \Delta\mathbf{x}}{\sqrt{4x_1\Delta\mathbf{x}}} \ , \\
f(\mathbf{x}) &:= l(\mathbf{x}) + t(\mathbf{x}) \ .
\end{aligned}$$

Then, $\forall \mathbf{x} \in K, f(\mathbf{x}) \geq 0$.

Thus, we considered non-linear functions, defined in terms of semi-algebraic operations and univariate transcendental functions ($\cos$, $\arctan$, $\exp$, etc).

Such transcendental functions can be classically approximated by polynomials, which leads to semi-algebraic optimization problems, which can be solved by sparse sum of squares techniques leading to SDP formulations. However, in practice, this approach is limited by the growth of the size of the SDP instances to be solved, whichs grows quickly with the degree of polynomial approximations.

In this PhD, we developed an alternative method, which consists in bounding some constituents of the non-linear function to be optimized by suprema of quadratic forms with well chosen Hessians. This is based on the idea of “maxplus approximation” initially introduced in optimal control, and also, on abstract interpretation (the template method introduced by Manna et al. in static analysis). In this way, we end up with a new global optimization technique, which takes advantage of the precision of sum of squares and of the scalability of abstraction methods.

These methods have been implemented in a software tool: `NLCertify`. This tool generates certificates from semi-algebraic and sum of square certificates. Its interface with Coq allows one to take benefit of the certified arithmetics available in this proof assistant, and so, to obtain estimators and valid bounds for each approximation.

The performances of this certification tool have been shown on several global optimization problems from the litterature, as well as on essentially tight inequalities taken from Hales’ proof (Flyspeck project).

This work is presented in [25], [26].

6.5.4. Vérification de systèmes temps-réels/Verification of real-time systems

Participants: Xavier Allamigeon, Uli Fahrenberg [IRISA], Stéphane Gaubert, Ricardo Katz [Conicet], Axel Legay [IRISA].

Dans [141], Lu, Madsen, Milata, Ravn, Fahrenberg et Larsen ont montré que les polyèdres tropicaux peuvent être utilisés dans le cadre de l’analyse d’accessibilité d’automates temporisés. En effet, les polyèdres tropicaux expriment naturellement des invariants non-convexes, qui sont en fait des disjonctions d’invariants fournis par des DBM (*difference bound matrices*). A ce titre, les polyèdres tropicaux devraient permettre de réduire le nombre de disjonctions réalisées pendant l’analyse d’automates temporisés. Une limitation importante de cette approche est cependant que les polyèdres tropicaux sont topologiquement fermés, et qu’ils ne peuvent donc pas exprimer de contraintes d’inégalités strictes. Ces dernières sont néanmoins fondamentales dans l’analyse de systèmes temps-réels.

Nous avons donc développé dans [44] une généralisation des polyèdres tropicaux permettant d'exprimer des contraintes mixtes, *i.e.* strictes ou larges. Notre approche repose sur l'utilisation d'inégalités tropicales linéaires à coefficients dans un (quotient du) semi-anneau de germes affines. Afin de réaliser des opérations sur cette nouvelle classe de polyèdres tropicaux, nous avons défini deux nouveaux algorithmes. Le premier est un analogue tropical de l'élimination de Fourier-Motzkin. Celle-ci s'applique plus généralement à des systèmes d'inégalités linéaires sur des semi-anneaux idempotents et totalement ordonnés. Le second algorithme permet de tester si un système de contraintes mixtes admet une solution. Nous montrons en effet que ce problème est équivalent en temps polynomial à la résolution d'un problème de jeux déterministes à somme nulle. Ces deux contributions nous permettent de définir les primitives requises pour l'analyse d'accessibilité d'automates temporisés.

English version

Lu, Madsen, Milata, Ravn, Fahrenberg and Larsen have shown in [141] that tropical polyhedra can be applied to the reachability analysis of timed automata. Indeed, tropical polyhedra naturally express non-convex invariants, which correspond to disjunctions of invariants provided by DBM (*difference bound matrices*). Consequently, tropical polyhedra should allow to reduce the number of disjunctions arising during the analysis of timed automata. An important limitation of this approach is that tropical polyhedra are topologically closed, and thus they cannot express strict inequality constraints. However, such constraints plays an important role in the analysis of real-time systems.

As a result, we have developed in [44] a generalization of tropical polyhedra, in order to express mixed constraints, *i.e.* strict or loose ones. Our approach relies on tropical linear inequalities with coefficients in a (quotient of) the semiring of affine germs. In order to perform operations on this new class of polyhedra, we have introduced two new algorithms. The first one is a tropical analog of Fourier-Moztkin elimination. In fact, it applies more generally to systems of linear inequalities over totally ordered and idempotent semirings. The second algorithm allows to test the feasability of a mixed constraint system. We indeed show that this problem is polynomial-time equivalent to solving mean payoff games. These two contributions allow to define the primitives required by the reachability analysis of timed automata.

MC2 Project-Team

6. New Results

6.1. Cancer modelling

We have improved our generic mathematical models describing tumor growth. These models were then specialized for several types of cancer (thyroidal lung nodules, brain tumors). The algorithm used to recover the parameters of these models from medical images has also been greatly improved and is now adapted to run on HPC architectures.

- Secondary tumors in the lung:
  The mathematical models describing the growth of secondary in the lungs have now settled and are well understood. The main focus of the year was to keep on using these models on patient data. New clinical case were selected by clinicians from the Institut Bergonié, there are currently under study. The model is currently able to reproduce the growth observed on 5 clinical cases. Huge improvements to the calibration algorithms were made. The initial seeding of the algorithms was a weak point of the procedure and the robustness regarding the time-derivative of the observations is now much more accurate. A complete rewrite of the routines was done to improve their versatility and efficiency. Previously, the numerical simulations and calibration were performed in 2D (clinicians selected the most relevant slice showing the evolution of the tumor). Work is now ongoing to switch to full 3D computations and calibration. A newly hired engineer is testing our calibration technique on a dozen of clinical cases.
- Metastasis to the liver of a GIST
  We have derived a continuous model describing the growth of a GIST metastasis to the liver treated with Glivec and Sutent. This model is able to qualitatively reproduce the evolution observed on two different patients. Work has also started on developing new markers computed from the texture of the tumor seen on images to be able to detect any change in the response to treatment. This was the subject of an internship in the General Electric Healthcare research center. The results are promising so far.
- Modeling glioblastomas:
  In 2011, a hierarchy of models describing the growth of brain tumors was developed (and described in a submitted paper) in collaboration with University of Alabama at Birmingham. As we wished to obtain models that could be calibrated from patient data and yet be reasonably accurate, we believe that these models are suitable trade-offs between the simplicity of the Swanson's model (the only one used on patient data of brain tumors so far) and the accuracy of more complex models (that cannot really produce quantitative results). We have derived a new model that allows us to study the efficency of anti-angiogenic therapies. It seems to predict that the efficacy of these treatments is limited, this could be confirmed by a world-wide ongoing clinical study. Work is ongoing with this model to develop new marker to quantify patient anti-angiogenic drugs as soon as possible. This collaboration will be made stronger by a new Phd in UAB co-advised by T. Colin.
- Modelling of electrochemotherapy :
  Two articles related to the electrical cell modelling have been done ([59], [56]) . The first one deals with the influence of the ionic fluxes on the transmembrane voltage potential and on the cell volume. The main insight of the results consists in linking the transmembrane potential with the cell volume: it has been observed experimentally that cells with a low voltage potential do divide, whereas cells with high voltage potential do not, and the obtained relationship between voltage potential and cell volume can provide an explanation. The second article deals with a new model of cell electroporation essentially based on the experimental results of the I.G.R. In this paper we describe precisely the model, which takes into account the main experimental results in the electroporation process, and we present a variationnal formulation inherent to the model that leads to new efficient schemes in order to numerically solve the involved P.D.E.

The article describing a new electrical model of classical has been published in Journal of Math Biology. This new phenomenological model involves much less parameters than the usual models, but it still provides the qualitatively good description of the electroporation. The main feature of this model lies in the fact that it provides an intrinsic behavior of the cell membrane, which seems in accordance with the preliminary experimental results of the IGR partner. We also adapted the finite difference method developed by L. Weynans and M. Cisternino for elliptic interface problems to the electropermeabilization model developed recently by C. Poignard with O. Kavian. The new method has been validated by convergence tests and comparison with other models. We have proven that in one dimension the numerical solution converges to the solution of the exact problem. A paper describing these results has been submitted. A second-order Cartesian method for the simulation of electropermeabilization cell models, Leguebe M., Poignard C., Weynans L., Inria Resaerch Report RR-8302).

- Cell Migration modelling:
  The collaboration with IECB (University of Bordeaux) has continued with the postdocatoral position of Julie Joie. We have obtain a continuous model of cell density evolving on micropatterned polymers. The research report RR 7998 will be published in Math. Biosci. and Eng. A discrete model describing the single cells motility is being written.

  We also have started a collaboration with the University of Osaka (Japan), thanks to a PHC Sakura project, on the invadopodia. C. Poignard has been invited at Osaka in february by Prof. Suzuki. A model describing the destruction of the extracellular matrix by the MMP enzyme, and then the cell migration has been obtained. R. Mahumet, a PhD student of Prof. Suzuki is developing a code to simulate the model.
- Adaptive radiotherapy: a new work is also ongoing with Institut Bergonié to quantify the movement and deformations of organs of patients with sarcoma treated with radiotherapy. The preliminary results highlight that these mouvements are much larger than clinicians expected. We are now working on improving our workflow and developing a new segmentation technique to have this monitoring automatically performed (for the time being clinicians have to delineate each structure which is a very time consuming task). The ultimate goal is to change the therapeutical protocol to take these movements into account (which is currently not the case). Preliminary contacts have been made with a company developing a dose computing software to evaluate the efficacy of an adaptive planning of the Radio-Therapy compared to the constant dose plan currently given to the patient.
- Modeling meningioma growth and their responses to radiotherapy: In collaboration with Institut Bergonié, we have started developing new models for studying meningioma. Our model is able to reproduce the characteristic shape of these tumors which is in itself a very satisfactory result. Furthermore, from this model we have derived a simpler non-spatial model able to reproduce the 4 different types of response to radiotherapy observed by clinicians on the more than 70 patients they have selected for this study.
- Theoretical biology of the metastatic process
  We proposed a theoretical study of systemic inhibition of angiogenesis (SIA) in a population of tumors by deriving a model from biophysical considerations and simulating a novel mathematical model able to describe the development of a population of tumors in mutual inhibitory interactions at the organism scale. We showed that the model could explain experimental data on metastatic development and tested the hypothesis of global dormancy (cancer without disease) resulting from the net inhibitory action of stimulatory and inhibitory signaling interactions among the lesions comprising the total tumor burden. We found SIA alone is not sufficient for global dormancy but could suppress the growth of the total metastatic burden. See [41]. The resulting model is a nonlinear partial differential transport equation with nonlocal boundary condition that describes organism-scale population dynamics under the influence of three processes: birth (dissemination of secondary tumors), growth and inhibition (through angiogenesis). The asymptotic behavior of the model was numerically investigated in a second publication [40] and revealed interesting dynamics ranging

from convergence to a steady state to bounded non-periodic or periodic behaviors, possibly with complex repeated patterns.

In link with the previous study and in order to base the modeling on relevant biological data studying the basic phenomenon of growth interactions among tumors, we first performed a rational, quantitative and discriminant analysis of the descriptive and predictive properties of classical ODE-based models for tumor volume kinetics, which has been summarized in a publication that is currently under revision for PloS Computational Biology [42].

A collaboration with John Ebos from the Roswell Park Cancer Institute in Buffalo (NY, USA) has been initiated that deals with the objectives to quantify metastatic aggressiveness of several cancer cell linesand to rationally define a neo-adjuvant (i.e., prior to resection of the primary lesion) efficacy score of several anti-cancer chemical agents. The PhD of Etienne Baratchart has been initiated, in close collaboration with the "Angiogenesis and cancer microenvironment laboratory" directed by the Pr Andreas Bikfalvi, about the initiation, development and role of the pre-metastatic and metastatic niche.

6.2. Newtonian fluid flows simulations and their analysis

- Simulations of water distribution systems :Water losses may constitute a large amount of the distributed total water volume throughout water distribution systems. Here, a new model method is proposed that intends to minimize the total water volume distributed through leakage reduction. Our group has worked on the derivation of advection-reaction-diffusion type equations with an explicit relationship between the local pressure and the leakage rate. An original splitting technique to solve this type of hydraulic problem was then achieved. This technique allows pressure-dependent leakage to be taken into account, whereas in most models leakage is assumed to be uniform along a pipe. Finally, a constrained optimization problem was formulated for leakage reduction in WDS. The control variable had the mean of a local head loss and is considered in the Boundary Conditions to avoid dealing with discontinuities in the governing equations. The objective function to minimize was a regularization of the total water volume distributed. Specific operational constraints were added to ensure enough pressure at consumption points. The direct solution for this minimization problem was sought with a Gradient type method. The leakage reduction was proven to be significant in a case study. The percentage of leakage reduced from 24% to 10% in the linear relationship between pressure and leakage flow rate. With other leakage exponents, the same rate of reduction was achieved . The method was applied on a real network in the South-West of France. Controlling the pressure at two different strategic points permits a significant amount of the total distributed water to be saved (5%). This work was performed in collaboration with Cemagref Bordeaux . Future work will consist of applying a sensibility analysis of control location points to optimize the method.
- Incompressible flows : modeling and simulation of moving and deformable bodies. The incompressible Navier-Stokes equations are discretized in space onto a fixed cartesian mesh. The deformable bodies are taken into using a first order penalization method and/or second order immersed boundary method. The interface between the solid and the fluid is tracked using a level-set description so that it is possible to simulate several bodies freely evolving in the fluid. A turbulence model based on Samgorinsky model has been added to the numerical code. The numerical code written in the C langage is massively parallel. The large linear systems (over than 100 millions of dofs) are solved using the Petsc Library. As an illustration of the methods, fish-like locomotion is analyzed in terms of propulsion efficiency. Underwater maneuvering and school swimming are also explored. We were able to simulate the three-dimensional flow about a swimmer for realistic physical configurations. Another application is the turbulent 3D flow around complex wind turbine (see http://www.math.u-bordeaux1.fr/~mbergman and http://www.math.u-bordeaux1.fr/MAB/mc2/analysis.html for simulation movies). Wake flows generated by boat propellers are also modeled and simulated.
  We recently take in account a simplified elasticity model of the swimmer (elastic caudal tail of a fish). Some elastic parameters allows to increase the swimming efficiency around 20%-30%.

Recent developments on multiphase flows have been performed. We are able to simulate water/air interactions with interface regularization. The interface with a boat is also taken into account. See http://www.math.u-bordeaux1.fr/~mbergman for simulations.
We are also able to compute the strong (implicit) fluid structure interactions between the fluid and an elastic medium. For instance we have simulated a fish with an elastic tail and highlighted the fact that a given flexibility of tail allows to increase significantly the swimming efficiency.

- Turbulence flow on an hemisphere : Participants: Charles-Henri Bruneau, Patrick Fischer (MCF Bordeaux 1), Yong Liang Xiang (PostDoc)
ANR Cyclobulle lead by Hamid Kellay Soap hemi-bubble film experiments have shown some links between the formation of vortices when the hemi-bubble is heated at the equator and the formation of tornados in the earth atmosphere. Two-dimensional simulations using a stereographic map are used to compare to these experimental results and confirm the results when Coriolis force and heat source terms are added.
- Compressible flows: Immersed boundary methods. We are concerned with immersed boundary methods, i.e., integration schemes where the grid does not fit the geometry, and among this class of methods, more specifically with cartesian grid methods, where the forcing accounting for the presence of boundaries is performed at the discrete level. We have developed a simple globally second order scheme inspired by ghost cell approaches to solve compressible flows, inviscid as well as viscous. In the fluid domain, away from the boundary, we use a classical finite-volume method based on an approximate Riemann solver for the convective fluxes and a centered scheme for the diffusive term. At the cells located on the boundary, we solve an ad hoc Riemann problem taking into account the relevant boundary condition for the convective fluxes by an appropriate definition of the contact discontinuity speed. This method can easily be implemented in existing codes and is suitable for massive parallelization. It has been validated in two dimensions for Euler and Navier-Stokes equations, and in three dimensions for Euler equations. The order of convergence is two in L^2 norm for all variables, and between one and two in L^∞ depending on the variables. The 3D code has been parallelized with MPI. The case of a moving solid has been tested (flapping wing) and gives results for the drag and the lift in agreement with the references in the literature.
The Oldroyd B constitutive model is used to study the role of the viscoelasticity of dilute polymer solutions in two-dimensional flows past a bluff body using numerical simulations. This investigation is motivated by the numerous experimental results obtained in quasi two dimensional systems such as soap film channels. The numerical modeling is novel for this case and therefore a comprehensive comparison is carried out to validate the present penalization method and artificial boundary conditions. In particular we focus on flow past a circular object for various values of the Reynolds number, Weissenberg number, and polymer viscosity ratio. Drag enhancement and drag reduction regimes are discussed in detail along with their flow features such as the pattern of vortex shedding, the variation of lift as well as changes in pressure, elongational rates, and polymer stress profiles. A comprehensive study of the flow behavior and energy balance are carefully carried out for high Reynolds numbers. Flow instabilities in both numerical and experimental results are discussed for high Weissenberg numbers .
- Elliptic problems: We have developed a new cartesian method to solve elliptic problems with immersed interfaces. These problems appear in numerous applications, among them: heat transfer, electrostatics, fluid dynamics, but also tumour growth modelling, or modelling of electric potential in biological cells This method is second order accurate in the whole domain, notably near the interface. The originality of the method lies on the use of additionnal unknows located on interface points, on which are expressed flux equalities. Special care is dedicated to the discretization near the interface, in order to recover a stable second order accuracy. Actually, a naive discretization could lead to a first order scheme, notably if enough accuracy in the discretization of flux transmission condtions is not provided. Interfaces are represented with a distance level-set function discretized on the grid points. The method has been validated on several test-cases with complex interfaces in 2D. A parallel version has been developed using the PETSC library.

- Simulations of fluid-solid interactions : The interaction of an elastic structure and an fluid occurs in many phenomena in physics. To avoid the difficulty of coupling lagrangian elasticity with an eulerian fluid we consider a whole eulerian formulation. The elasticity of the structure is computed with retrograde caracteristics which satisfy a vectorial transport equation. We derive the associated fluid-structure models for incompressible and compressible media. The equations are discretized on a cartesian mesh with finite differences and finite volumes schemes. The applications concern the bio-locomotions and the study of air-elastic interaction.
- Vortex methods : The aim of this work is to couple vortex methods with the penalization methods in order to take advantage from both of them. This immersed boundary approach maintains the efficiency of vortex methods for high Reynolds numbers focusing the computational task on the rotational zones and avoids their lack on the no-slip boundary conditions replacing the vortex sheet method by the penalization of obstacles. This method that is very appropriate for bluff-body flows is validated for the flow around a circular cylinder on a wide range of Reynolds numbers. Its validation is now extended to moving obstacles (axial turbine blades) and three-dimensional bluff-bodies (flow around a sphere). See [72]. Moreover, using the global properties of the penalization method, this technique permits to include porous media simultaneously in the flow computation. We aim to adapt the porous media flows to our new method and to apply it in order to implement passive control techniques using porous layers around bluff-bodies.
- Domain decomposition : Domain decomposition methods are a way to parallelize the computation of numerical solutions to PDE. To be efficient, domain decompositions methods should converge independently on the number of subdomains. The classical convergence result for the additive Schwarz preconditioner with coarse grid is based on a stable decomposition. The result holds for discrete versions of the Schwarz preconditioner, and states that the preconditioned operator has a uniformly bounded condition number that depends only on the number of colors of the domain decomposition, and the ratio between the average diameter of the subdomains and the overlap width. Constants are usually non explicit and are only asserted to depend on the "shape regularity" of the domain decomposition.
  two years ago, we showed the result holds the additive Schwarz preconditioner can also be defined at the continuous level and provided completely explicits estimates. Last year, we established that a similar result also holds for non shape regular domain decompositions where the diameter of the smallest subdomain is significantly smaller than the diameter of the largest subdomain. The constants are also given explicitely and are independent of the ratio between the diameter of the largest sudomain and the diameter of the smallest subdomain.
  This year, we have studied explored new coarse spaces algorithms for domain decomposition methods. Coarse spaces are necessary to get a scalable algorithm whose convergence speed does not deteriorate when the number of subdomains increases. For domains decomposition methods with discontinuous iterates, we showed that continuous coarse spaces can never be an optimal choice. As an alternative, we introduced both the use of discontinuous coarse spaces(DCS) and a new coarse space algorithm using these discontinuous coarse spaces.

6.3. Flow control and shape optimization

- Flow control : Participants: Charles-Henri Bruneau, Iraj Mortazavi, Emmanuel Creusé (Lille), Patrick Gilliéron (Paris).

  An efficient active control of the two- and three-dimensions flow around the 25 degrees rear window Ahmed body has been performed. A careful theoretical and numerical study of the trajectories of the vortices allows to adapt the control in order to improve its efficiency and get a better drag reduction.

MCTAO Project-Team

6. New Results

6.1. Optimal control for quantum systems: the contrast problem in NMR

These studies aim at optimizing the contrast in Nuclear Magnetic Resonance imaging using advanced optimal control. As said in section 4.2 , our work on this problem is based on experiments conducted in Prof. S. Glaser in Munich, see [29].

6.1.1. Theoretical aspects

Participants: Bernard Bonnard, John Marriott, Monique Chyba [University of Hawaii], Gautier Picot [University of Hawaii], Olivier Cots, Jean-Baptiste Caillau.

This is done in collaboration with University of Hawaii, and deals with many theoretical aspects of the contrast problem in NMR: analysis of the optimal flow [5], feedback classification in relation with the relaxation times of the species [10], [4]. John Marriott defended his PhD thesis on this topic, on August 28, 2013.

6.1.2. Numerical aspects

Participants: Bernard Bonnard, Jean-Baptiste Caillau, Olivier Cots, Mathieu Claeys [LAAS CNRS, Toulouse], Pierre Martinon [COMMANDS team, Inria].

We performed, in a collaboration with Pierre Martinon (COMMANDS team, Inria) and Mathieu Claeys (LAAS CNRS, Toulouse), a thorough comparison of the various available numerical methods in optimal control on this important physical problem. Direct and indirect methods (implemented in the Bocop and Hampath sofwares) were tested in the contrast problem, and LMI techniques were used to obtain global bounds on the extremum (in the contrast problem there are many local optima and the global optimality is a complicated issue). This successful collaboration is accounted for in [15] and was presented at the CDC conference [12]

6.2. Conjugate and cut loci computations and applications

Participants: Bernard Bonnard, Olivier Cots, Jean-Baptiste Caillau.

One of the most important results obtained by B. Bonnard and his collaborators concern the explicit computations of conjugate and cut loci on surfaces. This has obvious applications in optimal control to compute the global optimum; it is also relevent in optimal transport where regularity properties of the transport map in the Monge problem is related to convexity properties of the tangent injectivity domains.

In [3], we complete the previous results obtained in [27] (we bring them from ellipsoids to general revolutions surfaces).

The conjugate and cut loci in Serret-Andoyer metrics and dynamics of spin particles with Ising coupling are analized in [7], this is a first step towards the computation of conjugate and cut loci on left invariant Riemannian and sub Riemannian metrics in S0(3) with applications for instance to the attitude control problem of a spacecraft.

An analysis of *singular* metrics on revolution surfaces, motivated by the average orbital transfer problem when the thrust direction is restricted, is proposed in [2].

Finally, [8] determines cut and conjugate loci in an enegy minimizing problem that is related to the quantum systems mentionned in the first paragraph of section 4.2 .

6.3. Averaging in control

Participants: Bernard Bonnard, Helen-Clare Henninger, Jean-Baptiste Pomet.

A paper on the construction and properties of an "average control system" [1] appeared this year, it is based on Alex Bombrun's doctoral work (2007). It connects solutions of highly oscillating control systems to those of an average control system, when the frequency of oscillation goes high. It also gives a better ground to averaging for minimum time.

This average system in the case of minimum time for low thrust orbit transfer in the two body problem is currently being explored, in particular the study of its inherent singularities. In [16] we give some properties of this system, like geodesic convexity, and compare it with the one obtained for minimum energy, and Helen Henninger's PhD aims at going further in this direction and then apply this local study to real missions, possibly in a three-body environment.

6.4. Optimal transport

Participants: Ludovic Rifford, Alice Erlinger, Alessio Figalli [U. of Texas at Austin, USA], Thomas Gallouet [Inria, SIMPAF team], Bernard Bonnard, Jean-Baptiste Caillau, Lionel Jassionesse, Robert Mc Cann [U. of Toronto].

- The very general condition for continuity of the transport map given in [41] motivated exploration of conditions for convexity of the tangent injectivity domain [42], [3]. Lionel Jassionnesse's PhD is in part devoted to Ma-Trudinger-Wang tensor that also plays an important role in this matter. Ludovic Rifford has an ongoing collaboration with Alession Figalli and Thomas Gallouet on the link between this MTW tansor and the convexity of injectivity domains; They already improved a result by Loeper and Villani (the preprint "Ma-Trudinger-Wang condition vs. convexity of injectivity domains" is available from the authors) and aim at proving a conjecture due to Villani, that would hold in the case of anlaytic surfaces.
- The goal of Alice Erlinger's PhD, joint with University of Toronto, is to explore Optimal Transport's application to modeling in economics. She unfortunately stopped her PhD, but some results have already been obtained.

6.5. Applications of control methods to dynamical systems

Participants: Gonzalo Contreras, Alessio Figalli, Ayadi Lazrag, Ludovic Rifford, Raffael Ruggiero.

Ludovic Rifford and collaborators have been applying with success, techniques from geometric control theory to open problems in dynamical systems, mostly on genericity properties and using controllability methods to build suitable perturbations.

This has been applied to closing geodesics [64]. Ayadi Lazrag's PhD also deals with such problems; applying techniques close to these in [65], one goal is to establish a version of Francks' lemma for geodesic flows and to apply this to persitence problems. The approach relies on control theory results, with order 2 conditions. See [18] and another preprint ("Franks' lemma for C^2-Mañé perturbations of Riemannian metrics and applications to persistence" by Lazrag, Rifford and Ruggiero, available from the authors).

In [17], a non trivial conjecture on generic hyperbolicity of the so-called Aubry set of a Hamiltonian is solved on compact surfaces and in the C^2 topology (for genericity).

MESCAL Project-Team

6. New Results

6.1. Simulation

6.1.1. Simulation of Parallel Computing Systems

Researchers in the area of distributed computing conduct many of their experiments in simulation. While packet-level simulation is often used to study network protocols, it can be too costly to simulate network communications for large-scale systems and applications. The alternative chosen in SimGrid and a few other simulation frameworks is to simulate the network based on less costly flow-level models. Surprisingly, in the literature, validation of these flow-level models is at best a mere verification for a few simple cases. Consequently, although distributed computing simulators are widely used, their ability to produce scientifically meaningful results is in doubt. In [13] we focus on the validation of state-of-the-art flow-level network models of TCP communication on Wide Area Networks, via comparison to packet-level simulation. While it is straightforward to show cases in which previously proposed models lead to good results, instead we systematically seek cases that lead to invalid results. Careful analysis of these cases reveal fundamental flaws and also suggest improvements. One contribution of this work is that these improvements lead to a new model that, while far from being perfect, improves upon all previously proposed models. A more important contribution, perhaps, is provided by the pitfalls and unexpected behaviors encountered in this work, leading to a number of enlightening lessons. In particular, this work shows that model validation cannot be achieved solely by exhibiting (possibly many) ”good cases.” Confidence in the quality of a model can only be strengthened through an invalidation approach that attempts to prove the model wrong.

The previous results assume steady-state and provide thus a reasonable model when message size is very large. Although, such assumptions may be reasonable when studying grid applications, when simulating HPC applications message sizes are often much smaller and phenomenon like slow-start or how communications and computations overlap have to be accurately modeled. Simulation and modeling for performance prediction and profiling is yet essential for developing and maintaining HPC code that is expected to scale for next-generation exascale systems. In [15], [34] we describe an implementation of a flow-based hybrid network model that accounts for factors such as network topology and contention, which are commonly ignored by the LogP models. Although, this may seem like a strange choice, we focus on large-scale, Ethernet-connected systems, as these currently compose 37.8% of the TOP500 index, and this share is expected to increase as higher-speed 10 and 100GbE become more available. Furthermore, the European Mont-Blanc project to study exascale computing by developing prototype systems with low-power embedded devices will also use Ethernet-based interconnect [28]. Our model is implemented within SMPI, an open-source MPI implementation that connects real applications to the SimGrid simulation framework. SMPI provides implementations of collective communications based on current versions of both OpenMPI and MPICH. SMPI and SimGrid also provide methods for easing the simulation of large-scale systems, including shadow execution, memory folding, and support for both online and offline (i.e., post-mortem) simulation. We validate our proposed model by comparing traces produced by SMPI with those from real world experiments, as well as with those obtained using other established network models. Our study shows that SMPI has a consistently better predictive power than classical LogP-based models for a wide range of scenarios including both established HPC benchmarks and real applications.

6.1.2. Perfect Simulation

Perfect simulation is a very efficient technique that uses coupling arguments to provide a sample from the stationary distribution of a Markov chain in a finite time without ever computing the distribution. In [7], we consider Jackson queueing networks (JQN) with finite buffer constraints and analyze the efficiency of sampling from their stationary distribution. In the context of exact sampling, the monotonicity structure of JQNs ensures that such efficiency is of the order of the coupling time (or meeting time) of two extremal sample paths. In the context of approximate sampling, it is given by the mixing time. Under a condition on the drift of the stochastic process underlying a JQN, which we call *hyper-stability*, in our main result we show that the coupling time is polynomial in both the number of queues and buffer sizes. Then, we use this result to show that the mixing time of JQNs behaves similarly up to a given precision threshold. Our proof relies on a recursive formula relating the coupling times of trajectories that start from network states having 'distance one', and it can be used to analyze the coupling and mixing times of other Markovian networks, provided that they are monotone. An illustrative example is shown in the context of JQNs with blocking mechanisms.

In [35], we extend the technique to handle situations with infinite space state. We consider open JQN with losses with mixed finite and infinite queues and analyze the efficiency of sampling from their exact stationary distribution. Although the underlying Markov chain may have an infinite state space, we show that perfect sampling is possible. The main idea is to use a JQN 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 for acyclic or hyperstable networks. 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. We also extend our approach to non-monotone networks such as queueing networks with negative customers.

6.2. Interactive Analysis and Visualization of Large Distributed Systems

6.2.1. Interactive Visualization

High performance applications are composed of many processes that are executed in large-scale systems with possibly millions of computing units. A possible way to conduct a performance analysis of such applications is to register in trace files the behavior of all processes belonging to the same application. The large number of processes and the very detailed behavior that we can record about them lead to a trace size explosion both in space and time dimensions. The performance visualization of such data is very challenging because of the quantities involved and the limited screen space available to draw them all. If the amount of data is not properly treated for visualization, the analysis may give the wrong idea about the behavior registered in the traces.

In [33], we detail data aggregation techniques that are fully configurable by the user to control the level of details in both space and time dimensions. We also present two visualization techniques that take advantage of the aggregated data to scale. These features are part of the Viva and Triva open-source tools and framework.

The performance of parallel and distributed applications is also highly dependent on the characteristics of the execution environment. In such environments, the network topology and characteristics directly impact data locality and movements as well as contention, which are key phenomena to understand the behavior of such applications and possibly improve it. Unfortunately few visualizations available to the analyst are capable of accounting for such phenomena. In [26], we propose an interactive topology-based visualization technique based on data aggregation that enables to correlate network characteristics, such as bandwidth and topology, with application performance traces. We claim that such kind of visualization enables to explore and understand non trivial behavior that are impossible to grasp with classical visualization techniques. We also claim that the combination of multi-scale aggregation and dynamic graph layout allows our visualization technique to scale seamlessly to large distributed systems. We support these claims through a detailed analysis of a high performance computing scenario and of a grid computing scenario.

6.2.2. Entropy Based Analysis

Although the previous approaches already improve upon state of the art and are useful on current scenarios, it is clear that at very large scale they would probably not be as effective, which led us to change perspective and to investigate how entropy can help building tractable macroscopic descriptions. Indeed, data aggregation can provide such abstractions by partitioning the systems dimensions into aggregated pieces of information. This process leads to information losses, so the partitions should be chosen with the greatest caution, but in an acceptable computational time. While the number of possible partitions grows exponentially with the size of the system, we propose in [25] an algorithm that exploits exogenous constraints regarding the system semantics to find best partitions in a linear or polynomial time. We detail two constrained sets of partitions that are respectively applied to temporal and spatial aggregation of an agent-based model of international relations. The algorithm succeeds in providing meaningful high-level abstractions for the system analysis.

Our approach is able to evaluate geographical abstractions used by the domain experts in order to provide efficient and meaningful macroscopic descriptions of the world global state [23]. We also successfully applied this technique to identify international media events by spatially and temporally aggregating RSS Flows of Newspapers [22], in particular with the case of the Syrian civil war between May 2011 and December 2012 [31], [21].

We also applied this technique to the analysis of large distributed systems and combined it with the treemap visualization technique [40], [14]. These features have been integrated in the Viva and Triva open-source tools and framework.

6.3. Trace Management and Analysis

6.3.1. Embedded Systems

The growing complexity of embedded system hardware and software makes their behavior analysis a challenging task. In this context, tracing provides relevant information about the system execution and appears to be a promising solution. However, trace management and analysis are hindered by several issues like the diversity of trace formats, the incompatibility of trace analysis methods, the problem of trace size and its storage as well as by the lack of visualization scalability. In [42], [27], [41], we present FrameSoC, a new trace management infrastructure that solves all the above issues together. It provides generic solutions for trace storage and defines interfaces and plugin mechanisms for integrating diverse analysis tools. We illustrate the benefit of FrameSoC with a case study of a visualization module that enables representation scalability of large traces by using an aggregation algorithm. Temporal aggregation techniques based on entropy are also currently integrated to the FrameSoC framework.

6.3.2. Jobs Resource Utilization

In HPC community the System Utilization metric enables to determine if the resources of the cluster are efficiently used by the batch scheduler. This metric considers that all the allocated resources (memory, disk, processors, etc) are full-time utilized. To optimize the system performance, we have to consider the effective physical consumption by jobs regarding the resource allocations. This information gives an insight into whether the cluster resources are efficiently used by the jobs. In [20], [30], we propose an analysis of production clusters based on the jobs resource utilization. The principle is to collect simultaneously traces from the job scheduler (provided by logs) and jobs resource consumption. The latter has been realized by developing a job monitoring tool, whose impact on the system has been measured as lightweight (0.35% speed-down). The key point is to statistically analyze both traces to detect and explain underutilization of the resources. This could enable to detect abnormal behavior, bottlenecks in the cluster leading to a poor scalability, and justifying optimizations such as gang scheduling or best effort scheduling. This method has been applied to two medium sized production clusters on a period of eight months.

6.4. Reconstructing the Software Environment of an Experiment

In the scientific experimentation process, an experiment result needs to be analyzed and compared with several others, potentially obtained in different conditions. Thus, the experimenter needs to be able to redo the experiment. Several tools are dedicated to the control of the experiment input parameters and the experiment replay. In parallel concurrent and distributed systems, experiment conditions are not only restricted to the input parameters, but also to the software environment in which the experiment was carried out. It is therefore essential to be able to reconstruct this type of environment. The task can quickly become complex for experimenters, particularly on research platforms dedicated to scientific experimentation, where both hardware and software are in constant rapid evolution. In [19] we discuss the concept of the reconstructability of software environments and propose a tool, Kameleon, for dealing with this problem.

6.5. Performance Evaluation

6.5.1. Computing the Throughput of Probabilistic and Replicated Streaming Applications

In [8], we investigate how to compute the throughput of probabilistic and replicated streaming applications. We are given (i) a streaming application whose dependence graph is a linear chain; (ii) a one-to-many mapping of the application onto a fully heterogeneous target platform, where a processor is assigned at most one application stage, but where a stage can be replicated onto a set of processors; and (iii) a set of random variables modeling the computation and communication times in the mapping. We show how to compute the throughput of the application, i.e., the rate at which data sets can be processed, under two execution models, the Strict model where the actions of each processor are sequentialized, and the Overlap model where a processor can compute and communicate in parallel. The problem is easy when application stages are not replicated, i.e., assigned to a single processor: in that case the throughput is dictated by the critical hardware resource. However, when stages are replicated, i.e., assigned to several processors, the problem becomes surprisingly complicated: even in the deterministic case, the optimal throughput may be lower than the smallest internal resource throughput. The first contribution of the paper is to provide a general method to compute the throughput when mapping parameters are constant or follow I.I.D. exponential laws. The second contribution is to provide bounds for the throughput when stage parameters (computation and communication times) form associated random sequences, and are N.B.U.E. (New Better than Used in Expectation) variables: the throughput is bounded from below by the exponential case and bounded from above by the deterministic case. An extensive set of simulation allows us to assess the quality of the model, and to observe the actual behavior of several distributions.

6.5.2. Optimization of Cloud Task Processing with Checkpoint-Restart Mechanism

In [17], we explain how to optimize fault-tolerance techniques based on a checkpointing/restart mechanism, in the context of cloud computing. Our contribution is three-fold. (1) We derive a fresh formula to compute the optimal number of checkpoints for cloud jobs with varied distributions of failure events. Our analysis is not only generic with no assumption on failure probability distribution, but also attractively simple to apply in practice. (2) We design an adaptive algorithm to optimize the impact of checkpointing regarding various costs like checkpointing/restart overhead. (3) We evaluate our optimized solution in a real cluster environment with hundreds of virtual machines and Berkeley Lab Checkpoint/Restart tool. Task failure events are emulated via a production trace produced on a large-scale Google data center. Experiments confirm that our solution is fairly suitable for Google systems. Our optimized formula outperforms Young's formula by 3-10 percent, reducing wallclock lengths by 50-100 seconds per job on average.

6.6. Game Theory and Applications

6.6.1. Fair Scheduling in Large Distributed Computing Sytems

Fairly sharing resources of a distributed computing system between users is a critical issue that we have investigated in two ways.

Our first proposal specifically addresses the question of designing a distributed sharing mechanism. A possible answer resorts to Lagrangian optimization and distributed gradient descent. Under certain conditions, the resource sharing problem can be formulated as a global optimization problem, which can be solved by a distributed self-stabilizing demand and response algorithm. In the last decade, this technique has been applied to design network protocols (variants of TCP, multi-path network protocols, wireless network protocols) and even distributed algorithms for smart grids. In [9], we explain how to use this technique for scheduling Bag-of-Tasks (BoT) applications on a Grid since until now, only simple mechanisms have been used to ensure a fair sharing of resources amongst these applications. Although the resulting algorithm is in essence very similar to previously proposed algorithms in the context of flow control in multi-path networks, we show using carefully designed experiments and a thorough statistical analysis that the grid context is surprisingly more difficult than the multi-path network context. Interestingly, we can show that, in practice, the convergence of the algorithm is hindered by the heterogeneity of application characteristics, which is completely overlooked in related theoretical work. Our careful investigation provides enough insights to understand the true difficulty of this approach and to propose a set of non-trivial adaptations that enable convergence in the grid context. The effectiveness of our proposal is proven through an extensive set of complex and realistic simulations.

Our second proposal is centralized but more fine grain as it does drop the steady-state hypothesis and considers sequences of campaigns. Campaign Scheduling is characterized by multiple job submissions issued from multiple users over time. The work in [18] presents a new fair scheduling algorithm called OStrich whose principle is to maintain a virtual time-sharing schedule in which the same amount of processors is assigned to each user. The completion times in the virtual schedule determine the execution order on the physical processors. Then, campaigns are interleaved in a fair way by OStrich. For independent sequential jobs, we show that OStrich guarantees the stretch of a campaign to be proportional to campaign's size and the total number of users. The theoretical performance of our solution is assessed by simulating OStrich compared to the classical FCFS algorithm, issued from synthetic workload traces generated by two different user profiles. This is done to demonstrate how OStrich benefits both types of users, in contrast to FCFS.

6.6.2. Fundamentals of Continuous Games

We have made the following contributions:

1. Continuous-time game dynamics are typically first order systems where payoffs determine the growth rate of the players' strategy shares. In [12], we investigate what happens beyond first order by viewing payoffs as higher order forces of change, specifying e.g., the acceleration of the players' evolution instead of its velocity (a viewpoint which emerges naturally when it comes to aggregating empirical data of past instances of play). To that end, we derive a wide class of higher order game dynamics, generalizing first order imitative dynamics, and, in particular, the replicator dynamics. We show that strictly dominated strategies become extinct in n-th order payoff-monotonic dynamics n orders as fast as in the corresponding first order dynamics; furthermore, in stark contrast to first order, weakly dominated strategies also become extinct for $n \geq 2$. All in all, higher order payoff-monotonic dynamics lead to the elimination of weakly dominated strategies, followed by the iterated deletion of strictly dominated strategies, thus providing a dynamic justification of the well-known epistemic rationalizability process of Dekel and Fudenberg. Finally, we also establish a higher order analogue of the folk theorem of evolutionary game theory, and we show that convergence to strict equilibria in n-th order dynamics is n orders as fast as in first order.
2. In [37] we introduce a new class of game dynamics made of a pay-off replicator-like term modulated by an entropy barrier which keeps players away from the boundary of the strategy space. We show that these *entropy-driven* dynamics are equivalent to players computing a score as their on-going exponentially discounted cumulative payoff and then using a quantal choice model on the scores to pick an action. This dual perspective on *entropy-driven* dynamics helps us to extend the folk theorem on convergence to quantal response equilibria to this case, for potential games. It also provides the main ingredients to design a discrete time effective learning algorithm that is fully distributed and only requires partial information to converge to QRE. This convergence is resilient to stochastic perturbations and observation errors and does not require any synchronization between the players.

6.6.3. Application to Wireless Networks

We have made the following contributions:

1. Starting from an entropy-driven reinforcement learning scheme for multi-agent environments, we develop in [36] a distributed algorithm for robust spectrum management in Gaussian multiple-input, multiple-output (MIMO) uplink channels. In continuous time, our approach to optimizing the transmitters' signal distribution relies on the method of matrix exponential learning, adjusted by an entropy-driven barrier term which generates a distributed, convergent algorithm in discrete time. As opposed to traditional water-filling methods, the algorithm's convergence speed can be controlled by tuning the users' learning rate; accordingly, entropy-driven learning algorithms in MIMO systems converge arbitrarily close to the optimum signal covariance profile within a few iterations (even for large numbers of users and/or antennas per user), and this convergence remains robust even in the

presence of imperfect (or delayed) measurements and asynchronous user updates.

2. Consider a wireless network of transmitter-receiver pairs where the transmitters adjust their powers to maintain a target SINR level in the presence of interference. In [46], we analyze the optimal power vector that achieves this target in large, random networks obtained by "erasing" a finite fraction of nodes from a regular lattice of transmitter-receiver pairs. We show that this problem is equivalent to the so-called Anderson model of electron motion in dirty metals which has been used extensively in the analysis of diffusion in random environments. A standard approximation to this model is the so-called coherent potential approximation (CPA) method which we apply to evaluate the first and second order intra-sample statistics of the optimal power vector in one- and two-dimensional systems. This approach is equivalent to traditional techniques from random matrix theory and free probability, but while generally accurate (and in agreement with numerical simulations), it fails to fully describe the system: in particular, results obtained in this way fail to predict when power control becomes infeasible. In this regard, we find that the infinite system is always unstable beyond a certain value of the target SINR, but any finite system only has a small probability of becoming unstable. This instability probability is proportional to the tails of the eigenvalue distribution of the system which are calculated to exponential accuracy using methodologies developed within the Anderson model and its ties with random walks in random media. Finally, using these techniques, we also calculate the tails of the system's power distribution under power control and the rate of convergence of the Foschini-Miljanic power control algorithm in the presence of random erasures.

MEXICO Project-Team

6. New Results

6.1. Diagnosis

- For non-diagnosable discrete event systems, *active* diagnosis aims at synthesizing a partial-observabion based control for the system in order to make it diagnosable. While some solutions had already been proposed for the active diagnosis problem, their complexity remained to be improved. In [40], we solved both the active diagnosability decision problem and the active diagnoser synthesis problem, proving that (1) our procedures are optimal w.r.t. to computational complexity, and (2) the memory required for the active diagnoser produced by the synthesis is minimal. Furthermore, focusing on the minimal delay before detection, we establish that the memory required for any active diagnoser achieving this delay may be highly greater than the previous one. So we refine our construction to build with the same complexity and memory requirement an active diagnoser that realizes a delay bounded by twice the minimal delay. An extension to *probabilistic* systems has been accepted to *FoSSaCS 2014*.
- In [41], we present a methodology for fault diagnosis in concurrent, partially observable systems with additional fairness constraints. In this *weak* diagnosis, one asks whether a concurrent chronicle of observed events allows to determine that a non-observable fault will inevitably occur, sooner or later, on any maximal system run compatible with the observation. The approach builds on strengths and techniques of unfoldings of safe Petri nets, striving to compute a compact prefix of the unfolding that carries sufficient information for the diagnosis algorithm. Our work extends and generalizes the unfolding-based diagnosis approaches by Benveniste et al. as well as Esparza and Kern. Both of these focused mostly on the use of sequential observations, in particular did not exploit the capacity of unfoldings to reveal inevitable occurrences of concurrent or future events studied by Balaguer et al. [19]. Our diagnosis method captures such indirect, revealed dependencies. We develop theoretical foundations and an algorithmic solution to the diagnosis problem, and present a SAT solving method for practical diagnosis with our approach. The algorithms to check diagnosability of concurrent systems are usually performed by local diagnoses of twin plant communicating with each other, directly or through a co- ordinator, and by that means pooling together the observations. Parallel analysis of diagnosability [43] takes advantage of the distribution of the system allowing to decide the diagnosability of the whole system in terms of the diagnosability of smaller systems.

6.2. Testing for Concurrent Systems

6.2.1. Model Based Testing with Labeled Event Structures

In [52], we have developped a complete testing framework for concurrent systems, which included the notions of test suites and test cases. We studied what kind of systems are testable in such a framework, and we have proposed sufficient conditions for obtaining a complete test suite as well as an algorithm to construct a test suite with such properties. However complete test suites are usually infinite in practice. In [44] (and a submitted journal version), we have proposed several testing criteria based on dedicated notions of complete prefixes that selects a manageable test suite together with a coverable criterion that allows to compare them.

6.3. Petri Nets

6.3.1. A Modular Approach for Reusing Formalisms in Verification Tools of Concurrent Systems

Over the past two decades, numerous verification tools have been successfully used for verifying complex concurrent systems, modelled using various formalisms. However, it is still hard to coordinate these tools since they rely on such a large number of formalisms. Having a proper syntactical mechanism to interrelate them through variability would increase the capability of effective integrated formal methods. In [28], we propose a modular approach for defining new formalisms by reusing existing ones and adding new features and/or constraints. Our approach relies on standard XML technologies; their use provides the capability of rapidly and automatically obtaining tools for representing and validating models. It thus enables fast iterations in developing and testing complex formalisms. As a case study, we applied our modular definition approach on families of Petri nets and timed automata.

6.3.2. Computation of summaries using net unfoldings

In [38], we study the following summarization problem: given a parallel composition A = A1 || ... || An of labelled transition systems communicating with the environment through a distinguished component Ai, efficiently compute a summary Si such that E || A and E || Si are trace-equivalent for every environment E. While Si can be computed using elementary automata theory, the resulting algorithm suffers from the state-explosion problem. We present a new, simple but subtle algorithm based on net unfoldings, a partial-order semantics, give experimental results. Our algorithm can also handle divergences and compute weighted summaries with minor modifications.

6.3.3. Complexity Analysis of Continuous Petri Nets

At the end of the eighties, continuous Petri nets were introduced for: (1) alleviating the combinatory explosion triggered by discrete Petri nets and, (2) modelling the behaviour of physical systems whose state is composed of continuous variables. Since then several works have established that the computational complexity of deciding some standard behavioural properties of Petri nets is reduced in this framework. In [39], we first establish the decidability of additional properties like boundedness and reachability set inclusion. We also design new decision procedures for the reachability and lim-reachability problems with a better computational complexity. Finally we provide lower bounds characterising the exact complexity class of the boundedness, the reachability, the deadlock freeness and the liveness problems.

6.3.4. Contextual Merged Processes

In [45], we integrate two compact data structures for representing state spaces of Petri nets: merged processes and contextual prefixes. The resulting data structure, called contextual merged processes (CMP), combines the advantages of the original ones and copes with several important sources of state space explosion: concurrency, sequences of choices, and concurrent read accesses to shared resources. In particular, we demonstrate on a number of benchmarks that CMPs are more compact than either of the original data structures. Moreover, we sketch a polynomial (in the CMP size) encoding into SAT of the model-checking problem for reachability properties.

6.3.5. A Canonical Contraction for Safe Petri Nets

Under maximal semantics, the occurrence of an event a in a concurrent run of an occurrence net may imply the occurrence of other events, not causally related to a, in the same run. In recent works, we have formalized this phenomenon as the *reveals* relation, and used it to obtain a contraction of sets of events called *facets* in the context of occurrence nets. In [36], we extend this idea to propose a canonical contraction of general safe Petri nets into pieces of partial-order behaviour which can be seen as “macro-transitions” since all their events must occur together in maximal semantics. On occurrence nets, our construction coincides with the facets abstraction. Our contraction preserves the maximal semantics in the sense that the maximal processes of the contracted net are in bijection with those of the original net.

6.4. Composition

6.4.1. Specification of Asynchronous Component Systems with Modal I/O-Petri Nets

In collaboration with Professor Rolf Hennicker from LMU and M.H. Møller, a PhD student from Aalborg University, we have studied the asynchronous composition of systems where the internal channels remain observable. In [42], we have modelled such systems by Petri nets enlarged with communication channels, we have defined several channel properties and shown these properties are compositional, and proved their decidability. In TGC 2013 (not yet in HAL), we have extended the previous models with modalities *must* and *may* "à la Larsen" and generalized most of the results in this framework.

6.4.2. Bounding models families for performance evaluation in composite Web services

One challenge of composite Web service architectures is the guarantee of the Quality of Service (QoS). Performance evaluation of these architectures is essential but complex due to synchronizations inside the orchestration of services. In (ADD WHEN IN HAL), we propose methods to automatically derive from the original model a family of bounding models for the composite Web response time. These models allow to find the appropriate trade-off between accuracy of the bounds and the computational complexity. The numerical results show the interest of our approach w.r.t. complexity and accuracy of the response time bounds.

6.5. Stochastic Systems

6.5.1. Simulation-based Verification of HASL (Hybrid Automata Stochastic Logic) Formulas for Stochastic Symmetric Nets

The Hybrid Automata Stochastic Logic (HASL) has been recently defined as a flexible way to express classical performance measures as well as more complex, path-based ones (generically called "HASL formulas"). The considered paths are executions of Generalized Stochastic Petri Nets (GSPN), which are an extension of the basic Petri net formalism to define discrete event stochastic processes. The computation of the HASL formulas for a GSPN model is demanded to the COSMOS tool, that applies simulation techniques to the formula computation. Stochastic Symmetric Nets (SSN) are an high level Petri net formalism, of the colored type, in which tokens can have an identity, and it is well known that colored Petri nets allow one to describe systems in a more compact and parametric form than basic (uncolored) Petri nets. In [27], we propose to extend HASL and COSMOS to support colors, so that performance formulas for SSN can be easily defined and evaluated. This requires a new definition of the logic, to ensure that colors are taken into account in a correct and useful manner, and a significant extension of the COSMOS tool.

6.5.2. Steady-state control problem for Markov decision processes

We address in (ADD CITATION WHEN IN HAL) a control problem for probabilistic models in the setting of Markov decision processes (MDP). We are interested in the steady-state control problem which asks, given an ergodic MDP M and a distribution δ, whether there exists a (history-dependent randomized) policy π ensuring that the steady-state distribution of M under $\imath$ is exactly δ. We first show that stationary randomized policies suffice to achieve a given steady-state distribution. Then we infer that the steady-state control problem is decidable for MDP, and can be represented as a linear program which is solvable in PTIME. This decidability result extends to labeled MDP (LMDP) where the objective is a steady-state distribution on labels carried by the states, and we provide a PSPACE algorithm. We also show that a related steady-state language inclusion problem is decidable in EXPTIME for LMDP. Finally, we prove that if we consider MDP under partial observation (POMDP), the steady-state control problem becomes undecidable.

6.6. Timed Systems

6.6.1. Back in Time Petri Nets

The time progress assumption is at the core of the semantics of real-time formalisms. It is also the major obstacle to the development of partial-order techniques for real-time distributed systems since the events are ordered both by causality and by their occurrence in time. Anyway, extended free choice safe time Petri nets (TPNs) were already identified as a class where partial order semantics behaves well. In [37], we show that, for this class, the time progress assumption can even be dropped (time may go back in case of concurrency), which establishes a nice relation between partial-order semantics and time progress assumption.

6.6.2. Expressiveness of Timed Models

In coopération with Nantes and UPMC, an in-depth study of the expressiveness of time Petri nets was completed [20]. With roughly the same partners, we have extended th ITA (Interrupt Timed Automata) by parametrizing both guards and clock rates while preserving the decidability results (RP 2013, not yet in HAL).

6.7. Weighted Systems

6.7.1. Specification and Verification of Quantitative Properties via Expressions, Logics, and Automata

Alongside boolean properties, automatic verification of *quantitative* properties such as lifespan of an equipment, energy consumption of an application or reliability of a program is gaining importance rapidly. In the thesis [14] and the articles [32], [14] , several weight-enabled formalisms for specification of such properties were examined, including denotational ones such as regular expressions, first-order logic with transitive closure, or temporal logics, as well as more operational ones such as navigating automata, possibly extended with pebbles. A unified framework of graph structures allows to compare these formalisms with respect to expressiveness, using efficient translations from denotational to operational formalisms. Several decidability and complexity results for the algorithmic questions that arise were obtained, depending on the underlying semiring from which weights are chosen, and on the structures (words, trees, ...) considered.

6.8. Dynamic Communicating Systems

6.8.1. Specification and Verification of Dynamic Message-Passing Systems

In [31], we study dynamic communicating automata (DCA), an extension of classical communicating finite-state machines that allows for dynamic creation of processes. The behavior of a DCA can be described as a set of message sequence charts (MSCs). While DCA serve as a model of an implementation, we propose branching high-level MSCs (bHMSCs) on the specification side. Our focus is on the implementability problem: given a bHMSC, can one construct an equivalent DCA? As this problem is undecidable, we introduce the notion of executability, a decidable necessary criterion for implementability. We show that executability of bHMSCs is EXPTIME-complete. We then identify a class of bHMSCs for which executability effectively implies implementability.

6.9. Concurrent Recursive Programs

6.9.1. The Complexity of Model Checking Concurrent Recursive Programs

In [34], we consider the linear-time model checking problem for boolean concurrent programs with recursive procedure calls. While sequential recursive programs are usually modeled as pushdown automata, concurrent recursive programs involve several processes and can be naturally abstracted as pushdown automata with multiple stacks. Their behavior can be understood as words with multiple nesting relations, each relation connecting a procedure call with its corresponding return. To reason about multiply nested words, we consider the class of all temporal logics as defined in the book by Gabbay, Hodkinson, and Reynolds (1994). The unifying feature of these temporal logics is that their modalities are defined in monadic second-order (MSO) logic. In particular, this captures numerous temporal logics over concurrent and/or recursive programs that have been defined so far. Since the general model checking problem is undecidable, we restrict attention to phase bounded executions as proposed by La Torre, Madhusudan, and Parlato (LICS 2007). While the MSO model checking problem in this case is non-elementary, our main result states that the model checking (and satisfiability) problem for all MSO-definable temporal logics is decidable in elementary time. More precisely, it is solvable in $(n + 2)$-EXPTIME where n is the maximal level of the MSO modalities in the monadic quantifier alternation hierarchy. We complement this result and provide, for each level n, a temporal logic whose model checking problem is n-EXPSPACE-hard.

6.9.2. Model Checking Concurrent Recursive and Communicating Programs via Split-Width

The work described in the following was done by Aiswarya Cyriac in collaboration with Paul Gastin and K. Narayan Kumar, and it is part of Aiswarya Cyriac's PhD thesis, which has recently been defended. It is a generalisation of our CONCUR'12 paper where split-width is introduced to address the decidability of MSO specifications for multi-pushdown systems.

We consider generic systems which incorporate shared-variable communication and communication via channels. We are considering physically distributed machines which communicate via (possibly several) reliable first-in-first-out queues. Each of these machines are capable of running potentially recursive multi-threaded programs. These programs within a machine use shared variable for communication. Such a machine consisting of a set of threads communicating by shared memory can be formally modelled as a multi-pushdown system. Thus we have a network of multi-pushdown systems communicating via FIFO queues. Moreover, these programs may use stacks and queues as data-structures to aid their local computation. We call such a system a system of concurrent processes with data-structures (CPDS).

We introduce and study a new technique called split-width for the under-approximate verification of CPDS. This parameter is based on simple shuffle and merge operations and gives us a divide-conquer-way to prove the bound of languages. When parametrised by a bound on split-width, we obtain decidability for various verification problems. We provide a uniform decision procedure for various verification problems with optimal complexities.

We expose the power of split-width in several ways. We show that our simple algebra is powerful enough to capture any class of CPDS which admits decidability for MSO model checking, and yardstick graph metrics such as tree-width and clique-width. We also show that various restrictions well-studied in the literature for obtaining decidability of reachability for the particular cases of multi-pushdown systems and message passing systems admit a bound on split-width. In fact, we propose generic controllers which subsume many of these cases.

Distributed controller design amounts to designing a controller (which is another CPDS) which, when run sychronously with a system ensures bounded split-width. These controllers are distributed in nature and are independent of the system it is controlling. Thus such a controller respects the privacy of the system (by not reading their states, for instance). Moreover, thanks to split-width such a controlled system offers efficient (in most cases optimal) decision procedures for the verification of the controlled system. We propose a generic approach to define controllable classes of CPDS in terms of quotient graphs, which admit a "suitable" acyclicity restriction. We also give a generic controller for several of the classes definable in this framework. The controllers we propose are sound and complete for the respective class, meaning that they allow all and only the behaviours of this class. Moreover, our technique for proving the bound on split-width of the controlled systems is also generic and systematic, hence may easily extend to generalisations and other classes as well.

The decidability results for the controllable classes proposed in the thesis are new while they capture, as special cases, several restrictions studied in the literature like bounded phase, bounded scope, poly-forest topology etc.

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 and S. Lahbabi 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 collaboration with M. Lewin, E. Cancès and S. Lahbabi have introduced 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.

D. Gontier has obtained a complete, explicit, characterization of the set of spin-polarized densities for finite molecular systems. This problem was left open in the pionnering work of von Barth and Hedin setting up the Kohn-Sham density functional theory for magnetic compounds.

On the numerical side, E. Cancès, L. He (ENPC), Y. Maday (University Paris 6) and R. Chakir (IFSTTAR) have designed and analyzed a two-grid methods for nonlinear elliptic eigenvalue problems, which can be applied, in particular, to the Kohn-Sham model. Some numerical tests demonstrating the interest of the approach have been performed with the Abinit software.

Implicit solvation models aims at computing the properties of a molecule in solution (most chemical reactions take place in the liquid phase) by replacing all the solvent molecules but the few ones strongly interacting with the solute, by an effective continuous media accounting for long-range electrostatics. E. Cancès, Y. Maday (Paris 6), and B. Stamm (Paris 6) have recently introduced a very efficient domain decomposition method for the simulation of large molecules in the framework of the so-called COSMO implicit solvation models. A collaboration with F. Lipparini (Paris 6), B. Mennucci (Department of Chemistry, University of Pisa) and J.-P. Picquemal (Paris 6) is in progress to implement this algorithm in widely used computational softwares (Gaussian and Tinker), and to extend this method to other implicit solvation models.

Claude Le Bris, in collaboration with Pierre Rouchon (Ecole des Mines de Paris), has pursued the study of a new efficient numerical approach, based on a model reduction technique, to simulate high dimensional Lindblad type equations at play in the modelling of open quantum systems. The specific case under consideration is that of oscillation revivals of a set of atoms interacting resonantly with a slightly damped coherent quantized field of photons. The approach may be employed for other similar equations. Current work is directed towards other numerical challenges for this type of problems.

5.2. Computational Statistical Physics

Participants: Claude Le Bris, Frédéric Legoll, Tony Lelièvre, Francis Nier, Mathias Rousset, Gabriel Stoltz.

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 [39], T. Lelièvre and G. Stoltz, in collaboration with physicists from CEA Saclay (especially, M. Athenes) studied a new adaptive technique of ABF type to compute on-the-fly the free energy of a system, without evaluating the second derivatives of the reaction coordinate. The method uses a Bayesian reinterpretation of an extended system where the reaction coordinate is considered as an additional variable.

In [44], G. Fort (Telecom Paris), B. Jourdain (CERMICS), E. Kuhn (INRA), T. Lelièvre and G. Stoltz have studied the efficiency of the Wang-Landau algorithm, building on a previous study where they proved the convergence of this method. The aim was to obtain precise estimates of the exit times out of metastable states. This was done in two ways: a theoretical study in the simplest possible metastable situation, a system with three states; and a numerical study in a more realistic situation (a two-dimensional double well potential).

5.2.2. 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.

Recently, the quasi stationary distribution has been identified by the team as a good mathematical tool to analyze metastable trajectories, and to make a link between a continuous state space dynamics (Langevin dynamics) and a discrete state space dynamics (kinetic Monte Carlo models), see for examplelelievre-13. This perspective can also be used to analyze accelerated dynamics techniques which have been proposed by A. Voter in the late nineties, to simulate very efficiently the state-to-state dynamics associated with metastable trajectories. For example, in [33], T. Lelièvre with D. Aristoff (University of Minnesota) propose a mathematical analysis of the Temperature Accelerated Dynamics. In [49], T. Lelièvre and F. Nier have studied the quasi-stationary distribution 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 accurately compute the spatial exit law along the boundary and non perturbative accurate formulas when the potential is changed inside the domain. This gives some insight into the foundations of the hyperdynamics method.

Finally, 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 [19] the distribution of the lengths of these trajectories, using large deviation techniques.

5.2.3. Nonequilibrium systems

Let us also mention that the article [22] on a derivation of a Langevin-type dynamics for a heavy particle in a non-zero background flow, co-authored by M. Dobson, F. Legoll, T. Lelièvre, and G. Stoltz, has been published.

5.2.4. Sampling techniques

In [29], T. Lelièvre studies with F. Nier and G. Pavliotis (Imperial College, London) the interest of using non-reversible dynamics (overdamped Langevin dynamics with a non-gradient drift term) to efficiently sample a given Boltzmann-Gibbs distribution.

5.2.5. Numerical analysis of simulation methods

Together with B. Leimkuhler and Ch. Matthews (Edinburgh University), G. Stoltz studied in [48] the discretization errors in the computation of average properties with Langevin dynamics integrated with splitting strategies. The average properties are either static (average of a given observable) or dynamic (transport coefficients). The main tool used in this analysis is the expansion of the transition operator in powers of the time step, with exact integral remainders; as well as fine estimates on the resolvent of the Langevin operators, especially in the so-called overdamped limit where the friction goes to infinity. Transport coefficients are studied either through errors in Green-Kubo formulae or errors in the linear response of nonequilibrium systems.

5.2.6. Coarse-graining of molecular systems

G. Stoltz, in collaboration with J.-B. Maillet and G. Faure, developed in [43] a potential energy function depending on the local density of the molecular fluid. The local density is evaluated with a three dimensional Voronoi tesselation, which proves more rigorous than the standard local averages typically found in the literature. The new potential allows to describe the compressibility of mesoparticles representing several molecules in a coarse-grained description of the atomic system. The quality of the potential has been assessed by reproducing equations of state and Hugoniot curves of model energetic materials.

5.2.7. 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. The work from F. Legoll and X. Blanc (Université Pierre et Marie Curie) is now published [12].

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. The work from F. Legoll in collaboration with E. Tadmor, W. K. Kim, L. Dupuy and R. Miller on the analysis of such an approach is now also published [32].

5.2.8. 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. The work from M. Dobson, C. Le Bris, and F. Legoll developing integrators for Hamiltonian systems with high frequencies (derived using homogenization techniques applied to the Hamilton-Jacobi PDE associated to the Hamiltonian ODE) is now published [22].

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. The work from C. Le Bris and F. Legoll, in collaboration with X. Dai and Y. Maday, studying several variants of the original plain parareal in time algorithm, is now also published [21].

5.2.9. 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. In this context, S. Lahbabi and F. Legoll have studied in [8] the case when the fine-scale, reference dynamics is a kinetic Monte Carlo model with small and fast time scales, and proved a path-wise convergence to a coarse kinetic Monte Carlo model only retaining slow degrees of freedom.

Another question is how to use a coarse-grained description (involving only the slow degrees of freedom) as a predictor for the dynamics of the actual reference system, involving all degrees of freedom. Together with G. Samaey (KU Leuven), F. Legoll and T. Lelièvre have addressed this question in the parareal framework, and shown in [28] that the precise coupling between both models should be done carefully in order for the algorithm to be efficient. In that case, the algorithm converges to the reference full dynamics.

5.3. Complex fluids

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

In his PhD under the supervision of Claude Le Bris and Tony Lelièvre, David Benoit 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 [10], 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, see [34].

Let us also mention that the paper [28] on a parareal algorithm to efficiently simulate micro-macro models which has been published this year.

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. Possibly discontinuous solutions were numerically simulated with a new finite-volume scheme of relaxation type that satisfies a discrete counterpart of the natural dissipation [13].

This work has been pursued for other fluid models and other flow regimes, with a view to better understanding the reduction mechanism leading from a physically detailed model to a useful one for numerical simulations at large (geophysical) scales [35].

On the other hand, note that it is often possible to consider only models for *incompressible* fluids (at low Mach numbers). Now, it is both important and delicate to understand how to numerically discretize the incompressibility constraint, a long-standing issue in numerical fluid mechanics. In collaboration with M. Picasso (EPFL), S. Boyaval has thus investigated the possibility to numerically quantify *a posteriori* the quality of a well-known, "simple" numerical method discretizing the incompressibility constraint, in a simple case [36]. This is part of another effort toward useful numerical simulations of complex flows, inline with current questions focused on discretization methods..

5.4. Application of greedy algorithms

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

Model reduction techniques are very important tools for applications. They consist in deriving from a high-dimensional problem, a low-dimensional model, which gives very quickly reliable results. We are in particular interested in two techniques: Proper Generalized Decomposition (greedy algorithms) and Reduced Basis techniques.

Concerning the Proper Generalized Decomposition, current research concerns the approximation of high-dimensional spectral problems, see [38]. Prototypical applications include electronic structure calculations or the computation of buckling modes in mechanics. We also explored in the PhD of J. Infante Acevedo the application of these techniques to option pricing problems, see [45].

Finally, in [40], Fabien Casenave (CERMICS), Alexandre Ern (CERMICS), Guillaume Sylvand (EADS IW) and Tony Lelièvre propose a new non intrusive implementation of the reduced basis technique using the Empirical Interpolation Method. The interest if the method is illustrated on aeroacoustic problems.

5.5. Homogenization and related topics

Participants: Virginie Ehrlacher, Claude Le Bris, Frédéric Legoll, François Madiot, William Minvielle.

The homogenization of (deterministic) non periodic systems is a well known topic. Although well explored theoretically by many authors, it has been less investigated from the standpoint of numerical approaches (except in the random setting). In collaboration with X. Blanc and P.-L. Lions, C. Le Bris has introduced a possible theory, giving rise to a numerical approach, for the simulation of multiscale nonperiodic systems. The theoretical considerations are based on earlier works by the same authors (derivation of an algebra of functions appropriate to formalize a theory of homogenization). The numerical endeavour is completely new. 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.

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.

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 previously shown that the variance of this matrix can be reduced using the technique of antithetic variables. In [47], F. Legoll and W. Minvielle have extended this technique to nonlinear, convex homogenization problems.

In addition, F. Legoll and W. Minvielle have investigated 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.

Yet another approach to reduce the variance is the so-called Multi Level Monte Carlo (MLMC) approach, which is based on using a surrogate model for the quantity of interest. The MLMC approach consists in using many realizations of the surrogate model (which is cheap to evaluate) and few realizations of the reference model (which is more expensive to evaluate). In collaboration with Y. Efendiev and C. Kronsbein, F. Legoll has explored in [41] how this approach can be used in random homogenization.

We have discussed above approaches to efficiently compute the homogenized coefficient, assuming we have a complete knowledge of the microstructure of the material. We have recently started to consider a related inverse problem, and more precisely a parameter fitting problem. Knowing the homogenized quantities, is it possible to recover some features of the microstructure properties? Obviously, since homogenization is an averaging procedure, not everything can be recovered from macroscopic quantities. A realistic situation is the case when we assume a functional form of the distribution of the microscopic properties, but with some unknown parameters that we would like to determine. In collaboration with A. Obliger and M. Simon, F. Legoll and W. Minvielle have started to address that problem, determining the unknown parameters of the microscopic distribution on the basis of macroscopic (e.g. homogenized) quantities. The preliminary results that have been obtained are very encouraging.

From a numerical perspective, the Multiscale Finite Element Method (MsFEM) is a classical strategy to address the situation when the homogenized problem is not known (e.g. in difficult nonlinear cases), or when the scale of the heterogeneities, although small, is not considered to be zero (and hence the homogenized problem cannot be considered as an accurate enough approximation).

The MsFEM has been introduced more than 10 years ago. However, even in simple deterministic cases, there is actually still room for improvement in many different directions. In collaboration with A. Lozinski (University of Besançon), F. Legoll and C. Le Bris have introduced and studied a variant of MsFEM that considers Crouzeix-Raviart type elements on each mesh element. The continuity across edges (or facets) of the (multiscale) finite element basis set functions is enforced only weakly, using fluxes rather than point values. The approach has been analyzed (combining classical arguments from homogenization theory and finite element theory) and tested on simple, but highly convincing cases [27]. In particular, an elliptic problem set on a domain with a huge number of perforations has been considered in [37]. The variant developed outperforms all existing variants of MsFEM.

A follow up on this work, in collaboration with U. Hetmaniuk (University of Washington in Seattle) and A. Lozinski (University of Besançon), consists in the study of multiscale advection-diffusion problems. Such problems are possibly advection dominated and a stabilization procedure is therefore required. How stabilization interferes with the multiscale character of the equation is an unsolved mathematical question worth considering for numerical purposes. This is the aim of the PhD thesis of François Madiot, which began in October 2013.

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? We have addressed some of these questions in [25], where we have in particular shown that this best constant matrix converges to the homogenized matrix, in the limit of infinitely rapidly oscillatory coefficients. Our approach can therefore be considered as an alternative way to compute the homogenized matrix. This is particularly interesting in random cases, where the standard approach is very expensive. Current work is directed towards extending the approach, in order to compute other quantities of interest than the homogenized coefficient.

To conclude this section, we mention the project undertaken by V. Ehrlacher during her six months postdoctoral position in the Cluster of Excellence Engineering of Advanced Materials (Erlangen University). This project, in collaboration with C. Le Bris, F. Legoll, G. Leugering and M. Stingl, aims at optimizing the shape of some materials (modelled as structurally graded linear elastic materials) in order to achieve the best mechanical response at the minimal cost. As often the case in shape optimization, the solution tends to be highly oscillatory, thus the need of homogenization techniques. We thus consider an initial microstructured material composed of steel and void and whose microstructure pattern is periodic (think e.g. of a periodic honeycomb structure). We next consider materials which are obtained from this initial material through a macroscopic deformation, and look for the optimal deformation achieving the best mechanical response. Encouraging first results have been obtained.

5.6. Coupling methods and 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 [30], the pobabilistic derivation of the chemotaxis equation from the individual motion of bacteriae have been carried out in a rigorous way. In [31], a numerical method simulating the individual evolution of bacteriae with "asymptotic" variance reduction have been proposed.

Motivated by the asymptotic variance reduction of DSMC methods (particle Monte-Carlo methods simulating low density fluids modeled by kinetic equations), the work in [50], M. Rousset considers space homogenous Boltzmann kinetic equations in dimension d with Maxwell collisions (and without Grad's cut-off). An explicit Markov coupling of the associated conservative (Nanbu) stochastic N-particle system is constructed, using plain parallel coupling of isotropic random walks on the sphere of two-body collisional directions. The resulting coupling is almost surely decreasing, and the L_2-coupling creation is computed explicitly. Some quasi-contractive and uniform in N coupling / coupling creation inequalities are then proved, relying on 2 + α-moments ($\alpha > 0$) of velocity distributions; upon N-uniform propagation of moments of the particle system, it yields a N-scalable α-power law trend to equilibrium. The latter are based on an original sharp inequality, which bounds from above the coupling distance of two centered and normalized random variables $(U, V) \in \mathbb{R}^d$, with the average square parallelogram area spanned by $(U - U_*, V - V_*)$, (U_*, V_*) denoting an independent copy. Two counter-examples proving the necessity of the dependance on > 2-moments and the impossibility of strict contractivity are provided. The paper, (mostly) self-contained, does not require any propagation of chaos property and uses only elementary tools.

MIMETIC Project-Team

6. New Results

6.1. Biomechanics and Motion Analysis

6.1.1. Modeling gesture in sports: tennis serve

Participants: Nicolas Bideau, Guillaume Nicolas, Benoit Bideau, Richard Kulpa.

In the midst of the INSEP project and the PhD of Caroline Martin, the tennis serve has been studied with biomechanical analyses. To this end, we have done kinematic and dynamic analyses based on motion capture, force plate and electromyographic systems. They provided information on how the gesture is performed and how it is related to injuries. Moreover, these analyses have been done on several level of players including top-level ones. A comparison of the kinematic and dynamic data can then be done. Our objective is to use these data in virtual reality to study the interaction between a tennis server and a receiver. We are creating a tool that displays a virtual server in front of a real receiver. The control of the virtual server is then done based on these biomechanical data. The objective is to analyze the reaction of the receiver depending on the movement of the server and its level of expertise.

6.1.2. Motion modeling in clinical applications

Participant: Armel Crétual.

We have developed a new index of gait quantification based on muscular activity called KeR-EGI. After having proved that this index is consistent and complementary with kinematics-based indices, we have shown that it is reproducible in patients even when their impairement level is high. This index is now used in clinical routine in adults. It will be also used in pediatrics in the next few months.

In orthopedics, we have proposed a novel method to quantify shoulder's global mobility called SCSV. It is based on the reachable volume in the whole configuration space of the shoulder, i.e. a 3-dimensional angular space. Clinical evaluations of shoulder's range of motion are quite always based on the analysis of only one axis, and the most usual refers to maximal external rotation from rest posture (ER1). Considering several mono-axial amplitudes, we have shown that ER1 is actually the worst choice to estimate global mobility. Instead of the ER1 procedure, we proposed to use the sum of 3 mono-axial amplitudes: external/internal amplitude at 90° lateral elevation, abduction and flexion/extension.

As shoulder is actually a complex of three articulations (gleno-humeral, scapulo-thoracic and sterno-calvicular), we have evaluated the contribution of each of them on global mobility. This has been done through a cadaveric study where we measured SCSV in any possible blocking conditions of these three articulations (from 0 to 3).

6.2. VR and Ergonomics

Participants: Charles Pontonnier [contact], Georges Dumont, Franck Multon, Pierre Plantard.

The use of virtual reality tools for ergonomics applications is a very important challenge in order to genezalize the use of such devices for the design of workstations.

First, the development of motion analysis tools is mandatory in order to provide additional information to the ergonomists and help them to analyse the work environment. Particularly, an analysis of the muscle forces involved in the motion generation is a very important information with regard to the ergonomics of a task. Several methods can lead to an estimation of these muscle forces. In a study we developed, we tried to assess the level of confidence for results obtained with an inverse dynamics method from real captured work tasks. The chosen tasks were meat cutting tasks, well known to be highly correlated to musculoskeletal troubles appearance in the slaughter industry.

The experimental protocol consisted in recording three main data during meat cutting tasks, and analysing their variation when some of the workstation design parameters were changing.

Then the motion was replayed in the AnyBody modeling system (AnyBody, Aalborg, Denmark) in order to obtain muscle forces generated during the motion. A trend comparison has been done, comparing recorded and computed muscle activations. Results showed that most of the computed activations were qualitatively close from the recorded ones (similar shapes and peaks), but quantitative comparison leaded to major differences between recorded and computed activations (the trend followed by the recorded activations in regard of a workstation design parameter, such as the table height, is not obtained with the computed activations) [15]. We currently explore those results to see if the fact that co-contraction of single joints muscles is badly estimated by classical inverse dynamics method can be a reason of this issue. We also work on the co-contration simulation in order to improve the results.

This work has been done in collaboration with the Center for Sensory-motor Interaction (SMI, Aalborg University, Aalborg, Denmark), particularly Mark de Zee (Associate Professor) and Pascal Madeleine (Professor).

Furthermore, the fidelity of the VR simulator has to be evaluated. For example, a simulator for assembly task has been evaluated in comparing different types of interaction : real, virtual and virtual + force feedback. Objective and subjective metrics of discomfort led to highlight the influence of the environment on motor control and sensory feedback, changing more or less deeply the way the task is performed. The results particularly showed a distorsion between the user's subjective rating of discomfort and the objective value associated to the postures they reached during the task execution. Nevertheless, scores obtained in real and virtual environments for objective and subjective indicators of discomfort were highly correlated [17], [16]. It indicates that despite the differences, the gap between real and virtual environments can be fullfiled. This work has been done within the frame of the european project FP7 VISIONAIR.

At last we proposed in collaboration with Thierry Duval (Hybrid team, Rennes) a new architecture for information sharing and bridging in collaborative virtual environements in application to ergonomics studies. This work has been awarded with a best paper award at The 4th IEEE conference on Cognitive Infocommunications (CogInfoCom 2013) [28].

6.3. Motion Sensing and analysis

Participant: Franck Multon [contact].

Sensing human activity is a very active field of research, with a wide range of applications ranging from entertainment and serious games to personal ambient living assistance, including rehabilitation. MimeTIC aims at proposing original methods to process raw motion capture data in order to compute relevant information according to the application.

In rehabilitation, we have collaborated with University of Montreal, Saint-Justine Hospital which main activity is rehabilitation of children with pathologies of the pyramidal control system. In this domain, defining metrics and relevant measurement to diagnose pathologies and to monitor patients during treatment is a key point. In gait, most of the previous works focus on gait spatio-temporal parameters (such as step length, frequency, stride duration, global speed) which could be measured with two main families of systems: 1) one-point measurement with a force plate, one accelerometer or dedicated devices (such as a Gait Ride), or 2) multi-point measurement systems with motion sensors or markers placed over the patient's skin. The former provides the clinician with compact but incomplete knowledge whereas the latter provides him with numerous data which are sometimes difficult to analyze and to get (specific technical skills are required). The first step to any type of analysis is to detect the main gait events, such as foot strikes and toe offs. In treadmill walking, widely used in rehabilitation as it enables the clinician to analyze numerous gait cycles in a limited place with a controlled speed, automatically detecting such gait events requires complex devices with specific technical skills (such as calibration and post-processing with motion capture systems).

Recent papers have demonstrated that low-cost and easy-to-use depth cameras (such as a Kinect from Microsoft) look promising for serious applications requiring motion capture. However there exist some confusion between the feet and the ground at foot strike and foot off leading to bad estimation of the gait cycle events. We have proposed an alternative approach that consists in using the strong correlation between knee and foot trajectories to deduce foot strikes thanks to knee movements. The extremes of the distance between the two knees along the longitudinal axis provides us with very accurate gait events detection compared to previous works.

A second contribution consisted in defining a global gait asymmetry index according to depth images provided by a Kinect. In previous works this index relied on computing ratio between joint angles. With a Kinect, joint angles may be very noisy that could affect the asymmetry index. We have introduced a new index which is directly deduced from depth images without any joint angle estimation nor skeleton fitting. The method consists in building a model of the gait cycle of the patient by averaging depth images recorded along several cycles. As a consequence the noise within the instantaneous depth images is filtered leading to accurate surfaces of the patient gait (leading to a 3D+time data structure). The main vertical axis of the surface is used to define a symmetry plane. Consequently surfaces of the right part of the body can be symmetrized to be compared to the left part at compatible times in the gait cycle (such as a right foot strike is symmetrized to be compared to a left foot strike). The comparison between the two surfaces leads to a promising asymmetry index. The results (see Figure 4) demonstrate that this method is able to significantly distinguish asymmetrical gaits obtained by adding a 5cm sole under one of the feet of healthy subjects. Ongoing works consist in comparing this index to previously published ones which were based on accurate motion capture data. It will also be applied to unimpaired gaits of pathological subjects.

Figure 4. Longitudinal (DAI) and lateral (LAI) Asymmetry indexes computed thanks to depth images for normal gait and two artificial modification (adding a 5cm sole below the left or the right foot). The asymmetry index is computed all along the gait cycle and was able to statistically distinguish asymmetrical gaits.

6.4. VR and Sports

Participants: Richard Kulpa [contact], Benoit Bideau, Franck Multon.

Previous works in MimeTIC have shown the advantage of using VR to design and carry-out experiments on perception-action coupling in sports, especially for duels between two opponents. However the impact of using various technical solutions to carry-out this type of experiment in sports is not clear. Indeed immersion is performed by using interfaces to capture the motion/intention of the user and to deliver various multi-sensory feedbacks. These interfaces may affect the perception-action loop so that results obtained in VR cannot be systematically transferred to real practice.

Most of the applications in VR provide the user with visual feedbacks in which the avatar of the user can be more or less simplified (sometimes limited to a hand or the tools he his carrying). In first person view in caves the user generally does not need accurate avatars as he can perceive his real body but some authors have shown that the perception of distances is generally modified. Some authors have also demonstrated that first-person view was less efficient that third person view with avatars when performing accurate tasks such as reaching objects in constrained environments. We proposed an experiment to evaluate which type of feedback was the most appropriate one for complex precision tasks, such as basketball free-throw. In basketball free-throw the user has to throw a ball into a small basket placed at over 4.5m far from him. Thus perception of distance is actually a key point in such a task. Beginners and experts carried-out a first experiment in real in order to measure their motion and performance in real situation. Then beginners were asked to perform free throws with a real ball in hands, but in three conditions in a Cave (Immersia Room, Rennes): 1) first-person view (see Figure 5), 2) third-person view with the visual feedback of the ball's position, and 3) third-person view the virtual ball and additional rings modeling the perfect trajectory for the ball to get in the basket. Results show that significant difference exists in ball speed between first-person view condition compared to real condition whereas no difference exist in third-person view conditions. If we focus on successful throws only, ball speed in the last condition 3) was very similar to real condition whereas all the other VR conditions (1) and 2)) lead to significant differences compared to real situation. In all VR conditions the height of ball release was significantly higher in VR compared to real situation. These results show that VR conditions lead to adaptations in the way people perform such a precision task, especially for ball speed and height of ball release. However this difference is significantly higher with first person view and tends to zero in condition 3). Future works will tend to evaluate new conditions with avatars and complementary points of view (such as lateral and frontal views together as suggested by some authors). It will also be important to more clearly understand the problem of perception of distances in such an environment. This work has been performed in cooperation with University of Brassov in Romania.

Another key feedback is the external forces associated with the task. In most sports applications such forces are strongly linked to performance. However delivering these forces in virtual environments is still a challenge as it required haptic devices that could affect the way the users perform the task (with a different grip compared to real situation and limitations in dynamic response of the device). Pseudohaptics has been introduced in the early 2000. It consists in using visual feedbacks to make people perceive the forces linked to a task. However this approach has not been tested for whole-body interaction. In collaboration with Hybrid team in Inria Rennes, we studied how the visual animation of a self-avatar could be artificially modified in real-time in order to generate different haptic perceptions. In our experimental setup participants could watch their self-avatar in a virtual environment in mirror mode. They could map their gestures on the self-animated avatar in real-time using a Kinect. The experimental task consisted in a weight lifting with virtual dumbbells that participants could manipulate by means of a tangible stick. We introduce three kinds of modification of the visual animation of the self-avatar: 1) an amplification (or reduction) of the user motion (change in C/D ratio), 2) a change in the dynamic profile of the motion (temporal animation), or 3) a change in the posture of the avatar (angle of inclination). An example is depicted in Figure 6 . Thus, to simulate the lifting of a "heavy" dumbbell, the avatar animation was distorted in real-time using: an amplification of the user motion, a slower dynamics, and a larger angle of inclination of the avatar. We evaluated the potential of each technique using an ordering task with four different virtual weights. Our results show that the ordering task could be well achieved with every technique. The C/D ratio-based technique was found the most efficient. But participants globally appreciated all the different visual effects, and best results could be observed in the combination configuration. Our results pave the way to the exploitation of such novel techniques in various VR applications such as for sport training, exercise games, or industrial training scenarios in single or collaborative mode.

../../../../projets/mimetic/IMG/basket.jpg

Figure 5. First-person view condition in the basket free-throw performed in a cave (Immersia Room, France).

../../../../projets/mimetic/IMG/DavidGomez.jpg

Figure 6. Weight discrimination task: the animation of the avatar showed a lifting effort according to the weight of the virtual dumbbell and the user has to rank the conditions from the lightest to the heaviest mass.

6.5. Autonomous Virtual Humans

6.5.1. Space and Time Constrained Task Scheduling for Crowd Simulation

Participants: Carl-Johan Jorgensen, Fabrice Lamarche [contact].

Crowd distribution in cities highly depends on how people schedule their daily activities. When performing an intended activity, people decisions and behavior mainly consist in scheduling tasks that compose this activity, planning paths between locations where these tasks should be performed, navigating along the planned paths and performing the scheduled tasks.

We proposed a task scheduling model aims at selecting where, when and in which order several tasks, representing an intended activity, should be performed. The proposed model handles spatial and temporal constraints relating to the environment and to the agent itself. Personal preferences, characterizing the agent, are also taken into account. Produced task schedules are optimized on the long term and exhibit adequate choices of locations and times with respect to the agent intended activity and its environment. Once computed, these task schedules are relaxed and used to drive a microscopic crowd simulation in which observable flows of pedestrians emerge from the scheduled individual activities. Such simulations are easy to produce and do not require the use of a complex decisional model. In terms of validation, we conducted an experiment that shows that our algorithm produces task schedules which are representative of humans' ones.

This work is part of the iSpace&Time project in which virtual cities are populated with virtual pedestrians and vehicles.

6.5.2. Long term planning and opportunism

Participants: Philippe Rannou, Fabrice Lamarche [contact].

Autonomous virtual characters evolve in dynamic virtual environments in which changes may be unpredictable. One main problem when dealing with long term action planning in dynamic environment is that an agent should be able to behave properly and adapt its behavior to perceived changes while still fulfilling its goals.

We propose a system that combines long term action planning with failure anticipation and opportunism. The system is based on a modified version of an HTN planning algorithm. It generates plans enriched with information that enable a monitor to detect relevant changes of the environment. Once such changes are detected, a plan adaptation is triggered. Such adaptations include modifying the plan to react to a predicted failure and more importantly to exploit opportunities offered by the environment.

This system has been extended to better take into account the relationship between action planning and the environment. It is now combined with our space and time constrained tasks scheduling system (Cf. 6.5) to optimize the choice of locations where actions should be performed.

6.6. Interactive Virtual Cinematography

Participants: Marc Christie [contact], Christophe Lino, Cunka Sanokho.

The domain of Virtual Cinematography explores the operationalization of rules and conventions pertaining to camera placement, light placement and staging in virtual environments. Within the context of the ANR CHROME project, we have tackled the problem of portraying events in complex crowd simulations using steering behaviors. The system we proposed relies on Reynolds' model of steering behaviors to control and locally coordinate a collection of camera agents similar to a group of reporters. In our approach, cameras are either in a scouting mode, searching for relevant events to convey, or in a tracking mode following one or more unfolding events. The key benefit, in addition to the simplicity of the steering rules, holds in the capacity of the system to adapt to the evolving complexity of crowd simulations by self-organizing the cameras to track most of the events. The results have been presented as the Motion in Games conference [21].

We have also created a table-top interactive application to offer collaborative and high-level control on multi-dimensional and temporal data. This has been applied to the collaborative control of cinematographic parameters in a virtual movie, using our cinematographic engine [26].

In the ANR project Cinecitta, we have proposed means to evaluate the sens of balance in synthetic shots. Balance represents the equilibrium of visual weights in the screen, *i.e.* equilibrium of the visual interests one perceives. Balance is a key criteria in the aesthetics of a shot, and only a few approaches have seriously tackled this issue. In our approach, we rely on a dataset of well-balanced shot extracted from real movies to construct a balance feature space. The balance feature space is then used to estimate the sense of balance in new synthetic shot. We have furthermore extended the approach by automatically recomputing viewpoints to improve balance. A journal paper is under submission in Computer Graphics Forum.

6.7. Interactive Stroytelling

Participants: Marc Christie [contact], Hui-Yin Wu.

In 2013, within the Inria Associate Team FORMOSA (see 8.3.1.1), we have proposed a framework for the creation of parametrable and personalized stories in interactive storytelling. In any kind of storytelling, the success of the story relies both on the intricate plot design and control of the author as well as the emotional feedback of the user. With the assistance of computing algorithms combined with the maturing understanding of narrative structures, it is possible for interactive stories to create a more personalized, engaging, and well-controlled narrative content to users than traditional linear narrative. And with the emergence of new storytelling technologies, critical issues concern the creation of such complex narratives in virtual 3D environments, and the coherent simulation of these interactive narratives.

In the framework we proposed, the author can specify characteristics on the story structure and fragments (pieces of story) in order to generate variations of interactive stories. The characteristics we consider are genre, story complexity, and Chatman's modes of plot (eg a good hero fails). The story generation model we devised combines a branching story structure with a three-step graph traversal algorithm that filters and recombines story fragments from the characteristics, generating a high-level interactive script that satisfies all authorial constraints, and provides sufficient abstraction from the technical implementation. The script is then simulated in a real-time storytelling system, featuring autonomous characters and automatic camera control. The work has been presented as a short paper in the CASA conference [30].

We then extended this approach to handle temporal aspects of discourse in stories (i.e. how to temporally rearrange fragments of a story while maintaining consistency and logic whatever the user's choices). By rewriting our graph traversal algorithm (which filters inconsistent branches, and propagates constraints along the branches), and performing the graph traversal on each choice selected by the user, we enable the simulation of consistent temporal variations in stories. This typically allows the creation of flashbacks, flashforwards, parallel and embedded stories. Early results have been presented as a poster as Motion in Games 2013 [21].

6.8. Haptic Cinematography

Participant: Marc Christie [contact].

In 2013, we have demonstrated an approach to Haptic Cinematography in very selective events (2013 CHI conference [40], 2013 Siggraph Emerging Technology [41], 2013 UIST conference [39]). This is joint work with members of the Hybrid team (Anatole Lécuyer, Fabien Danieau) and members of the Technicolor Company (Philippe Guillotel, Nicolas Mollet, Julien Fleureau). Haptic cinematography consists in enhancing our audio-visual experience of movies by adding haptic effects related to the semantics of camera motions. Camera motions in movies, which are typically non-diegetic elements in a narrative, tend to enhance user experience both visually and emotionnaly. The questions we address here are (i) whether the coupling between camera motions and haptic motions improve this experiences and (ii) what are the rules and recommendations for coupling these motions. Results, that we published in the IEEE Multimedia journal [8], demonstrate that (i) the coupling is effective when precisely synchonized, (ii) the direction of motions between the camera and the haptic motions do not need to be correlated, and (iii) haptic metaphores can easily be perceived by the spectators. This opens great perspectives as to how haptic devices can enhance audio-visual contents in more sublte ways than straightforward mappings between diegatic elements and haptic motions.

6.9. Biomechanics for avatar animation

Participants: Julien Pettré [contact], Charles Pontonnier, Georges Dumont, Franck Multon, Ana Lucia Cruz Ruiz, Steve Tonneau.

Bio-inpired controllers and planners are compelling for avatar animation. We are currently engaging several works on the subject within the frame of the ENTRACTE project 8.1.5 .

Ana-Lucia Cruz-Ruiz has been recruited as a PhD student since november 2013 to begin to work on musculoskeletal-based methods for avatar animation. More precisely, the goal of this thesis is to define and evaluate a modular and multiscale whole-body musculoskeletal model usable to analyze and human movement and synthetize realistiv avatar animations. The specificity of the subject is hidden in the words "modular" and "multiscale". "Modular" says that the model has to be easily tunable to be modified in accordance with the investigated motor control theories (uncontrolled manifold, motor synergies,...). "Multiscale" means that the model has to exhibit multiple levels of details cohabiting at the same time, depending on the region of interest investigated. At last, the model have to be easily scalable, in order to be applied to different morphologies. Moreover, she currently explores musculoskeletal-based simplified joint behaviors to improve torque-based dynamics applications.

We also address the problem of planning human motion in constrained environment. In previous approach, planning human motion is performed based on robotics planning algorithms the objective of which is to avoid obstacles. In our approach, we suggest that creating contacts with the obstacles of the environment is actually a mean to perform a motion tasks. We thus model human motion as a sequence of contacts between humans and obstacles. A contact planner is being developed, and results being prepared for publication.

6.10. Crowds

Participants: Julien Pettré [contact], Anne-Hélène Olivier, Julien Bruneau, Jonathan Perrinet, Kevin Jordao, David Wolinski.

6.10.1. Analysis of Locomotion Trajectories during Collision Avoidance

The experimental observation of physical interactions between real walkers is for us a great source of inspiration for the design of realistic microscopic models of crowd simulation. This year, we have continued analysing locomotion trajectories of real walkers during collision avoidance tasks. Analysis focused on individual strategies and role set to solve such a reciprocal interaction. Our analysis revealed that walkers combine re-orientation and speed adaptations to avoid collisions, but more importantly, that the strategies, as well as the global amount of adaptations is dependent on the role each one has in the avoidance (e.g., passing first, giving way). Our results are reported in [13]. In addition, we inspected the role of psychologic factors on the metrics of interactions [27].

6.10.2. Evaluation of Locomotion Trajectories performed in Virtual Reality

Virtual Reality rooms are physically limited in space, and prevent users virtually walking by really walking in larger virtual spaces: a locomotion interface is employed to overcome this issue. The interface is composed of a peripheral device, such as joystick, as well as of a software component which transform users' actions on the peripheral device into a virtual locomotion. In this work, we wondered if users where performing similar trajectories in virtuo than in vivo: such question is important when aiming at using VR form motion analysis purpose. We evaluated the bias introduced by several couples of devices and software components during the execution of goal directed locomotion tasks. As reported in [7], impressive similarities on the formed trajectories even when the device control motions are radically different in comparison with walking motions.

6.10.3. Virtual Populations for large-scale digital environments and Cultural Heritage Applications

We are developing techniques dedicated to the animation of large virtual populations at very low computational cost based on the crowd patches techniques. Crowd patches can be described as 3D animated textures that small groups animations. They are composed in space to form large population. This year, we coupled the crowd patches approach with mutable shape models: such association enable users cdesigning patches composition in an interactive manner, as introduced in [22]. We applied those techniques to design populations of some old Malaysian trading ports [25].

6.10.4. Macroscopic derivations of microscopic simulation models

Crowd phenomenon exhibit macroscopic structures which derive from the combination of local interactions between individuals. Together with the IMT in Toulouse in the frame of the ANR-Pedigree project (term. 2012), the microscopic models developed in our team has been derived into macroscopic models to demonstrate their ability to provoke the mergence of some typical macroscopic structures [35], [36].

MINT Project-Team

6. New Results

6.1. Human limits in small unidirectional mouse movements

Participants: Jonathan Aceituno [correspondant], Géry Casiez, Nicolas Roussel.

Computer mouse sensors keep increasing in resolution. The smallest displacement they can detect gets smaller, but little is known on our ability to control such small movements. Small target acquisition has been previously tackled, but the findings do not apply to the problem of finding the useful resolution of a user with a mouse, which corresponds to the smallest displacement (s)he can reliably produce with that device. In [16], we detail this definition and provide an associated experimental protocol to measure the useful resolution. We then report on the results of a study suggesting that high-end mice are not likely to be used to their full potential. We further comment on the different strategies used by participants to acheive best performance, and derive implications for user interfaces.

6.2. Small, Medium, or Large? Estimating the User-Perceived Scale of Stroke Gestures

In [27], we show that a large consensus exists among users in the way they articulate stroke gestures at various scales (i.e., small, medium, and large) and formulate a simple rule that estimates the user-intended scale of input gestures with 87% accuracy. Our estimator can enhance current gestural interfaces by leveraging scale as a natural parameter for gesture input, reflective of user perception (i.e., no training required). Gesture scale can simplify gesture set design, improve gesture- to-function mappings, and reduce the need for users to learn and for recognizers to discriminate unnecessary symbols.

6.3. Métamorphe : a shape changing keyboard

Métamorphe is a keyboard with mobile keys [21]. Whether keys are pressed or released, they can be at their usual height, or raised. This mechanism allows both to provide haptic feedback to ease eyes-free interaction, and to access the side of the keys. The sides of the keys can be pushed, like the top of the keys. Therefore each key can be mapped to several actions. For instance this could be useful for command selection.

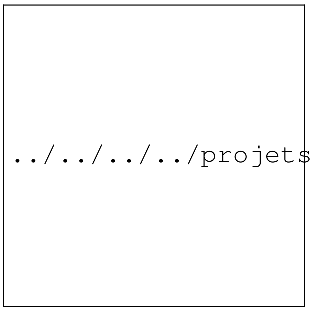
6.4. Designing Intuitive Multi-touch 3D Navigation Techniques

Participants: Géry Casiez, Damien Marchal [correspondant], Nicolas Roussel, Clement Moerman.

Multi-touch displays have become commonplace over recent years. Numerous applications take advantage of this to support interactions that build on users' knowledge and correspond to daily practices within the real world. 3D applications are also becoming more common on these platforms, but the multi-touch techniques for 3D operations often lag behind 2D ones in terms of intuitiveness and ease of use. Intuitive navigation techniques are particularly needed to make multi-touch 3D applications more useful, and systematic approaches are direly needed to inform their design: existing techniques are still too often designed in ad-hoc ways. In [25], we propose a methodology based on cognitive principles to address this problem. The methodology combines standard user-centered design practices with optical flow analysis to determine the mappings between navigation controls and multi-touch input. It was used to design the navigation technique of a specific application for our industrial partner Idées3Com. The resulting technique proved to be more efficient and preferred by users when compared to existing ones, which provides a first validation of the approach.

../../../../projets/mint/IMG/metamorphebanner.png

Figure 1. a) Métamorphe concept: the user presses the control key, keys corresponding to hotkeys rise b) key mounted on a solenoid, with force sensors on the sides c) press on the top of the key d) press on the right of the key e) press on the left of the key.



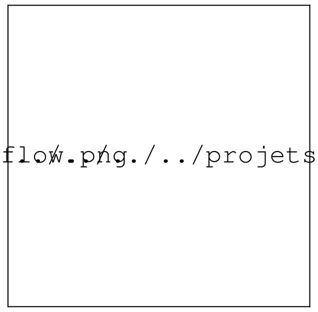


Figure 2. (Left) The optical flow for camera movements are used to design the shape of the interaction gestures. (Right) The evaluation scenario used to compare several state of the art navigation techniques.

6.5. Mockup Builder: 3D modeling on and above the surface

Mockup Builder [11] is a semi-immersive environment for conceptual design which allows virtual mockups to be created using gestures. Our goal is to provide familiar ways for people to conceive, create and manipulate three-dimensional shapes. To this end, we developed on-and-above-the-surface interaction techniques based on asymmetric bimanual interaction for creating and editing 3D models in a stereoscopic environment. Our approach combines both hand and finger tracking in the space on and above a multi-touch surface. This combination brings forth an alternative design environment where users can seamlessly switch between interacting on the surface or above it to leverage the benefit of both interaction spaces. A formal user evaluation conducted with experienced users shows very promising avenues for further work towards providing an alternative to current user interfaces for modeling.

6.6. Towards Many Gestures to One Command: A User Study for Tabletops

Participants: Yosra Rekik, Laurent Grisoni [correspondant], Nicolas Roussel.

This work has been accepted as a long paper at Interact 2013. Multi-touch gestures are often thought by application designers for a one-to-one mapping between gestures and commands, which does not take into account the high variability of user gestures for actions in the physical world; it can also be a limitation that leads to very simplistic interaction choices. Our motivation is to make a step toward many-to-one mappings between user gestures and commands, by understanding user gestures variability for multi-touch systems; for doing so, we set up a user study in which we target symbolic gestures on tabletops. From a first phase study we provide qualitative analysis of user gesture variability; we derive this analysis into a taxonomy of user gestures, that is discussed and compared to other existing taxonomies. We introduce the notion of atomic movement; such elementary atomic movements may be combined throughout time (either sequentially or in parallel), to structure user gesture. A second phase study is then performed with specific class of gesture-drawn symbols; from this phase, and according to the provided taxonomy, we evaluate user gesture variability with a fine grain quantitative analysis. Our findings indicate that users equally use one or two hands, also that more than half of gestures are achieved using parallel or sequential combination of atomic movements. We also show how user gestures distribute over different movement categories, and correlate to the number of fingers and hands engaged in interaction. Finally, we discuss implications of this work to interaction design, practical consequences on gesture recognition, and potential applications.

6.7. Sub-space gestures: elements of design for mid-air interaction with distant displays

Participants: Hanae Rateau, Laurent Grisoni [correspondant], Bruno de Araujo.

(Research report, accepted to publication in a modified version to IUI 2014). Multi-touch gestures are often thought by application designers for a one-to-one mapping between gestures and commands, which does not take into account the high variability of user gestures for actions in the physical world; it can also be a limitation that leads to very simplistic interaction choices. Our motivation is to make a step toward many-to-one mappings between user gestures and commands, by understanding user gestures variability for multi-touch systems; for doing so, we set up a user study in which we target symbolic gestures on tabletops. From a first phase study we provide qualitative analysis of user gesture variability; we derive this analysis into a taxonomy of user gestures, that is discussed and compared to other existing taxonomies. We introduce the notion of atomic movement; such elementary atomic movements may be combined throughout time (either sequentially or in parallel), to structure user gesture. A second phase study is then performed with specific class of gesture-drawn symbols; from this phase, and according to the provided taxonomy, we evaluate user gesture variability with a fine grain quantitative analysis. Our findings indicate that users equally use one or two hands, also that more than half of gestures are achieved using parallel or sequential combination of atomic movements. We also show how user gestures distribute over different movement categories, and correlate to the number of fingers and hands engaged in interaction. Finally, we discuss implications of this work to interaction design, practical consequences on gesture recognition, and potential applications.

6.8. Merging two tactile stimulation principles: Electrovibration and Squeeze film effect

Participants: Michel Amberg, Frédéric Giraud, Clément Nadal, Betty Semail [correspondant].

Electrovibration and squeeze film effect can modify the perception a user has of a flat surface, with opposite action. In fact, electrovibration increases the friction of the finger on the surface, while the squeeze film reduces it. These two stimulation principles are compatible, and in this work [23], we wanted to merge them in a tactile stimulator, in order to enhance the control of the lateral force. Our approach was to identify the effect of each tactile stimulation, and we proposed its modelling: the dynamic of the mecahnical response of the fingerpulp has to be taken into account between the programmed stimulus and the resulting lateral force. We have shown also that the two techniques may be used simultaneously accounting to a few precautions. From the first experimental trials, the conclusion here is that the squeeze film effect is able to reduce tangential forces generated by the electrostatic forces, by going on acting on the friction coefficient.

../../../../projets/mint/IMG/figure8.png ../../../../projets/mint/IMG/electrovibr

Figure 3. (Left) The experimental test bench to measure the forces produced during the stimulation. (Right) The tactile stimulator merging the two stimulation principles.

MISTIS Project-Team

6. New Results

6.1. Mixture models

6.1.1. Parameter estimation in the heterogeneity linear mixed model

Participant: Marie-José Martinez.

Joint work with: Emma Holian (National University of Ireland, Galway)

In studies where subjects contribute more than one observation, such as in longitudinal studies, linear mixed models have become one of the most used techniques to take into account the correlation between these observations. By introducing random effects, mixed models allow the within-subject correlation and the variability of the response among the different subjects to be taken into account. However, such models are based on a normality assumption for the random effects and reflect the prior belief of homogeneity among all the subjects. To relax this strong assumption, Verbeke and Lesaffre (1996) proposed the extension of the classical linear mixed model by allowing the random effects to be sampled from a finite mixture of normal distributions with common covariance matrix. This extension naturally arises from the prior belief of the presence of unobserved heterogeneity in the random effects population. The model is therefore called the heterogeneity linear mixed model. Note that this model does not only extend the assumption about the random effects distribution, indeed, each component of the mixture can be considered as a cluster containing a proportion of the total population. Thus, this model is also suitable for classification purposes.

Concerning parameter estimation in the heterogeneity model, the use of the EM-algorithm, which takes into account the incomplete structure of the data, has been considered in the literature. Unfortunately, the M-step in the estimation process is not available in analytic form and a numerical maximisation procedure such as Newton-Raphson is needed. Because deriving such a procedure is a non-trivial task, Komarek et al. (2002) proposed an approximate optimization. But this procedure proved to be very slow and limited to small samples due to requiring manipulation of very large matrices and prohibitive computation.

To overcome this problem, we have proposed in [28], [52] an alternative approach which consists of fitting directly an equivalent mixture of linear mixed models. Contrary to the heterogeneity model, the M-step of the EM-algorithm is tractable analytically in this case. Then, from the obtained parameter estimates, we can easily obtain the parameter estimates in the heterogeneity model.

6.1.2. Taking into account the curse of dimensionality

Participants: Stéphane Girard, Alessandro Chiancone, Seydou-Nourou Sylla.

Joint work with: C. Bouveyron (Univ. Paris 1), M. Fauvel (ENSAT Toulouse) and J. Chanussot (Gipsa-lab and Grenoble-INP)

In the PhD work of Charles Bouveyron (co-advised by Cordelia Schmid from the Inria LEAR team) [64], we propose new Gaussian models of high dimensional data for classification purposes. We assume that the data live in several groups located in subspaces of lower dimensions. Two different strategies arise:

- the introduction in the model of a dimension reduction constraint for each group
- the use of parsimonious models obtained by imposing to different groups to share the same values of some parameters

This modelling yields a new supervised classification method called High Dimensional Discriminant Analysis (HDDA) [4]. Some versions of this method have been tested on the supervised classification of objects in images. This approach has been adapted to the unsupervised classification framework, and the related method is named High Dimensional Data Clustering (HDDC) [3]. Our recent work consists in adding a kernel in the previous methods to deal with nonlinear data classification.

6.1.3. Mixture modelling using skewed multivariate heavy tailed distributions with variable amounts of tailweight

Participants: Florence Forbes, Darren Wraith.

Clustering concerns the assignment of each of N, possibly multidimensional, observations $y_1, ..., y_N$ to one of K groups. A popular way to approach this task is via a parametric finite mixture model. While the vast majority of the work on such mixtures has been based on Gaussian mixture models in many applications the tails of normal distributions are shorter than appropriate or parameter estimations are affected by atypical observations (outliers). The family of location and scale mixtures of Gaussians has the ability to generate a number of flexible distributional forms. It nests as particular cases several important asymmetric distributions like the Generalised Hyperbolic distribution. The Generalised Hyperbolic distribution in turn nests many other well known distributions such as the Normal Inverse Gaussian (NIG) whose practical relevance has been widely documented in the literature. In a multivariate setting, we propose to extend the standard location and scale mixture concept into a so called multiple scaled framework which has the advantage of allowing different tail and skewness behaviours in each dimension of the variable space with arbitrary correlation between dimensions. The approach builds upon, and develops further, previous work on scale mixtures of Gaussians [25]. Estimation of the parameters is provided via an EM algorithm with a particular focus on NIG distributions. Inference is then extended to cover the case of mixtures of such multiple scaled distributions for application to clustering. Assessments on simulated and real data confirm the gain in degrees of freedom and flexibility in modelling data of varying tail behaviour and directional shape.

6.1.4. High-Dimensional Regression with Gaussian Mixtures and Partially-Latent Response Variables

Participant: Florence Forbes.

Joint work with: Antoine Deleforge and Radu Horaud from the Inria Perception team.

In this work we address the problem of approximating high-dimensional data with a low-dimensional representation. We make the following contributions. We propose an inverse regression method which exchanges the roles of input and response, such that the low-dimensional variable becomes the regressor, and which is tractable. We introduce a mixture of locally-linear probabilistic mapping model that starts with estimating the parameters of inverse regression, and follows with inferring closed-form solutions for the forward parameters of the high-dimensional regression problem of interest. Moreover, we introduce a partially-latent paradigm, such that the vector-valued response variable is composed of both observed and latent entries, thus being able to deal with data contaminated by experimental artifacts that cannot be explained with noise models. The proposed probabilistic formulation could be viewed as a latent-variable augmentation of regression. We devise expectation-maximization (EM) procedures based on a data augmentation strategy which facilitates the maximum-likelihood search over the model parameters. We propose two augmentation schemes and we describe in detail the associated EM inference procedures that may well be viewed as generalizations of a number of EM regression, dimension reduction, and factor analysis algorithms. The proposed framework is validated with both synthetic and real data. We provide experimental evidence that our method outperforms several existing regression techniques.

6.1.5. Acoustic space learning via variational EM for Sound-Source Separation and Localization

Participant: Florence Forbes.

Joint work with: Antoine Deleforge and Radu Horaud from the Inria Perception team.

In this paper we address the problems of modeling the acoustic space generated by a full-spectrum sound source and of using the learned model for the localization and separation of multiple sources that simultaneously emit sparse-spectrum sounds. We lay theoretical and methodological grounds in order to introduce the *binaural manifold* paradigm. We perform an in-depth study of the latent low-dimensional structure of the high-dimensional interaural spectral data, based on a corpus recorded with a human-like audiomotor robot head. A non-linear dimensionality reduction technique is used to show that these data lie on a two-dimensional (2D) smooth manifold parameterized by the motor states of the listener, or equivalently, the sound source directions. We propose a *probabilistic piecewise affine mapping* model (PPAM) specifically designed to deal with high-dimensional data exhibiting an intrinsic piecewise linear structure. We derive a closed-form expectation-maximization (EM) procedure for estimating the model parameters, followed by Bayes inversion for obtaining the full posterior density function of a sound source direction. We extend this solution to deal with missing data and redundancy in real world spectrograms, and hence for 2D localization of natural sound sources such as speech. We further generalize the model to the challenging case of multiple sound sources and we propose a variational EM framework. The associated algorithm, referred to as *variational EM for source separation and localization* (VESSL) yields a Bayesian estimation of the 2D locations and time-frequency masks of all the sources. Comparisons of the proposed approach with several existing methods reveal that the combination of acoustic-space learning with Bayesian inference enables our method to outperform state-of-the-art methods.

6.2. Statistical models for Neuroscience

6.2.1. Hemodynamically informed parcellation of cerebral fMRI data

Participants: Florence Forbes, Aina Frau-Pascual, Thomas Vincent.

Joint work with: Philippe Ciuciu from Team Parietal and Neurospin, CEA in Saclay.

Standard detection of evoked brain activity in functional MRI (fMRI) relies on a fixed and known shape of the impulse response of the neurovascular coupling, namely the hemodynamic response function (HRF). To cope with this issue, the joint detection-estimation (JDE) framework has been proposed. This formalism enables to estimate a HRF per region but for doing so, it assumes a prior brain partition (or parcellation) regarding hemodynamic territories (eg. [14]). This partition has to be accurate enough to recover accurate HRF shapes but has also to overcome the detection-estimation issue: the lack of hemodynamics information in the non-active positions. During the internship of A. Frau Pascual at Neurospin, we proposed an hemodynamically-based parcellation, consisting first of a feature extraction step, followed by a Gaussian Mixture-based parcellation, which considers the injection of the activation levels in the parcellation process, in order to overcome the detection-estimation issue and find the underlying hemodynamics. The work has been submitted to the ICASSP conference in 2014.

6.2.2. Variational variable selection to assess experimental condition relevance in event-related fMRI

Participants: Florence Forbes, Christine Bakhous, Lotfi Chaari, Thomas Vincent, Farida Enikeeva.

Joint work with: Michel Dojat (Grenoble Institute of Neuroscience) and Philippe Ciuciu from Neurospin, CEA in Saclay.

Brain functional exploration investigates the nature of neural processing following cognitive or sensory stimulation. This goal is not fully accounted for in most functional Magnetic Resonance Imaging (fMRI) analysis which usually assumes that all delivered stimuli possibly generate a BOLD response everywhere in the brain although activation is likely to be induced by only some of them in specific brain regions. Generally, criteria are not available to select the relevant conditions or stimulus types (e.g. visual, auditory, etc.) prior to activation detection and the inclusion of irrelevant events may degrade the results, particularly when the Hemodynamic Response Function (HRF) is jointly estimated as in the JDE framework mentioned in the previous section. To face this issue, we propose an efficient variational procedure that automatically selects the conditions according to the brain activity they elicit. It follows an improved activation detection and local HRF estimation that we illustrate on synthetic and real fMRI data. This approach is an alternative to our previous approach based on Monte-Carlo Markov Chain (MCMC) inference [63]. Corresponding papers [31], [45]. A synthesis can also be found in the PhD manuscript of C. Bakhous (Grenoble University, December 2013) [11].

6.2.3. Bayesian Joint Detection-Estimation of cerebral vasoreactivity from ASL fMRI data

Participants: Florence Forbes, Thomas Vincent.

In the context of ARC AINSI project, joint work with: Philippe Ciuciu from Neurospin, CEA in Saclay.

Functional MRI (fMRI) is the method of choice to non-invasively probe cerebral activity evoked by a set of controlled experimental conditions. A rising fMRI modality is Arterial Spin Labeling (ASL) which enables to quantify the cerebral perfusion, namely the cerebral blood flow (CBF) and emerges as a more direct biomarker of neuronal activity than the standard BOLD (Blood Oxygen Level Dependent) fMRI.

Although the study of cerebral vasoreactivity using fMRI is mainly conducted through the BOLD fMRI modality (see the two previous sections), owing to its relatively high signal-to-noise ratio (SNR), ASL fMRI provides a more interpretable measure of cerebral vasoreactivity than BOLD fMRI. Still, ASL suffers from a low SNR and is hampered by a large amount of physiological noise. Our contribution, described in [43], [44] aims at improving the recovery of the vasoreactive component from the ASL signal. To this end, a Bayesian hierarchical model is proposed, enabling the recovery of perfusion levels as well as fitting their dynamics. On a single-subject ASL real data set involving perfusion changes induced by hypercapnia, the approach is compared with a classical GLM-based analysis. A better goodness-of-fit is achieved, especially in the transitions between baseline and hypercapnia periods. Also, perfusion levels are recovered with higher sensitivity and show a better contrast between gray- and white matter.

6.2.4. Physiologically-inspired Bayesian analysis of BOLD and ASL fMRI data

Participants: Florence Forbes, Thomas Vincent, Jennifer Sloboda.

In the context of ARC AINSI project, joint work with: Philippe Ciuciu from Neurospin, CEA in Saclay.

The ASL modality is most commonly used as a static measure where the average perfusion is computed over a volume sequence lasting several minutes. Recently, ASL has been used in functional activation protocols and hence gives access to a dynamic measure of perfusion, namely the variations of CBF which are elicited by specific tasks. ASL MRI mainly consists of acquiring pairs of control and label images and looking at the average control-label difference. The Signal-to-Noise Ratio (SNR) of this difference is very low so that several hundreds of image pairs need to be acquired, thus increasing significantly the time spent by the subject in the scanner and making the acquisition very sensitive to the patient's movement. In addition, this averaging requires that the perfusion signal is at a steady state, limiting the scope of fMRI task experiments to baseline perfusion measurements or long block designs. In contrast, it is highly desirable to measure change in perfusion due to an effect of interest in activation paradigms from event-related designs. It is technically possible to collect event-related ASL data but most approaches to functional ASL data analysis use a standard linear model (GLM-based) formulation with regressors encoding differences in control/tag scans and both ASL and BOLD activation signals being associated with the same canonical response function. The canonical hemodynamic response function (HRF) is generally used although it has been been calibrated on BOLD experiments only, thus reflecting simultaneous variations of CBF, cerebral blood volume (CBV) and cerebral oxygen consumption (CMRO2). In contrast, the perfusion signal only reflects variation in CBF so that the associated response, the perfusion response function (PRF), is likely to differ from the HRF. In the internship proposal of Jennifer Sloboda, we proposed to recover both a hemodynamic (BRF for BOLD response function) and a perfusion (PRF) response functions from event-related functional ASL data. To do so, a joint detection estimation (JDE) formalism was used. In the BOLD context, the JDE framework has proven to successfully extract the HRF while also performing activation detection. We had recently extended this formalism (see Section 6.2.3 and [43], [44]) to model an additional perfusion component linked to the BOLD one through a common activation detection. The main issue addressed then was to characterize the link between BOLD and perfusion components. To establish this link, we proposed a methodological axis which consists of developing a physiologically-inspired approach. To do so, dynamical non-linear equations available in physiological models were linearized and approximated in a parsimonious way so as to establish prior relations between the perfusion and BOLD responses which can be injected in our Bayesian setting. The inference of the initial model is currently done through a Markov Chain Monte Carlo approach but a Variational Expectation-Maximization implementation is also conceivable. As such, the tasks were two-fold: (1) investigate the physiological model and (2) inject it into the JDE setting. Investigation of the physiological model allows for: (1) creation of artificial fMRI data, (2) investigation of the relationship between physiological changes and the resulting simulated BOLD or ASL signal, and (3) characterization of the link between BOLD and perfusion responses. Injection of the physiologically inspired prior into the JDE model, is to (1) improve perfusion response recovery, (2) determine physiologically quantified units to the JDE recovered values This work is going to serve as a preliminary investigation into the incorporation of physiological information in the Bayesian JDE setting from which to determine the trajectory of future model developments.

6.3. Markov models

6.3.1. Spatial modelling of plant diversity from high-throughput environmental DNA sequence data

Participants: Florence Forbes, Angelika Studeny.

This is joint work with: Eric Coissac and Pierre Taberlet from LECA (Laboratoire d'Ecologie Alpine) and Alain Viari from Inria team Bamboo.

This work [48] considers a statistical modelling approach to investigate spatial cross-correlations between species in an ecosystem. A special feature is the origin of the data from high-troughput environmental DNA sequencing of soil samples. Here we use data collected at the Nourague CNRS Field Station in French Guiana. We describe bivariate spatial relationships in these data by a separable linear model of coregionalisation and estimate a cross-correlation parameter. Based on this estimate, we visualise plant taxa co-occurrence pattern in form of 'interaction graphs' which can be interpreted in terms of ecological interactions. Limitations of this approach are discussed along with possible alternatives in [48].

6.3.2. Modelling multivariate counts with graphical Markov models.

Participants: Jean-Baptiste Durand, Florence Forbes, Marie-José Martinez, Angelika Studeny.

Joint work with: Pierre Fernique (Montpellier 2 University, CIRAD and Inria Virtual Plants), Yann Guédon (CIRAD and Inria Virtual Plants) and Iragaël Joly (INRA-GAEL and Grenoble INP).

Multivariate count data are defined as the number of items of different categories issued from sampling within a population, which individuals are grouped into categories. The analysis of multivariate count data is a recurrent and crucial issue in numerous modelling problems, particularly in the fields of biology and ecology (where the data can represent, for example, children counts associated with multitype branching processes), sociology and econometrics. Denoting by K the number of categories, multivariate count data analysis relies on modelling the joint distribution of the K-dimensional random vector $N = (N_0, ..., N_{K-1})$ with discrete components. Our work focused on I) Identifying categories that appear simultaneously, or on the contrary that are mutually exclusive. This was achieved by identifying conditional independence relationships between the K variables; II)Building parsimonious parametric models consistent with these relationships; III) Characterizing and testing the effects of covariates on the distribution of N, particularly on the dependencies between its components.

Our context of application was characterised by zero-inflated, often right skewed marginal distributions. Thus, Gaussian and Poisson distributions were not *a priori* appropriate. Moreover, the multivariate histograms typically had many cells, most of which were empty. Consequently, nonparametric estimation was not efficient.

To achieve these goals, we proposed an approach based on graphical probabilistic models (Koller & Friedman, 2009 [70]) to represent the conditional independence relationships in N, and on parametric distributions to ensure model parsimony [51]. The considered graphs were partially directed, so as to represent both marginal independence relationships and cyclic dependencies between quadruplets of variables (at least).

Graph search was achieved by a stepwise approach, issued from unification of previous algorithms presented in Koller & Friedman (2009) for DAGs: Hill climbing, greedy search, first ascent and simulated annealing. The search algorithm was improved by taking into account our parametric distribution assumptions, which led to caching overlapping graphs at each step. An adaptation to PDAGs of graph search algorithms for DAGs was developed, by defining new operators specific to PDAGs.

Comparisons between different algorithms in the literature for directed and undirected graphical models was performed on simulated datasets to: (i) Assess gain in speed induced by caching; (ii) Compare the graphs obtained under parametric and nonparametric distributions assumptions; (iii) Compare different strategies for graph initialization. Strategies based on several random graphs were compared to those based on a fast estimation of an undirected graph, assumed to be the moral graph.

First results were obtained in modelling individual daily activity program [50] and interactions between flowering and vegetative growth in plants (see sections below).

6.3.3. Statistical characterization of tree structures based on Markov tree models and multitype branching processes, with applications to tree growth modelling.

Participant: Jean-Baptiste Durand.

Joint work with: Pierre Fernique (Montpellier 2 University and CIRAD) and Yann Guédon (CIRAD), Inria Virtual Plants.

The quantity and quality of yields in fruit trees is closely related to processes of growth and branching, which determine ultimately the regularity of flowering and the position of flowers. Flowering and fruiting patterns are explained by statistical dependence between the nature of a parent shoot (*e.g.* flowering or not) and the quantity and natures of its children shoots – with potential effect of covariates. Thus, better characterization of patterns and dependences is expected to lead to strategies to control the demographic properties of the shoots (through varietal selection or crop management policies), and thus to bring substantial improvements in the quantity and quality of yields.

Since the connections between shoots can be represented by mathematical trees, statistical models based on multitype branching processes and Markov trees appear as a natural tool to model the dependencies of interest. Formally, the properties of a vertex are summed up using the notion of vertex state. In such models, the numbers of children in each state given the parent state are modeled through discrete multivariate distributions. Model selection procedures are necessary to specify parsimonious distributions. We developed an approach based on probabilistic graphical models (see Section 6.3.2) to identify and exploit properties of conditional independence between numbers of children in different states, so as to simplify the specification of their joint distribution [51], [32].

This work was carried out in the context of Pierre Fernique's first year of PhD (Montpellier 2 University and CIRAD). It was applied to model dependencies between short or long, vegetative or flowering shoots in apple trees. The results highlighted contrasted patterns related to the parent shoot state, with interpretation in terms of alternation of flowering (see paragraph 6.3.4). It was also applied to the analysis of the connections between cyclic growth and flowering of mango trees [32]. This work will be continued during Pierre Fernique's PhD thesis, with extensions to other fruit tree species and other parametric discrete multivariate families of distributions, including covariates and mixed effects.

6.3.4. Statistical characterization of the alternation of flowering in fruit tree species

Participant: Jean-Baptiste Durand.

Joint work with: Jean Peyhardi and Yann Guédon (Mixed Research Unit DAP, Virtual Plants team), Baptiste Guitton, Yan Holtz and Evelyne Costes (DAP, AFEF team), Catherine Trottier (Montpellier University)

A first study was performed to characterize genetic determinisms of the alternation of flowering in apple tree progenies [37], [21]. Data were collected at two scales: at whole tree scale (with annual time step) and a local scale (annual shoot or AS, which is the portions of stem that were grown during the same year). Two replications of each genotype were available.

Indices were proposed to characterize alternation at tree scale. The difficulty is related to early detection of alternating genotypes, in a context where alternation is often concealed by a substantial increase of the number of flowers over consecutive years. To separate correctly the increase of the number of flowers due to aging of young trees from alternation in flowering, our model relied on a parametric hypothesis for the trend (fixed slopes specific to genotype and random slopes specific to replications), which translated into mixed effect modelling. Then, different indices of alternation were computed on the residuals. Clusters of individuals with contrasted patterns of bearing habits were identified.

To model alternation of flowering at AS scale, a second-order Markov tree model was built. Its transition probabilities were modelled as generalized linear mixed models, to incorporate the effects of genotypes, year and memory of flowering for the Markovian part, with interactions between these components.

Asynchronism of flowering at AS scale was assessed using an entropy-based criterion. The entropy allowed for a characterisation of the roles of local alternation and asynchronism in regularity of flowering at tree scale.

Moreover, our models highlighted significant correlations between indices of alternation at AS and individual scales.

This work was extended by the Master 2 internship of Yan Holtz, supervised by Evelyne Costes and Jean-Baptiste Durand. New progenies were considered, and a methodology based on a lighter measurement protocol was developed and assessed. It consisted in assessing the accuracy of approximating the indices computed from measurements at tree scale by the same indices computed as AS scale. The approximations were shown sufficiently accurate to provide an operational strategy for apple tree selection.

As a perspective of this work, patterns in the production of children ASs (numbers of flowering and vegetative children) depending on the type of the parent AS must be analyzed using branching processes and different types of Markov trees, in the context of Pierre Fernique's PhD Thesis (see paragraph 6.3.3).

6.4. Semi and non-parametric methods

6.4.1. Modelling extremal events

Participants: Stéphane Girard, El-Hadji Deme.

Joint work with: L. Gardes (Univ. Strasbourg) and E. Deme (Univ. Gaston Berger, Sénégal)

We are working on the estimation of the second order parameter ρ (see paragraph 3.3.1). We proposed a new family of estimators encompassing the existing ones (see for instance [69], [68]). This work is in collaboration with El-Hadji Deme who obtained a grant (IBNI price) to work within the Mistis team on extreme-value statistics. The results are published in [18].
In addition to this work, we have established a review on the Weibull-tail distributions [29].

6.4.2. Conditional extremal events

Participants: Stéphane Girard, Gildas Mazo, Jonathan El-Methni.

Joint work with: L. Gardes (Univ. Strasbourg) and A. Daouia (Univ. Toulouse I and Univ. Catholique de Louvain)

The goal of the PhD thesis of Alexandre Lekina was to contribute to the development of theoretical and algorithmic models to tackle conditional extreme value analysis, *ie* the situation where some covariate information X is recorded simultaneously with a quantity of interest Y. In such a case, the tail heaviness of Y depends on X, and thus the tail index as well as the extreme quantiles are also functions of the covariate. We combine nonparametric smoothing techniques [66] with extreme-value methods in order to obtain efficient estimators of the conditional tail index and conditional extreme quantiles. When the covariate is functional and random (random design) we focus on kernel methods [16].

Conditional extremes are studied in climatology where one is interested in how climate change over years might affect extreme temperatures or rainfalls. In this case, the covariate is univariate (time). Bivariate examples include the study of extreme rainfalls as a function of the geographical location. The application part of the study is joint work with the LTHE (Laboratoire d'étude des Transferts en Hydrologie et Environnement) located in Grenoble.

6.4.3. Estimation of extreme risk measures

Participants: Stéphane Girard, Jonathan El-Methni, El-Hadji Deme.

Joint work with: L. Gardes and A. Guillou (Univ. Strasbourg)

One of the most popular risk measures is the Value-at-Risk (VaR) introduced in the 1990's. In statistical terms, the VaR at level $\alpha \in (0,1)$ corresponds to the upper α-quantile of the loss distribution. The Value-at-Risk however suffers from several weaknesses. First, it provides us only with a pointwise information: VaR(α) does not take into consideration what the loss will be beyond this quantile. Second, random loss variables with light-tailed distributions or heavy-tailed distributions may have the same Value-at-Risk . Finally, Value-at-Risk is not a coherent risk measure since it is not subadditive in general. A coherent alternative risk measure is the Conditional Tail Expectation (CTE), also known as Tail-Value-at-Risk, Tail Conditional Expectation or Expected Shortfall in case of a continuous loss distribution. The CTE is defined as the expected loss given that the loss lies above the upper α-quantile of the loss distribution. This risk measure thus takes into account the whole information contained in the upper tail of the distribution. It is frequently encountered in financial investment or in the insurance industry. In [36], we have established the asymptotic properties of the classical CTE estimator in case of extreme losses, *i.e.* when $\alpha \to 0$ as the sample size increases. We have exhibited the asymptotic bias of this estimator, and proposed a bias correction based on extreme-value techniques [36]. Similar developments have been achieved in the case of the Proportional Hazard Premium measure of risk [19]. In [22], we study the situation where some covariate information is available. We thus has to deal with conditional extremes (see paragraph 6.4.2). We also proposed a new risk measure (called the Conditional Tail Moment) which encompasses various risk measures, such as the CTE, as particular cases.

6.4.4. Multivariate extremal events

Participants: Stéphane Girard, Gildas Mazo, Florence Forbes, Van Trung Pham.

Joint work with: C. Amblard (TimB in TIMC laboratory, Univ. Grenoble I) and L. Menneteau (Univ. Montpellier II)

Copulas are a useful tool to model multivariate distributions [72]. At first, we developed an extension of some particular copulas [1]. It followed a new class of bivariate copulas defined on matrices [56] and some analogies have been shown between matrix and copula properties.

However, while there exist various families of bivariate copulas, much fewer has been done when the dimension is higher. To this aim an interesting class of copulas based on products of transformed copulas has been proposed in the literature. The use of this class for practical high dimensional problems remains challenging. Constraints on the parameters and the product form render inference, and in particular the likelihood computation, difficult. We proposed a new class of high dimensional copulas based on a product of transformed bivariate copulas [61]. No constraints on the parameters refrain the applicability of the proposed class which is well suited for applications in high dimension. Furthermore the analytic forms of the copulas within this class allow to associate a natural graphical structure which helps to visualize the dependencies and to compute the likelihood efficiently even in high dimension. The extreme properties of the copulas are also derived and an R package has been developed.

As an alternative, we also proposed a new class of copulas constructed by introducing a latent factor. Conditional independence with respect to this factor and the use of a nonparametric class of bivariate copulas lead to interesting properties like explicitness, flexibility and parsimony. In particular, various tail behaviours are exhibited, making possible the modeling of various extreme situations. A pairwise moment-based inference procedure has also been proposed and the asymptotic normality of the corresponding estimator has been established [53].

6.4.5. Level sets estimation

Participant: Stéphane Girard.

Joint work with: A. Guillou and L. Gardes (Univ. Strasbourg), G. Stupfler (Univ. Aix-Marseille) and A. Daouia (Univ. Toulouse I and Univ. Catholique de Louvain)

The boundary bounding the set of points is viewed as the larger level set of the points distribution. This is then an extreme quantile curve estimation problem. We proposed estimators based on projection as well as on kernel regression methods applied on the extreme values set, for particular set of points [10]. We also investigate the asymptotic properties of existing estimators when used in extreme situations. For instance, we have established in collaboration with G. Stupfler that the so-called geometric quantiles have very counter-intuitive properties in such situations [60] and thus should not be used to detect outliers.
In collaboration with A. Daouia, we investigate the application of such methods in econometrics [17]: A new characterization of partial boundaries of a free disposal multivariate support is introduced by making use of large quantiles of a simple transformation of the underlying multivariate distribution. Pointwise empirical and smoothed estimators of the full and partial support curves are built as extreme sample and smoothed quantiles. The extreme-value theory holds then automatically for the empirical frontiers and we show that some fundamental properties of extreme order statistics carry over to Nadaraya's estimates of upper quantile-based frontiers.
In collaboration with G. Stupfler and A. Guillou, new estimators of the boundary are introduced. The regression is performed on the whole set of points, the selection of the "highest" points being automatically performed by the introduction of high order moments [26], [27].

6.4.6. Retrieval of Mars surface physical properties from OMEGA hyperspectral images.

Participants: Stéphane Girard, Alessandro Chiancone.

Joint work with: S. Douté from Laboratoire de Planétologie de Grenoble, J. Chanussot (Gipsa-lab and Grenoble-INP) and J. Saracco (Univ. Bordeaux).

Visible and near infrared imaging spectroscopy is one of the key techniques to detect, to map and to characterize mineral and volatile (eg. water-ice) species existing at the surface of planets. Indeed the chemical composition, granularity, texture, physical state, etc. of the materials determine the existence and morphology of the absorption bands. The resulting spectra contain therefore very useful information. Current imaging spectrometers provide data organized as three dimensional hyperspectral images: two spatial dimensions and one spectral dimension. Our goal is to estimate the functional relationship F between some observed spectra and some physical parameters. To this end, a database of synthetic spectra is generated by a physical radiative transfer model and used to estimate F. The high dimension of spectra is reduced by Gaussian regularized sliced inverse regression (GRSIR) to overcome the curse of dimensionality and consequently the sensitivity of the inversion to noise (ill-conditioned problems) [57]. We have also defined an adaptive version of the method which is able to deal with block-wise evolving data streams [15].

6.4.7. High-dimensional change-point detection with sparse alternatives

Participant: Farida Enikeeva.

Joint work with: Zaid Harchaoui from LEAR team Inria Grenoble

The change-point problem is a classical problem of statistics that arises in various applications as signal processing, bioinformatics, financial market analysis. The goal of change-point problems is to make an inference about the moment of a change in the distribution of the observed data. We consider the problem of detection of a simultaneous change in mean in a sequence of Gaussian vectors.

The state-of-the-art approach to the change-point detection/estimation is based on the assumption of growing number of observations and fixed dimension of the signal. We work in high-dimensional setting assuming that the vector dimension tends to infinity and the length of the sequence grows slower than the dimension of the signal. Assuming that the change occurs only in a subset of the vector components of unknown cardinality we can reduce our problem to the problem of testing non-zero components in a sequence of sparse Gaussian vectors. We construct a testing procedure that is adaptive to the number of components with a change. This testing procedure is based on combination of two chi-squared type test statistics. This combined test provides an optimal performance of the test both in the cases of high and moderate sparsity. We obtain the detection boundary of the test and show its rate-optimality in minimax sense.

The results of the paper [59] were presented at

- NIPS 2013, Workshop on Modern Nonparametric Methods in Machine Learning (Dec. 2013)
- Conference on Structural Inference in Statistics, Potsdam, Germany (Sept. 2013)

6.4.8. Yield Improvement by the Redundancy Method for Component Calibration

Participant: Farida Enikeeva.

Joint work with: Dominique Morche (CEA-LETI) and Alp Oguz (CEA-LETI)

This work [23] was done in the framework of the Optimyst II project of MINALOGIC in collaboration with CEA-LETI and LJK-UJF. In this project we explore the benefits of the redundant channels methodology for the calibration of electronic components.

The demand for high data rate in communication puts stringent requirements on components' dynamic range. However, the extreme size reduction in advanced technology results inadvertently in increased process variability, which inherently limits the performances. The redundancy approach is based on the idea of dividing an elementary component (capacitor, resistor, transistor) into several subsets and then choosing an optimal combination of such subsets to provide the production of a component with very precise characteristics. For several years, the redundancy method has been identified as complementary to digital calibration to improve the performances. On practice, it is hard for a designer to select an optimal number of redundant components to provide the desired production yield and to minimize the area occupied by the components. The usual way to solve this problem is to resort to statistical simulations which are time consuming and sometimes misleading. We propose a normal approximation of the yield in order to estimate the number of redundant components needed to provide a minimal area occupied by the components.

MNEMOSYNE Team

6. New Results

6.1. Overview

Though our view is systemic, our daily research activities are concerned with the design, at a given scale of description, of models of neuronal structures, each concerned with a specific learning paradigm. Of course, a major challenge is to keep in mind the systemic view, to put a specific emphasis on the way each neuronal structure communicates with the rest of the system and to highlight how the learning paradigm interplays with other memory systems.

Among the numerous loops involving the brain, the body and the environment, a basic grid of description corresponds to distinguish “Perception Loops”, the goal of which is to extract from the inner and outer world sensory invariants helpful to identify and evaluate the current state and to make predictions from previous learning, and “Action Loops”, the goal of which is to rely on this sensory, emotional and motivational information to decide, plan and trigger actions for the benefit of the body.

Presently, our team is engaged on the following topics: Concerning perception loops, we are firstly considering the role of the hippocampus and of the posterior cortex in learning high level sensory cues that contribute to pavlovian conditioning in the amygdala. Secondly, we are investigating the role of the thalamus in attentional shifts in the cortex. This latter topic relies on recent advances we made in the Keops ANR project (*cf.* § 7.1) on a model of the retina that we also sum up here. Concerning Action loops, we are preparing a critical analysis of the current views of the interactions between the prefrontal cortex and the basal ganglia. Finally, we also report here more methodological achievements.

6.2. Pavlovian conditioning

The fundamental role of the amygdala in pavlovian conditioning is widely acknowledged, both on the motor, autonomic and hormonal expression of pavlovian responses and on the learning of the associations between conditional and unconditioned stimuli. This year, we have proposed models showing in a systemic view how errors of prediction in the amygdala might trigger, by cholinergic hormonal expression, episodic learning in the hippocampus [17] and semantic learning and attentional shift in the posterior cortex [8]. On this basis, we are currently studying how emotional values of sensory information are dynamically encoded in the basolateral nucleus of the amygdala [11], to fit with the multiple requirements for this kind of information in the brain.

6.3. The thalamus is more than a relay

Many recent results in neuroscience indicate that the role of the thalamus in the brain is certainly more important than it used to be considered, particularly concerning its relation with the cortex. Interestingly, we considered this question as a side effect of our work in the Keops project (*cf.* § 7.1) with our chilean neuroscientist colleagues studying non standard ganglion cells in the retina. Our modeling [6] and bibliographic studies led us to propose a biologically-founded algorithm [13] for the interplay between the modulatory and driving connections between the thalamus and the cortex, in the case of the projection of these ganglion cells on the thalamus [14]. This study has been carried out with the strong constraint of proposing a system working on a real visual flow. In the near future, we aim at developing this original view of the thalamus in the more general case of its pulvinar associative nucleus, learning to route to the cortex multimodal information flows.

6.4. Eyes are really smarter than believed

While it is known that the retina has not only standard ganglion cells devoted to contrast (parvo) or motion (magno) cues, but is able to perform sophisticated detection of spatial or temporal events in the visual scene (konio) [40], it was still to understand how such computation could be implemented as a robust effective processing of realistic natural image sequence and not only as a toy model of cartoonish stimuli. It has been shown in our group [19] that biologically plausible variational models of non-linear filtering coupled with optimized simple threshold mechanisms as derived from statistical learning mechanisms provides an efficient and realistic simulation of such non-standard retina output as observable at the level of konio-cells [14].

6.5. On the computational efficiency of Basal Ganglia models

Many valuable models have been proposed to capture the richness of the fundamental relations between the basal ganglia and prominent brain structures including the prefrontal cortex, the hippocampus and the superior colliculus. To choose among them the mechanisms on which to build the design of the motor pole of our brain-inspired system, a fundamental issue is to evaluate the efficiency of these models in more realistic cases than the ones which are generally considered by the authors. For this reason, we are presently preparing a comparative study of models, including a model developed in our lab with neuroscientist colleagues [2], in the more realistic case of large sensory and motor flows.

6.6. From distributed computing to distributed computing

One of the challenges in our systemic approach is to promote behaviors using distributed adaptive numerical computing that prevents (by definition) the specification of any behavior at a global level (no homunculus). This paradigm has been found to be very close to some class of problems in computer graphics where one tries to achieve a specific effect at the whole image level while the actual implementation is made at the pixel level (fragment). Transposing our expertise in distributed numerical computing, we thus experiment GPU programming (vertex/fragment programming) that offers hardware enforcement of distributed constraints: every fragment get the same program but their combination promotes a global image effect. This has been done for the case of text rendering [4] and dashed line rendering [5].

MOAIS Project-Team

6. New Results

6.1. Distributed Art Performance

Moais collaborated with partners fromI2cat, Barcelona, Psnc, Poznan and Grenoble-INP to setup a live distributed art performance for the ICT 2013 conference at Vilnius. This distributed performance gathered musicians located a Poznan, Barcelona and Vilnius, as well as a dancer modeled in 3D on the Grimage platform at Inria Grenoble. Though physically present in different cities these artists performed together for a numerical dance and music performance numerically assembled and transmitted in real-time at Vilnius. This joint effort relies on the FlowVR framework from Moais and the UltraGrid software from CESNET. This event received a significant attention from the medias (In France: FR3 and Tele-grenoble, France inter, etc.). A video is available at http://cyan1.grenet.fr/podcastmedia/Visionair/ICT2013_promo.m4v.

6.2. VTK Parallelization Framework

Moais developed a framework for the parallelization of scientific visualization algorithms based on on-demand task extraction and work stealing techniques. This work is developed for the VTK software and supports the OpenMP, Intel TBB and Kaapi runtime environments. Mathias Ettinger visited the Kitware company, NY, for two months to prepare the integration of this work in the next release of VTK. This work is performed in collaboration with the EDF company.

6.3. Parallel Sorting Algorithm

We developped a novel adaptive sorting algorithm, called PaVO, relying on a Packed Mermory Array data structure. Maintaining gaps in the array of elements enable to reduce the span of modifications needed when reordering elements. This is particularly relevant in a parallel context to reduce the data dependencies. Performance results on a NUMA architecute show that PaVo outperfoms standard parallel sorting algorithms even for a large amount of disorder.

6.4. High bandwidth IPSec gateways and ICMP

Internet Control Message Protocol (ICMP)i is essential for performance aspects in particular for Path Maximum Transmission Unit discovery but is also known to be a cause of attacks. In collaboration with Planet, we demonstrate, through a real exploit on a testbed, that an external attacker having eavesdropping and traffic injection capabilities in the black untrusted network, without any access to clear-text (thesis of Ludovic Jacquin). This impacts out current research on trusted outsourced computations.

6.5. Efficient Parallel multi-GPUs execution

We developped a novel scheduling algorithm in Kaapi to perform multi-GPUs execution of task' based program [19]. Performance results on Cholesky factorization on up to 8-GPUs shows that Kaapi outperfoms similar runtime systems and even hand code parallel version.

6.6. Porting Kaapi for Native Mode on Intel Xeon Phi

Kaapi was ported natively on Intel Xeon Phi co-processor. Specific memory hierarchy was managed transparently to the applicaiton by the development of specific hierarchical work stealing scheduler. Experimentations on dense linear algebra kernels (Cholesky, LU and QR factorization) shows a very promising gain compared to the standard parallel implementation available in the Intel MKL [16]. Extension of theses resultats are under publication process.

6.7. Adaptive loop scheduling in GCC OpenMP runtime library

We port an adaptive loop scheduler from Kaapi into the OpenMP runtime library of GCC called libGOMP [12]. The loop scheduler is consencious of the bloc data mapping to improve locality of computation.

6.8. Kaapi in EPX standard distribution

Kaapi software developped by the MOAIS team was included in the standard EPX distribution. EPX has won the 2013 Grand Prix SFEN (http://www-epx.cea.fr).

MODAL Project-Team

6. New Results

6.1. Resampling procedures

Participant: Alain Celisse.

The new deep understanding of cross-validation procedures in density estimation has been tackled with new results in terms of risk estimation and model selection [7]. This is the first step towards a fully data-driven and optimal choice of cross-validation strategy.

6.2. Kernel change-point

Participants: Alain Celisse, Guillemette Marot, Morgane Pierre-Jean.

On the basis of theoretical arguments, an empirical analysis has been carried out to assess the influence of the choice of the kernel in the kernel change-point strategy described in [2]. This assessment has been done in the biological context of copy number variation and allele B fraction. Several talks have been given in seminars (SSB seminar in Paris,...) and workshops (JSFDS, SMPGD,...)

6.3. Gaussian process in RKHS

Participants: Alain Celisse, Jérémie Kellner.

Since numerous papers make a Gaussian assumption for observations in the reproducing kernel Hilbert space (RKHS), it is important to be able to assess the validity of this crucial assumption. As long as it has been validated, the Gaussian framework can be further used to infer statistical properties of the population at hand (mean, variance,...).

A statistical test has been designed to address such questions at the RKHS level. It is fully computationally efficient and provides really good power in numerous settings. Theoretical properties for the test statistic have been derived as well.

6.4. Model for conditionally correlated categorical data

Participants: Christophe Biernacki, Matthieu Marbac-Lourdelle, Vincent Vandewalle.

It is a model-based clustering where categorical data are grouped into conditionally independent blocks. The corresponding block distribution is a parsimonious multinomial distribution where the few free parameters correspond to the most likely modality crossings, while the remaining probability mass is uniformly spread over the other modality crossings. The exact computation of the integrated complete-data likelihood allows to perform the model selection, by a Gibbs sampler, reducing the computing time consuming by parameter estimation and avoiding BIC criterion biases pointed out by our experiments.

This model was presented in a conference [13] with scientific committee and in a seminar [17]. An article will be soon submitted. Furthermore, a R package is currently under development.

6.5. Mixture model for mixed kind of data

Participants: Christophe Biernacki, Matthieu Marbac-Lourdelle, Vincent Vandewalle.

A mixture model of Gaussian copula allows to cluster mixed kind of data. Each component is composed by classical margins while the conditional dependencies between the variables is modeled by a Gaussian copula. The parameter estimation is performed by a Gibbs sampler. This model was presented in a conference [14]. Some technical points will be developed before providing an article.

6.6. Mixture of Gaussians with Missing Data

Participants: Christophe Biernacki, Vincent Vandewalle.

The generative models allow to handle missing data. This can be easily performed by using the EM algorithm, which has a closed form M-step in the Gaussian setting. This can for instance be useful for distance estimation with missing data. It has been proposed to improve the distance estimation by fitting a mixture of Gaussian distributions instead of a considering only one Gaussian component [21]. This is a joined work with Emil Eirola and Amaury Lendrasse .

A parallel work is in progress on the mixture degeneracy when considering mixture of Gaussians with missing data. It have been experimentally noticed that the degeneracy in this case is particularly slow. This behaviour is different from the usual setting of degeneracy with mixture of Gaussians which is usually rather fast. A first attempt of the theoretical characterization of this behaviour around a degenerated solution has been presented at a conference [16].

6.7. Transfert learning in model-based clustering

Participant: Christophe Biernacki.

In many situations one needs to cluster several datasets, possibly arising from different populations, instead of a single one, into partitions with identical meaning and described by similar features. Such situations involve commonly two kinds of standard clustering processes. The samples are clustered traditionally either as if all units arose from the same distribution, or on the contrary as if the samples came from distinct and unrelated populations. But a third situation should be considered: As the datasets share statistical units of same nature and as they are described by features of same meaning, there may exist some link between the samples. We propose a linear stochastic link between the samples, what can be justified from some simple but realistic assumptions, both in the Gaussian and in the t mixture model-based clustering context [26]. This is a joint work with Alexandre Lourme.

6.8. Gaussian Models Scale Invariant and Stable by Projection

Participant: Christophe Biernacki.

Gaussian mixture model-based clustering is now a standard tool to determine an hypothetical underlying structure into continuous data. However many usual parsimonious models, despite their appealing geometrical interpretation, suffer from major drawbacks as scale dependence or unsustainability of the constraints by projection. In this work we present a new family of parsimonious Gaussian models based on a variance-correlation decomposition of the covariance matrices. These new models are stable by projection into the canonical planes and, so, faithfully representable in low dimension. They are also stable by modification of the measurement units of the data and such a modification does not change the model selection based on likelihood criteria. We highlight all these stability properties by a specific geometrical representation of each model. A detailed GEM algorithm is also provided for every model inference. Then, on biological and geological data, we compare our stable models to standard geometrical ones.

This joint work with Alexandre Lourme is now published in [6].

6.9. Clustering and variable selection in regression

Participants: Christophe Biernacki, Loïc Yengo, Julien Jacques.

A new framework is proposed to address the issue of simultaneous linear regression and clustering of predictors where regression coefficients are assumed to be drawn from a Gaussian mixture distribution. Prediction is thus performed using the conditional distribution of the regression coefficients given the data, while clusters are easily derived from posterior distribution in groups given the data. This work is now published in [28]

6.10. An AIC-like criterion for semi-supervised classification

Participants: Christophe Biernacki, Vincent Vandewalle.

In semi-supervised classification, generative models take naturally into account unlabeled data and parameter estimation can be easily performed through the EM algorithm. However, traditional model selection criteria either does not take into consideration the predictive purpose (AIC or BIC criteria) or involve a high computational cost because of the EM mechanism (cross validation criteria). Alternatively, we propose the penalized model selection criterion AICcond which aims to estimate the predictive power of a generative model by approximating its predictive deviance. AICcond has similar computational cost to AIC, owns good consistency theoretical properties and highlights encouraging behaviour for variable and model selection in comparison to other standard criteria.

This joint work with Gilles Celeux and Gérard Govaert is now published in[16].

6.11. Consistency of a nonparametric conditional mode estimator for random fields

Participant: Sophie Dabo-Niang.

Sophie Dabo-Niang settled the consistency of a nonparametric conditional mode estimator for random fields, Statistical Methods and Applications [9].

6.12. Spatial linear models

Spatial linear models only capture global linear relationships between locations. However, in many circumstances the spatial dependency is not linear. It is, for example, the classical case where one deals with the spatial pattern of extreme events such as in the economic analysis of poverty, in the environmental science,... This leads naturally to consider nonparametric modeling.

6.13. Auto-associative models

Serge Iovleff gave a complete treatment of the Auto-Associative models in the semi-linear case and wrote a software for estimating these models (hal-00734070, version 1).

6.14. BlockCluster

Serge Iovleff has submitted a paper on the BlockCluster package in collaboration with Parmeet Bathia.

6.15. Rmixmod

Serge Iovleff has contributed to a paper submitted to JSS (hal-00919486, version 1) in collaboration with R. Lebret, F. Langrognet, C. Biernacki, G. Celeux, and G. Govaert.

6.16. Clustering for functional data

Participants: Julien Jacques, Cristian Preda.

In Jacques & Preda 2014 (CSDA), we propose a model-based clustering algorithm for multivariate functional data, based on multivariate functional principal components analysis. A review on clustering for functional data has also be published in Jacques & Preda 2014 (ADAC). Variable selection in high-dimensional regression Participants: Julie Hamon, Julien Jacques, Clarisse Dhaenens. In the context of genomic analysis, dealing with high-throughput genotyping data, we develop a genetic algorithm which looks for the best subset of variables (of given size) to predict some quantitative feature.

6.17. Wavelet based clustering using mixed effects functional models

Participant: Guillemette Marot.

The paper related to the wavelet based clustering procedure presented in the activity report from MODAL team in 2012 was published in Biometrics [22].

6.18. Differential meta-analysis of RNA-seq data from multiple studies

Participant: Guillemette Marot.

An adaptation of meta-analysis methods intially proposed for microarray studies has been proposed for RNA-seq data. The R package metaRNASeq is available on the R Forge and the preprint of the paper is available on Arxiv [48].

6.19. Toxoplasma transcription factor TgAP2XI-5 regulates the expression of genes involved in parasite virulence and host invasion

Participant: Guillemette Marot.

The use of peak detection methods implemented in the Bioconductor package `Ringo` has enabled to better understand part of the gene regulation process in T. Gondii parasite. The new findings in Biology have been published in *Walker (2013)*.

MODEMIC Project-Team

6. New Results

6.1. Macroscopic models

6.1.1. About species coexistence

Participants: Fabien Campillo, Jérôme Harmand, Claude Lobry, Alain Rapaport, Tewfik Sari.

The so called "Principle of Competitive Exclusion" states that in the chemostat model, in presence of p substrates only p species can coexist. By contrast, in a bioreactor used for decontamination, hundreds to thousand different species are observed in presence of just very few substrates. Actually the classical chemostat models rely on assumptions: perfect mixing, substrate-dependent growth rate, constant environment, only asymptotic results are considered, deterministic continuous models...

A long term objective since Mere Inria project-team is to revisit the chemostat model in the absence of one or more of these hypotheses having in view the question of coexistence. In our "major publications" we proved coexistence in absence of the second hypothesis [6] or during long transient [9]. In [57], we consider the case where the environment (in some sense) is periodic in time. Our results concerning non continuous and/or stochastic models (see Section 6.2.3) are also a first step in avoiding the fifth hypothesis.

6.1.2. Modeling and analysis of bioprocesses

Participants: Boumédiène Benyahia, Radhouane Fekih-Salem, Jérôme Harmand, Claude Lobry, Guilherme Pimentel, Alain Rapaport, Tewfik Sari.

Within the supervision of the PhD thesis of R. Fekih-Salem, we have studied a chemostat model where the species are present in two forms, isolated and aggregated individuals, such as attached bacteria in biofilm or bacteria in flocks [22]. We show that our general model contains a lot of models that were previously considered in the literature [77]. Assuming that flocculation and deflocculation dynamics are fast compared to the growth of the species, we construct and analyse a reduced chemostat-like model in which both the growth functions and the apparent dilution rate depend on the density of the species.

Within the framework of the PhD thesis of B. Benyahia, we have included the fouling dynamics of membranes into the AM2 (or AMOCO) model and we have analyzed the resulting model (called the AM2b) [15]. In particular, we have integrated into this model the production and the degradation of Soluble Microbial Products (SMP), which are known to play an important role in the membrane fouling phenomenon. We show that under some general assumptions, the AM2b model has the same number of equilibria as the AM2 model and can exhibit bi-stability. However, under certain operating conditions or if biological parameters values are slightly modified, the AM2b model exhibits equilibria bifurcations and multi-stability properties.

The available anaerobic digestion models used for control purposes do usually only consider soluble matter. In fact, part of the pollutants are not soluble but are under a particulate form. In order to establish whether adding the dynamics of such matter into the models is important for the system behavior or not, we have studied new anaerobic models and established that depending on the kinetics of this additional reaction step, the qualitative behavior of the process may be significantly modified [44].

This year, G. Pimentel as started a PhD co-supervised with the University of Mons, about modeling of the membrane fouling in bioreactors in view of control. The objective is to represent cake formation and air cross-flow as a manipulated variable in the models, in view of future studies of control strategies for improving the efficiency of MBR processes [35], [47], [52], [53].

6.1.3. Ecosystem functioning in heterogeneous environments

Participants: Céline Casenave, Jérôme Harmand, Alain Rapaport.

This year, we have carried out a study of particular spatial interconnections such as "buffered" configurations, and its ecological impacts in terms of setup of a species in environments that are unfavorable when perfectly mixed. We have extended our previous results about the design of configurations for obtaining a global stability [28]. New conditions have been obtained for a species to setup when it is impossible in a perfectly mixed environment. At the opposite, we have characterized configurations that could destabilize bioprocesses.

With UMR Géosciences (Univ. of Rennes 1), we have carried on our analysis of the equivalence of two soil fracture models in terms of transfer functions [19]: the MINC (Multiple INteractiong Continua) and MRMT (Multiple Rate Mass Transfer) models that are quite popular in soil hydrodynamics. We have shown that a strict equivalence can be obtained if one considers different volumes in the discretization of the MINC model. For the moment, this study concerns the transfer of abiotic substances only.

In soil ecosystems, it often happens that several functional groups can be detected to operate concomitantly. We have investigated the mathematical properties of a relatively simple model that has been proposed by the UREP lab (Inra Clermont), that distinguishes explicitly two functional groups of micro-organisms: the decomposers of SOM (soil organic mater) and the producers of SOM, and compared it with a single microbial compartment model in terms of prediction of the so-called "priming effect" [27].

Together with agronomists of the UMR Eco & Sols (Cirad, Inra, IRD, Montpellier SupAgro) and the supervision of the MSc thesis of C. Droin, we have proposed and started to study a new model of consumer/resource for soil microbial ecosystems, in which we explicitly distinguish available and recalcitrant resources [71].

6.2. Stochastic and hybrid models

6.2.1. Stochastic macroscopic models

Participants: Fabien Campillo, Marc Joannides.

We continued our study of stochastic modeling of the chemostat. In a first study we establish the Fokker-Planck equation of the law of the diffusion process. This equation features relevant boundary condition for the washout. We propose specific finite difference schemes to account for this feature [18]. In a second work we adopt the same approach to more accurately study the logistic model [64] which allowed us to propose estimation procedures to take into account the extinction (see Section 6.3.2).

6.2.2. From microscopic models to macroscopic laws

Participants: Fabien Campillo, Coralie Fritsch, Jérôme Harmand, Claude Lobry.

We proposed a chemostat model where the bacterial population is individually-based (IBM), each bacterium is explicitly represented and has a mass evolving continuously over time, and where the substrate concentration is represented as a conventional ordinary differential equation. These two components are coupled with the bacterial consumption. Mechanisms acting on the bacteria are explicitly described (growth, division and uptake). Bacteria interact via consumption. We set the exact Monte Carlo simulation algorithm of this model and its mathematical representation as a stochastic process. We prove the convergence of this process to the solution of an integro-differential equation (IDE) when the population size tends to infinity. The IDE is discretized with the help of finite differences, with simulation as well as the IBM are developed in `Python` with the help of the Gamma-Team (UMR Mistea) [63].

Finally with O. Ovaskainen (Univ. of Helsinki) we developed an evolution model for the chemostat.

6.2.3. Simulation and analysis of hybrid models and the atto-fox problem

Participants: Fabien Campillo, Claude Lobry, Alain Rapaport.

We proposed a new "hybrid" model for the simulation of biofilm growth in a plug flow bioreactor, that combines information from three scales: a microscopic one for the individual bacteria, a mesoscopic or "coarse-grained" one that homogenizes at an intermediate scale the quantities relevant to the attachment/detachment process, and a macroscopic one in terms of substrate concentration. In contrast to existing partial differential equations models, this approach is based on a description of biological mechanisms at the individual scale, thus bringing in a biological justification of the attachment/detachment process responsible for the macroscopic behavior [20].

We pursue our study of the "atto-fox" question in the classical Rosenzweig-MacArthur model for a resource-consumer relationship: for certain values of parameters the system has a limit cycle such that the smallest value reached by the resource on this cycle is so small that the model validity is jeopardized [65].

6.3. Identification and control

6.3.1. Reconstruction methods of kinetics functions

Participant: Alain Rapaport.

A collaboration with Sisyphe Inria project-team has led to the development of a new identification method of the kinetics function in the chemostat model, without any a priori on the monotonicity of the function (thus allowing the consideration of bio-processes that are unstable in open loop) [29]. An extension of this method, that is based on singular perturbations, has been proposed for the extremum seeking problem with only two times scale (instead of three for the usual extremum seeking techniques [75]) [50].

6.3.2. Parameter estimation and particle filtering

Participants: Amine Boutoub, Fabien Campillo, Jérôme Harmand, Marc Joannides.

We consider a stochastic logistic growth model involving both birth and death rates in the drift and diffusion coefficients for which extinction eventually occurs almost surely. We then use the numerical integration of the Fokker-Planck equation presented in Section 6.2.1 to build a likelihood function for the unknown model parameters, when discretely sampled data is available. The existing estimation methods need adaptation in order to deal with the extinction problem. We propose such adaptations, based on the particular form of the Fokker-Planck equation, and we evaluate their performances with numerical simulations [64].

We develop particle approximation methods for the nonlinear filtering and parameter estimation with the help of the chemostat model [70].

6.3.3. Functional assignments methods

Participants: Jérôme Harmand, Alain Rapaport.

Following the philosophy of the work that was achieved within the framework of the former PhD thesis of M. Dumont [3], we have applied part of the proposed methodology for a better understanding of the dynamics of specific species of the anaerobic digestion [30], with Chilean collaborators (see Sections 7.4.1.1 and 7.4.2).

Using a combinatorial approach, we have also developed together with UMR Eco & Sols (Cirad, Inra, IRD, SupAgro – Montpellier) a new method to study the role of the interactions within bacterial species on the performance of an ecosystem. More precisely, based on the specific characteristics of the species of a community and the way they interact between each other, we propose a method to predict the behavior of the ecosystem with respect to its biodiversity [34], [25].

6.3.4. Stabilizing strategies for bioprocesses

Participants: Céline Casenave, Jérôme Harmand, Guilherme Pimentel, Alain Rapaport.

We have carrying on developments of stabilizing strategies for bio-processes, with specific characteristics:

- In [48], it has been shown how the buffered configuration of two chemostat models studied in Section 6.1.3 provides an efficient way to stabilize bioprocesses with inhibition. In the same spirit, its has been shown how the consideration of a "passive" buffer (i.e. without biological reaction) can play the role of a delay and enlarge the attraction basin of stable equilibria [49].
- If often happens in bio-processes that measurements are delayed. In [46], a new stabilizing strategy have been proposed to robustly cope with such delays for single chemostats with inhibition.
- For the stabilization of a series of reactors with multiple inputs, a control strategy based on a linearizing control law coupled with a state observer and an anti windup component has been proposed [66], [67], in view of its application in wine fermentation processes. The originality and difficulty of this multi-inputs problem are due to the inputs constraint that imposes that the manipulated dilution rate of each tank has to be less or equal than the one of the previous tank.

6.3.5. Optimal syntheses for bioprocesses control

Participants: Térence Bayen, Amel Ghouali, Jérôme Harmand, Claude Lobry, Alain Rapaport, Tewfik Sari.

We have continued our activities related to the development of optimal control laws for the optimization of bio-processes, notably in taking advantage of the presence of T. Bayen in the team in 2013. Three kinds of results, depending on the kind of processes under interest, were obtained.

a. Control of batch processes. Sequencing Batch bioReactors can be used to efficiently treat water containing both carbonaceous and nitrogenous pollutants. In such a case, an efficient control that can be used is the oxygen concentration. In such systems, oxic and anoxic bacterial are in competition for certain substances. For a simplified version of this complex situation, we have investigated the optimal strategies in order to minimize the energy to be introduced into the system under performance constraints. The originality of the approach lies in the fact that the original problem is transformed into a very general form. Thus, the optimal control problem is formulated and solved for a very general class of systems of ecological relevance [16].

b. Control of fed-batch processes. References [13], [36], [60] are devoted to the study of a bioreactor which is operated in fed-batch mode. We aim at finding an optimal control in feedback form (i.e. depending of the state) that steers the system in a minimal amount of time to a target (which typically has several interests in wastewater treatment). Finding an optimal control in feedback form is crucial from a practical point of view. In [13], previous works on the subject are extended to the case where the growth function depends on an additional product of the reaction. In the references [36], [60], we provide an optimal control in feedback form whenever mortality and recycling rates are considered, and in the case where the maximum dilution rate is not large enough to compete the growth of the species (in the latter case, this implies that the singular arc is non-necessary controllable implying difficulties in determining optimal controls). References [58], [61] are devoted to the study of optimal control problem governed by a chemostat-type model. In [58], an optimal feedback control law is provided in order to optimize the selection of a species in a chemostat model with one limiting substrate and two species. This brings an interesting issue in order to extend this result to the case where the number of species is larger than 3.

c. Coupled dynamics. References [59], [60] give the results of the study of an optimal control problem of a system coupling a culture of micro-algae limited by light and an anaerobic digester. The mathematical model for the dynamics of the reactors takes into account a periodic day-night model of the light in the culture of micro-algae and a chemostat model for the digester. Our aim is to optimize the production of methane in the digester during a certain number of days with respect to the dilution rate. In [59], some preliminary results on this problem are given for an optimal control problem governed by a one-dimensional Kolmogorov equation. In [60], the full system is analyzed by combining direct methods and indirect methods based on Pontryagin's Principle. In [62], we provide a complete characterization of optimal controls for a minimal time control problem where the system describes a two tanks gradostat model under a cascade inputs constraint. This model allows to create a gradient of resources that is expected to be more realistic to mimic real environment for studying micro-organisms growth.

6.4. Distributed delay systems

Participant: Céline Casenave.

In microbial ecosystems, time delays are often present. For a long time (especially with V. Volterra), distributed delay models (or integro-differential equations) have been proposed to take into account these delays in population models. Some dynamic problems dealing with integro-differential models can be tackled in an original way by using the methodology called "diffusive representation". Some works, which began during the PhD thesis of Céline Casenave, are still under development.

In [26], a new formulation of an integro-differential model of a porous media is proposed, based on this methodology. From this formulation, the dissipative and passive features of the porous wall are established, and numerical simulations are performed. A reduced order model is also proposed which summarizes the boundary behavior of the porous wall ([5]).

This work is done in collaboration with LAAS (Univ. Toulouse III) and the Gipsa-lab (CNRS, Grenoble-INP, Univ. Joseph Fourier, Univ. Stendhal). In the future, these works could be adapted to the case of microbial ecosystems.

6.5. Applications to wastewater treatment

Participants: Térence Bayen, Fabien Campillo, Radhouane Fekih-Salem, Amel Ghouali, Jérôme Harmand, Claude Lobry, Alain Rapaport, Tewfik Sari.

If an important part of our work has been done with the final objective of confronting models to data, the studies realized this year are rather theoretical (cf. research achieved within the framework of PhD theses by B. Benyahia, R. Fekih Salem, S. Hassam and G. Araujo Pimentel). In fact, they can be considered as prerequisites before being applied to real systems which, as for most Anaerobic MBRs, are still often found only at pilot scale and not yet applied on real sites.

Concerning the study of membrane fouling, we collaborate with our colleagues of the University Montpellier 2 within the framework of A. Charfi to characterize membrane fouling [43], [54], [42].

In association with the "Laboratoire d'Automatique de Tlemcen" (Univ. Aboubekr Belkaid) and the Gamma Team (UMR Mistea), Modemic launch the NuWat project (Numerics for water treatment research) in the Lirima network (see Section 7.4.2). The first visit of colleagues from the LAT allowed to make choices on the establishment of teachings for the Master in Tlemcen of general trainings and to define research priorities.

The collaboration with Moise Inria project-team has led to a patent application about an algorithm for "intelligent" pumps for the efficient treatment of large water resources [69]. The method relies on an extension of a former work [4] coupled with faithful simulations of the hydrodynamics of the resource and the pollutant dispersion. This typically applied for Chilean lakes, an application that we plan to launch within a common project with CIRIC Inria-Chile.

6.6. Applications to environmental microbiology

Participants: Céline Casenave, Jérôme Harmand, Alain Rapaport.

We have several ongoing works on the modeling of soil microbial ecosystems. The main characteristics of the models we develop with these partners, compared to aquatic microbial ecosystems, concern the availability of the resources, in terms of:

- spatial distribution and transfer of resources, using simple space representations, with the UMR Eco & Sols (Cirad, Inra, IRD, SupAgro – Montpellier) and the UMR Géosciences (Rennes). See Section 6.1.3 and references [28], [19];
- consideration of recalcitrant forms and recycling of nutrients, with UMR Eco & Sols (Cirad, Inra, IRD, SupAgro – Montpellier) and Inra UREP (Clermont). See Section 6.1.3 and [27], [71].

We have also proposed, together with the UMR Eco & Sols (Cirad, Inra, IRD, SupAgro – Montpellier), a new methodology to deduce from the observation of the performances of several assembling of reconstituted ecosystems, the number and the nature of species interactions (see Section 6.3.3 and [34], [25]).

The organization of a "research school" dedicated to the biologists of the marine research station of Banyuls (see Section 8.1) has led to a primary work about variable yield in marine microbial populations.

6.7. Applications to wine fermentation

Participants: Térence Bayen, Céline Casenave, Jérôme Harmand, Alain Rapaport.

[5] Two journal articles dealing with the identification of integro-differential models, and the controllability of some SISO Volterra models are still under revision.

We study the problem of the control of a Multi-State Continuous Fermentor (MSCF) composed of 4 reactors connected in series (the experimental pilot plant is located at Montpellier, in the UMR SPO (INRA, Montpellier SupAgro, Univ. Montpellier 1)). The goal is to control the sugar concentration of the four reactors with, as control inputs, the input flow rates of the four reactors. The originality of the problem comes from the cascade structure of the device which leads to a constraint on the control inputs. Two control strategies have been studied.

An output stabilizing control strategy. The linearizing control law proposed in [66], [67] (see Section 6.3.4) has been validated on numerical simulations, and then has been implemented (Labview-Matlab interface) on the experimental process. The obtained results are convincing; others experiments are scheduled in 2014 to refine the control law.

A minimal time state feedback strategy. The optimal state feedback studied in [62] (see Section 6.3.5) is of completely different nature, as it relies on bang-bang controls and singular arcs. We plan to couple this control law when far from the target with the previous stabilizing law when close from the target, in order to provide a practical sub-optimal strategy.

The first part of this work was conducted as a part of the European CAFE project (Computer-Aided Food processes for control Engineering) described in Section 7.3.1 , in collaboration with CESAME (Univ. Catholique de Louvain-la-neuve), and UMR SPO.

A new project, see Section 7.2 , in which the UMR SPO and the Unit Mistea are involved has begun in 2013. Preliminary work has been performed during the MSc thesis of S. Sekkat [74] about the modeling of the fermentation with addition of nitrogen in the MSCF.

6.8. Applications to micro-algae

Participants: Térence Bayen, Matthieu Sebbah.

An originality developed within the Biocore Inria project-team is to couple a bioreactor that cultivate micro-algae with an anaerobic digester, that uses micro-algae as a substrate that is then converted into valuable bio-gaz (methane). As micro-algae are micro-organisms whose growth is limited by light, one has to take into account periodic day-night model of the light. In [59], [61], control laws that maximize the biogaz production in this periodic framework have been proposed (see Section 6.3.5). In the framework of the Inria Project Lab “Algae in Silico” (see Section 7.2) and the Inria-CIRIC Center in Chile (see Section 7.4), several extensions and collaborations with Biocore Inris project-team are scheduled for the coming year.

6.9. Other results

This section contains some theoretical as well as applied results, that are not directly connected to the main field of the team.

Theoretical ecology
Participant: Tewfik Sari.

In [24], ecological trade-offs between species are studied to explain species coexistence in ecological communities. In our model, plant species compete for sites where each site has a fixed stress condition. Species differ both in stress tolerance and competitive ability. We derive the deterministic discrete-time dynamical system for the species abundances. We prove the conditions under which plant species can coexist in a stable equilibrium. We compare our model with a recently proposed, continuous-time dynamical system for a tolerance-fecundity trade-off in plant communities, and we show that this model is a special case of the continuous-time version of our model.

Calculus of variations
Participant: Térence Bayen.

The work [37] is devoted to the study of necessary and sufficient optimality conditions for weak and strong minima for optimal control problems governed by semi-linear parabolic equations; whereas in the field of calculus of variation, these conditions (such as Euler-Lagrange equation, Legendre's condition, Weierstrass's condition) have been deeply investigated, the study of strong solutions for optimal control problems of partial differential equations is new.

Ice cream crystallization
Participant: Céline Casenave.

We study the problem of the control of an ice cream crystallization process (the experimental pilot plant is located at Irstea Antony). The goal is to control the viscosity of the ice cream at the outlet of the continuous crystallizer. The problem has been studied in two steps.

Modeling, model reduction and parameter identification. On the basis of a population balance equation describing the evolution of the crystal size distribution (CSD) of the ice cream, and an energy balance equation, we have proposed an input-output reduced order model of the process, which is based on physical assumptions. The parameters of the model have been identified and the model has been validated from experimental data [68].

Design of the control law. Based on the reduced order model, a nonlinear control strategy based on an adaptive linearizing control law coupled with a Smith predictor to account for the measurement delay has been proposed [41], [66], [67]. The control has been implemented (Labview-Matlab interface) and then validated on the experimental pilot plant. During the industrial conference which has been organized in February 2013 by the European CAFE project (see Section 7.3.1), and which representatives of several industries in food processing attended, a live demonstration of the designed control law has been performed.

This work was conducted as a part of the European CAFE project, described in Section 7.3.1 , in collaboration with CESAME (Univ. Catholique de Louvain-la-neuve), Irstea Antony and AgroParisTech.

Semi-Markov land use dynamic
Participant: Fabien Campillo, Angelo Raherinirina.

We pursued the development of semi-Markov model for the inference of land use dynamic from data proposed by IRD. The thesis of A. Raherinirina was defended in August 2013 [12]. Later in the year during the stay of A. Raherinirina in Montpellier we completed an article accepted by the journal ARIMA and that will be published in 2014 [17].

MOISE Project-Team

6. New Results

6.1. Mathematical Modelling of the Ocean Dynamics

6.1.1. Coupling Methods for Oceanic and Atmospheric Models

Participants: Eric Blayo, Mehdi-Pierre Daou, Laurent Debreu, Florian Lemarié, Antoine Rousseau, Manel Tayachi.

6.1.1.1. Interface conditions for coupling ocean models

Many physical situations require coupling two models with not only different resolutions, but also different physics. Such a coupling can be studied within the framework of global-in-time Schwarz methods. However, the efficiency of these iterative algorithms is strongly dependent on interface conditions. As a first step towards coupling a regional scale primitive equations ocean model with a local Navier-Stokes model, a study on the derivation of interface conditions for 2-D $x-z$ Navier-Stokes equations has been performed in D. Cherel PhD thesis. It has been shown theoretically that several usual conditions lead to divergent algorithms, and that a convergent algorithm is obtained when using transmission conditions given by a variational calculation.

D. Cherel has implemented a Schwarz-based domain decomposition method, for which he developed optimized absorbing boundary conditions that mix the velocity and pressure variables on an Arakawa-C grid. The numerical results confirm the rate of convergence that has been obtained theoretically, thanks to a Fourier analysis of the semi-discretized problem. New optimized conditions offer much better convergence than classical Dirichlet-Dirichlet conditions, even when domains overlap. A paper is now ready for submission.

6.1.1.2. Coupling dimensionally heterogeneous models

The coupling of different types of models is gaining more and more attention recently. This is due, in particular, to the needs of more global models encompassing different disciplines (*e.g.* multi-physics) and different approaches (*e.g.* multi-scale, nesting). Also, the possibility to assemble different modeling units inside a friendly modelling software platform is an attractive solution compared to developing more and more global complex models. More specifically one may want to couple 1D to 2D or 3D models, such as Shallow Water and Navier Stokes models: this was the framework of our partnership with EDF in the project MECSICO, now extended with ARTELIA Group. In her PhD (defended Oct. 28th, 2013) [4], M. Tayachi built a theoretical and numerical framework to couple 1D, 2D and 3D models for river flows.

In [103] (now accepted for publication), we propose and analyze an efficient iterative coupling method for a dimensionally heterogeneous problem. We consider the case of a 2-D Laplace equation with non symmetric boundary conditions with a corresponding 1-D Laplace equation. We first show how to obtain the 1-D model from the 2-D one by integration along one direction, by analogy with the link between shallow water equations and the Navier-Stokes system. Then we focus on the design of a Schwarz-like iterative coupling method. We discuss the choice of boundary conditions at coupling interfaces. We prove the convergence of such algorithms and give some theoretical results related to the choice of the location of the coupling interface, and to the control of the difference between a global 2-D reference solution and the 2-D coupled one. These theoretical results are illustrated numerically. The extension of this work to shallow water equations has been started in 2013 with the PhD thesis of Medhi Pierre Daou (funded by ARTELIA). An extension to primitive equations is envisaged: a post-doc position has been proposed in 2013 (not funded) and will renewed in 2014.

6.1.1.3. Ocean-Atmosphere coupling

Coupling methods routinely used in regional and global climate models do not provide the exact solution to the ocean-atmosphere problem, but an approached one [75]. This finding has motivated a deep numerical analysis of multi-physics coupling problems, first on simplified academic cases based on diffusion equations. In this context, Schwarz-like iterative domain decomposition methods have been analyzed and efficient interface conditions have been determined to optimize the convergence rate of the method [19], [20], [53], [79]. This method has then been applied to the coupling of realistic oceanic and atmospheric models to simulate the propagation of a tropical cyclone (cyclone erica, Fig. 1). Sensitivity tests to the coupling method have been been carried out in an ensemblist approach. We showed that with a mathematically consistent coupling, compared to coupling methods en vogue in existing coupled models, the spread of the ensemble is reduced, thus indicating a much reduced uncertainty in the physical solution [24], [75].

The next step is now to complete the theoretical work done on a diffusion problem by including the formulation of physical parameterizations to tackle a problem more representative of the realistic models: a PhD thesis should start on this subject in fall 2014. In parallel, an important perspective is to assess the impact of our work on IPCC-like climate models, this task will be initiated in 2014 through a collaboration between the MOISE project-team and the LSCE (Laboratoire des Sciences du Climat et de l'Environnement).

In collaboration with geophysicists, number of studies to investigate important small-scales air-sea feedbacks are in progress [45]. Through those studies, the aim will be to mathematically derive a metamodel able to represent important processes in the marine atmospheric boundary layers. The medium term objective will be to use this metamodel to force high-resolution oceanic operational models for which the use of a full atmospheric model is not possible due to a prohibitive computational cost.

6.1.2. Numerical schemes for ocean modelling

Participants: Laurent Debreu, Jérémie Demange, Florian Lemarié.

Reducing the traditional errors in terrain-following vertical coordinate ocean models (or sigma models) has been a focus of interest for the last two decades. The objective is to use this class of model in regional domains which include not only the continental shelf, but the slope and deep ocean as well. Two general types of error have been identified: 1) the pressure-gradient error and 2) spurious diapycnal diffusion associated with steepness of the vertical coordinate. In ([92],[46]) we have studied the problem of diapycnal mixing. The solution to this problem requires a specifically designed advection scheme. We propose and validate a new scheme, where diffusion is split from advection and is represented by a rotated biharmonic diffusion scheme with flow-dependent hyperdiffusivity. The main numerical development was to render the biharmonic diffusion operator scheme unconditionally stable. This is particularly needed when the slopes between coordinates lines and isopycnal surfaces are important so that the rotation of the biharmonic leads to strong stability condition along the vertical coordinate where the grid size is relatively small. This work also extends more classical results on the stability of Laplacian diffusion with mixed derivatives.

In his PhD, Jérémie Demange works on advection-diffusion schemes for ocean models (Supervisors : L. Debreu, P. Marchesiello (IRD)). His work focuses on the link between tracers (temperature and salinity) and momentum advection and diffusion in the non hyperbolic system of equations typically used in ocean models (the so called primitive equations with hydrostatic and Boussinesq assumptions). We also investigated the use of a depth dependent barotropic mode in free surface ocean models. When most ocean models assume that this mode is vertically constant, we have shown that the use of the true barotropic mode, derived from a normal mode decomposition, allows more stability and accuracy in the representation of external gravity waves.

Salinity at 1000 m in the Southwest Pacific ocean is shown in figure 2 . The use of traditional upwind biased schemes (middle) exhibits a strong drift in the salinity field in comparison with climatology (left). The introduction of high order diffusion rotated along geopotential surfaces prevents this drift while maintaining high resolution features (right).

../../../../projets/moise/IMG/wrf_roms_domain_reduced2.png

Figure 1. Snapshots (March 12, 2003 at 8 p.m. GMT) of (a) oceanic sea surface temperature (b) atmospheric 10 meter winds during a coupled simulation.

../../../../projets/moise/IMG/diffusion.png

Figure 2. Salinity at 1000m in the Southwest Pacific ocean.

6.2. Data Assimilation for Geophysical Models

6.2.1. Development of a Variational Data Assimilation System for OPA9/NEMO

Participants: Arthur Vidard, Bénédicte Lemieux-Dudon, Pierre-Antoine Bouttier, Eric Blayo.

We are heavily involved in the development of NEMOVAR (Variational assimilation for NEMO). For several years now, we built a working group (coordinated by A. Vidard) in order to bring together various NEMOVAR user-groups with diverse scientific interests (ranging from singular vector and sensitivity studies to specific issues in variational assimilation). It has led to the creation of the VODA (Variational Ocean Data Assimilation for multi scales applications) ANR project (ended in 2012). A new project, part of a larger EU-FP7 project (ERA-CLIM2) has been submitted late 2012 and will start early 2014.
The project aims at delivering a common NEMOVAR platform based on NEMO platform for 3D and 4D variational assimilation. Following 2009-11 VODA activities, a fully parallel version of NEMOTAM (Tangent and Adjoint Model for NEMO) is now available for the community in the standard NEMO version. This version is based on the released 3.4.1 version of NEMO.

We are also investigating variational data assimilation methods applied to high resolution ocean numerical models (see figure 3). This part of the project is now well advanced and encouraging preliminary results are available on an idealised numerical configuration of an oceanic basin. Several novative diagnostics have been also developed in this framework as part of P.A. Bouttier's PhD that will be defended early 2014

Lastly, multi resolution algorithms have been developed to solve the variational problem, and preliminary results were presented in two international communications [52], [51].

6.2.2. Ensemble Kalman filtering for large scale ice-sheet models

Participants: Bertrand Bonan, Maëlle Nodet, Catherine Ritz.

In collaboration with C. Ritz (CNRS, Laboratoire de Glaciologie et Geophysique de l'Environnement (LGGE), Grenoble), we aim to develop inverse methods for ice cap models.

In the framework of global warming, the evolution of sea level is a major but ill-known phenomenon. It is difficult to validate the models which are used to predict the sea level elevation, because observations are heterogeneous and sparse.

Data acquisition in polar glaciology is difficult and expensive. Satellite data have a good spatial coverage, but they allow only indirect observation of the interesting data. Moreover, ice dynamics processes are highly non linear and involve many feedback loops, so that classical linear data assimilation give poor results.

B. Bonan defended his PhD [1] in November 2013 on this subject. We implemented the Ensemble Transform Kalman Filter (ETKF) algorithm for a flowline Shallow-Ice model, called Winnie, developed by C. Ritz at LGGE. On twin experiments we got interesting results. Figures 4 show the reconstruction of the bedrock topography for various ensemble sizes. We can see that the obtained bedrock is very close to the true one, even for small ensemble sizes. This is very promising for the future, as we want to implement this method into a full 3D model. A journal paper has been submitted on this subject, and the results have been presented at many conferences [27], [37], [48], [38], [39].

6.2.3. Inverse methods for full-Stokes glaciology models

Participants: Olivier Gagliardini, Maëlle Nodet, Catherine Ritz.

We are investigating the means to apply inverse modeling to another class of glaciology models, called full-Stokes model. Such a model is developed by LGGE and CSC in Finland, called Elmer/Ice. Contrary to large scale models, Elmer/Ice is based on the full Stokes equations, and no assumptions regarding aspect ratio are made, so that this model is well adapted to high resolution small scale modelling, such as glaciers (and more recently the whole Greenland ice-sheet).

../../../../projets/moise/IMG/seabass.jpg

Figure 3. Snapshot of the relative vorticity field (1/s) for an academic oceanic basin model at 1/100° horizontal resolution.

Figure 4.

Left: Bedrock topography after 20 years of the LETKF with inflation. The background (green) is compared to reference (blue) and the analyses for various ensemble sizes: 100 members (purple), 50 members (cyan) and 30 members (orange).
Right: Standard deviation of the errors (compared to the reference) for the bedrock topography after 20 years of the LETKF with inflation. The background (green) is compared to the analyses for various ensemble sizes: 100 members (purple), 50 members (cyan) and 30 members (orange).

In collaboration with O. Gagliardini, F. Gillet-Chaulet and C. Ritz (Laboratoire de Glaciologie et Géophysique de l'Environnement (LGGE), Grenoble), we investigated a new method to solve inverse problems for a Full-Stokes model of Groenland, which consisted in solving iteratively a sequence of Neumann and Dirichlet problems within a gradient descent algorithm. We also compared this method to an approximate variational algorithm, using the fact that the full Stokes equations are almost self-adjoint. These results have been presented at an international conference [44].

6.3. Development of New Methods for Data Assimilation

6.3.1. Variational Data Assimilation with Control of Model Error

Participants: Bénédicte Lemieux-Dudon, Arthur Vidard.

One of the main limitations of the current operational variational data assimilation techniques is that they assume the model to be perfect, mainly because of computing cost issues. Numerous researches have been carried out to reduce the cost of controlling model errors by controlling the correction term only in certain privileged directions or by controlling only the systematic and time correlated part of the error.

Both the above methods consider the model errors as a forcing term in the model equations. Trémolet (2006) describes another approach where the full state vector (4D field: 3D spatial + time) is controlled. Because of computing cost one cannot obviously control the model state at each time step. Therefore, the assimilation window is split into sub-windows, and only the initial conditions of each sub-window are controlled, the junctions between each sub-window being penalized. One interesting property is that, in this case, the computation of the gradients, for the different sub-windows, are independent and therefore can be done in parallel.

This method is now implemented in a realistic oceanic framework using OPAVAR/ NEMOVAR. An extensive documentation has been produced and we are now assessing the improvement over the previous scheme

6.3.2. Direct assimilation of sequences of images

Participants: François-Xavier Le dimet, Maëlle Nodet, Nicolas Papadakis, Arthur Vidard, Vincent Chabot.

At the present time the observation of Earth from space is done by more than thirty satellites. These platforms provide two kinds of observational information:

- Eulerian information as radiance measurements: the radiative properties of the earth and its fluid envelops. These data can be plugged into numerical models by solving some inverse problems.
- Lagrangian information: the movement of fronts and vortices give information on the dynamics of the fluid. Presently this information is scarcely used in meteorology by following small cumulus clouds and using them as Lagrangian tracers, but the selection of these clouds must be done by hand and the altitude of the selected clouds must be known. This is done by using the temperature of the top of the cloud.

MOISE was the leader of the ANR ADDISA project dedicated to the assimilation of images, and is a member of its follow-up GeoFluids (along with EPI FLUMINANCE and CLIME, and LMD, IFREMER and Météo-France) that just ended in 2013.

During the ADDISA project we developed Direct Image Sequences Assimilation (DISA) and proposed a new scheme for the regularization of optical flow problems [101]. Thanks to the nonlinear brightness assumption, we proposed an algorithm to estimate the motion between two images, based on the minimization of a nonlinear cost function. We proved its efficiency and robustness on simulated and experimental geophysical flows [77]. As part of the ANR project GeoFluids, we are investigating new ways to define distance between a couple of images. One idea is to compare the gradient of the images rather than the actual value of the pixels. This leads to promising results. Another idea, currently under investigation, consists in comparing main structures within each image. This can be done using, for example, a wavelet representation of images. Both approaches have been compared, in particular their relative merits in dealing with observation errors, in a submitted paper late 2013 [63] and presented in a national conference [34]

We are also part of TOMMI, another ANR project started mid 2011, where we are investigating the possibility to use optimal transportation based distances for images assimilation.

6.3.3. Image processing, Optimal transport

Participants: Maëlle Nodet, Nicolas Papadakis, Arthur Vidard, Nelson Feyeux.

Within the optimal transport project TOMMI funded by the ANR white program, a new optimization scheme based on proximal splitting method has been proposed to solve the dynamic optimal transport problem. This work allows the computation of generalized optimal transports and will be published in SIAM Journal on Imaging Sciences [96]. We investigate also the use of optimal transport based distances for data assimilation. N. Feyeux just started his PhD on this subject, and his PhD project has been presented in a regional workshop [49].

6.3.4. A Nudging-Based Data Assimilation Method: the Back and Forth Nudging

Participants: Maëlle Nodet, Jacques Blum, Didier Auroux.

The Back and Forth Nudging (BFN) algorithm has been recently introduced for simplicity reasons, as it does not require any linearization, nor adjoint equation, or minimization process in comparison with variational schemes. Nevertheless it provides a new estimation of the initial condition at each iteration.

Previous theoretical results showed that BFN was often ill-posed for viscous partial differential equations. To overcome this problem, we proposed a new version of the algorithm, which we called the Diffusive BFN, and which showed very promising results on one-dimensional viscous equations. Experiments on more sophisticated geophysical models, such as Shallow-Water equations and NEMO ocean model are still in progress, in collaboration with University of Nice, and have been presented at the MAMERN conference [30].

6.3.5. Multigrid methods for Variational Data Assimilation.

Participants: Laurent Debreu, François-Xavier Le Dimet, Arthur Vidard.

In order to lower the computational cost of the variational data assimilation process, we investigate the use of multigrid methods to solve the associated optimal control system. On a linear advection equation, we study the impact of the regularization term on the optimal control and the impact of discretization errors on the efficiency of the coarse grid correction step. We show that even if the optimal control problem leads to the solution of an elliptic system, numerical errors introduced by the discretization can alter the success of the multigrid methods. The view of the multigrid iteration as a preconditioner for a Krylov optimization method leads to a more robust algorithm. A scale dependent weighting of the multigrid preconditioner and the usual background error covariance matrix based preconditioner is proposed and brings significant improvements. This work is presented in a paper submitted to QJRMS ([68]).

6.3.6. Variational Data Assimilation and Control of Boundary Conditions

Participant: Eugène Kazantsev.

A variational data assimilation technique is applied to the identification of the optimal boundary conditions for two configurations of the NEMO model.

The first one is a full-physics low-resolution configuration, known as ORCA-2 model. In this experiment we identify optimal parametrizations of boundary conditions on the lateral boundaries as well as on the bottom and on the surface of the ocean [17]. The influence of boundary conditions on the solution is analyzed as in the assimilation window and beyond the window. It is shown that the influence of the lateral boundaries is not significant in this configuration, while optimal surface and bottom boundary conditions allow us to better represent the jet streams, such as Gulf Stream and Kuroshio. Analyzing the reasons of the jets reinforcement, we notice that data assimilation has a major impact on parametrization of the bottom boundary conditions for u and v [23].

The second configuration of the Nemo model is devoted to the identification of the optimal parametrization of lateral boundary conditions. The model in a rectangular box placed in mid-latitudes and subjected to the classical single or double gyre wind forcing is studied. The model grid can be rotated on a desired angle around the center of the rectangle in order to simulate the boundary approximated by a staircase-like coastlines. The solution of the model on the grid aligned with the box borders was used as a reference solution and as artificial observational data. It is shown that optimal boundary has a rather complicated geometry which is neither a staircase, nor a strait line. The boundary conditions found in the data assimilation procedure brings the solution toward the reference solution allowing to correct the influence of the rotated grid (see fig. 5).

Adjoint models, necessary to variational data assimilation, have been produced by the TAPENADE software, developed by the TROPICS team. This software is shown to be able to produce the adjoint code, that can be used in data assimilation after a memory usage optimization.

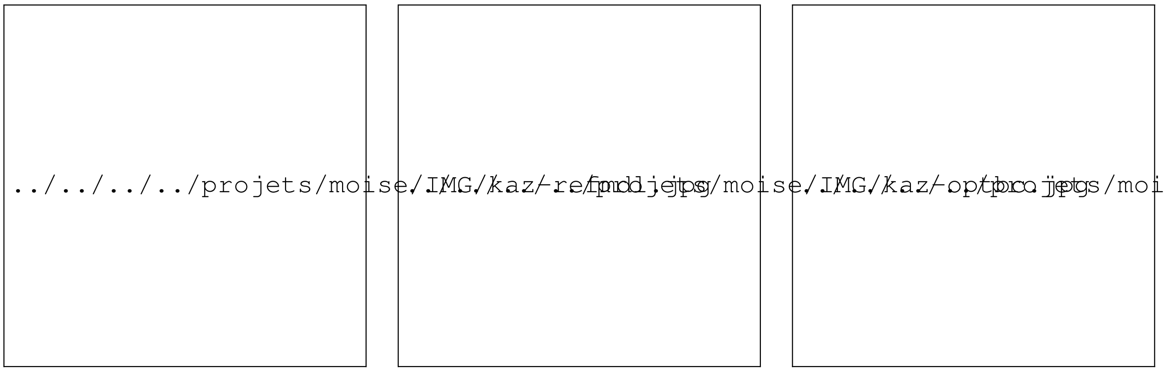


Figure 5. Sea surface elevation: reference solution on the aligned grid (left), solutions on the $30°$ *rotated grid with optimal (center) and classical (right) boundary conditions.*

6.4. Quantifying Uncertainty

6.4.1. Sensitivity analysis for forecasting ocean models

Participants: Anestis Antoniadis, Eric Blayo, Gaëlle Chastaing, Céline Helbert, Alexandre Janon, François-Xavier Le Dimet, Simon Nanty, Maëlle Nodet, Clémentine Prieur, Federico Zertuche, Simon Nanty, Laurent Gilquin.

6.4.1.1. Scientific context

Forecasting ocean systems require complex models, which sometimes need to be coupled, and which make use of data assimilation. The objective of this project is, for a given output of such a system, to identify the most influential parameters, and to evaluate the effect of uncertainty in input parameters on model output. Existing stochastic tools are not well suited for high dimension problems (in particular time-dependent problems), while deterministic tools are fully applicable but only provide limited information. So the challenge is to gather expertise on one hand on numerical approximation and control of Partial Differential Equations, and on the other hand on stochastic methods for sensitivity analysis, in order to develop and design innovative stochastic solutions to study high dimension models and to propose new hybrid approaches combining the stochastic and deterministic methods.

6.4.1.2. Estimating sensitivity indices

A first task is to develop tools for estimated sensitivity indices. In variance-based sensitivity analysis, a classical tool is the method of Sobol' [100] which allows to compute Sobol' indices using Monte Carlo integration. One of the main drawbacks of this approach is that the estimation of Sobol' indices requires the use of several samples. For example, in a d-dimensional space, the estimation of all the first-order Sobol' indices requires $d+1$ samples. Some interesting combinatorial results have been introduced to weaken this defect, in particular by Saltelli [98] and more recently by Owen [95] but the quantities they estimate still require $O(d)$ samples. In a recent work [104] we introduce a new approach to estimate for any k all the k-th order Sobol' indices by using only two samples based on replicated latin hypercubes. We establish theoretical properties of such a method for the first-order Sobol' indices and discuss the generalization to higher-order indices. As an illustration, we propose to apply this new approach to a marine ecosystem model of the Ligurian sea (northwestern Mediterranean) in order to study the relative importance of its several parameters. The calibration process of this kind of chemical simulators is well-known to be quite intricate, and a rigorous and robust — i.e. valid without strong regularity assumptions — sensitivity analysis, as the method of Sobol' provides, could be of great help. The computations are performed by using CIGRI, the middleware used on the grid of the Grenoble University High Performance Computing (HPC) center. We are also applying these estimates to calibrate integrated land use transport models. It is the first step in the PhD of Laurent Gilquin (started in October 2013). Laurent Gilquin is supervised by Clémentine Prieur and Elise Arnaud (EPI STEEP) and his PhD is funded by the ANR project CITIES.

We can now wonder what are the asymptotic properties of these new estimators, or also of more classical ones. In [67], the authors deal with asymptotic properties of the estimators. In [70], the authors establish also a multivariate central limit theorem and non asymptotic properties.

6.4.1.3. Intrusive sensitivity analysis, reduced models

Another point developed in the team for sensitivity analysis is model reduction. To be more precise regarding model reduction, the aim is to reduce the number of unknown variables (to be computed by the model), using a well chosen basis. Instead of discretizing the model over a huge grid (with millions of points), the state vector of the model is projected on the subspace spanned by this basis (of a far lesser dimension). The choice of the basis is of course crucial and implies the success or failure of the reduced model. Various model reduction methods offer various choices of basis functions. A well-known method is called “proper orthogonal decomposition" or “principal component analysis". More recent and sophisticated methods also exist and may be studied, depending on the needs raised by the theoretical study. Model reduction is a natural way to overcome difficulties due to huge computational times due to discretizations on fine grids. In [10], the authors present a reduced basis offline/online procedure for viscous Burgers initial boundary value problem, enabling efficient approximate computation of the solutions of this equation for parametrized viscosity and initial and boundary value data. This procedure comes with a fast-evaluated rigorous error bound certifying the approximation procedure. The numerical experiments in the paper show significant computational savings, as well as efficiency of the error bound. When a metamodel is used (for example reduced basis metamodel, but also kriging, regression, ...) for estimating sensitivity indices by Monte Carlo type estimation, a twofold error appears : a sampling error and a metamodel error. Deriving confidence intervals taking into account these two sources of uncertainties is of great interest. We obtained results particularly well fitted for reduced basis metamodels [90]. We are now considering problems related to more general models such as Shallow-Water models. In [15], the authors provide asymptotic confidence intervals in the double limit where the sample size goes to infinity and the metamodel converges to the true model.

Let us come back to the output of interest. Is it possible to get better error certification when the output is specified. A work in this sense has been submitted, dealing with goal oriented uncertainties assessment [89].

6.4.1.4. Sensitivity analysis with dependent inputs

An important challenge for stochastic sensitivity analysis is to develop methodologies which work for dependent inputs. For the moment, there does not exist conclusive results in that direction. Our aim is to define an analogue of Hoeffding decomposition [88] in the case where input parameters are correlated. Clémentine

Prieur supervised Gaëlle Chastaing's PhD thesis on the topic (defended in September 2013) [2]. We obtained first results [81], deriving a general functional ANOVA for dependent inputs, allowing defining new variance based sensitivity indices for correlated inputs. We then adapted various algorithms for the estimation of these new indices. These algorithms make the assumption that among the potential interactions, only few are significant. Two papers have been submitted [64], [66].

Céline Helbert and Clémentine Prieur supervise the PhD thesis of Simon Nanty (funded by CEA Cadarache). The subject of the thesis is the analysis of uncertainties for numerical codes with temporal and spatio-temporal input variables, with application to safety and impact calculation studies. This study implies functional dependent inputs. A first step is the modeling of these inputs.

6.4.1.5. Multy-fidelity modeling for risk analysis

Federico Zertuche's PhD concerns the modeling and prediction of a digital output from a computer code when multiple levels of fidelity of the code are available. A low-fidelity output can be obtained, for example on a coarse mesh. It is cheaper, but also much less accurate than a high-fidelity output obtained on a fine mesh. In this context, we propose new approaches to relieve some restrictive assumptions of existing methods ([91], [97]) : a new estimating method of the classical cokriging model when designs are not nested and a nonparametric modeling of the relationship between low-fidelity and high-fidelity levels. The PhD takes place in the REDICE consortium and in close link with industry. The first year was also dedicated to the development of a case study in fluid mechanics with CEA in the context of the study of a nuclear reactor.

The second year of the thesis was dedicated to the development of a new sequential approach based on a course to fine wavelets algorithm.

6.4.2. Evaluation of a posteriori covariance errors

In the context of data assimilation, taking into account the a priori covariance error on the prediction and on the observations, the model and the observations, an analysis can be obtained followed by a prediction. This one makes sense only if an estimation of the error can be provided. The tendency is to use "ensemble methods" i.e. to realize a large number of predictions and estimate statistics on the results. This method raises two problems: the high computational cost and the weak theoretical justification. We have proposed a new method based on the fact that in the linear case the covariance is the inverse of the Hessian. The principle of our method is to add a correcting term to the Hessian in the non linear case. This work has been published in 2013 [14]. This paper has also be presented at the 6th WMO Symposium on Data Assimilation held in College Park, MD, USA in October 2013 [73].

6.4.3. Second Order Information in Variational Data Assimilation

This theme is centered around sensitivity analysis with respect to the observations. The link between data and models is made only in the Optimality System. Therefore a sensitivity analysis on the observations must be carried out on the Optimality System thus using second order information. This research is done in cooperation with Victor Shutyaev (Institute of Numerical Mathematics, Moscow), Tran Thu Ha (Institute of Mechanics, Ha Noi, Vietnam). One paper is published in the Russ. J. Of Numerical Analysis [18]. Another application to identification of parameters in a hydrological model is submitted [105].

6.5. Tracking of mesoscale convective systems

Participants: Clémentine Prieur, Alexandros Makris.

6.5.1. Scientific context

We are interested in the tracking of mesoscale convective systems.A particular region of interest is West Africa. Data and hydrological expertise is provided by T. Vischel and T. Lebel (LTHE, Grenoble).

6.5.2. Results

A first approach involves adapting the multiple hypothesis tracking (MHT) model originally designed by the NCAR (National Centre for Atmospheric Research) for tracking storms [102] to the data for West Africa. With A. Makris (working on a post-doctoral position), we proposed a Bayesian approach [76], which consists in considering that the state at time t is composed on one hand by the events (birth, death, splitting, merging) and on the other hand by the targets' attributes (positions, velocities, sizes, ...). The model decomposes the state into two sub-states: the events and the targets positions/attributes. The events are updated first and are conditioned to the previous targets sub-state. Then given the new events the target substate is updated. A simulation study allowed to verify that this approach improves the frequentist approach by Storlie et al. (2009). It has been tested on simulations and must now be investigated in the specific context of real data on West Africa. Using PHD (probability hypothesis density) filters adapted to our problem, generalising recent developments in particle filtering for spatio-temporal branching processes (e.g. [80]) could be an interesting alternative to explore. The idea of a dynamic, stochastic tracking model should then provide the base for generating rainfall scenarios over a relatively vast area of West Africa in order to identify the main sources of variability in the monsoon phenomenon.

6.6. Multivariate risk indicators

Participants: Clémentine Prieur, Patricia Tencaliec.

6.6.1. Scientific context

Studying risks in a spatio-temporal context is a very broad field of research and one that lies at the heart of current concerns at a number of levels (hydrological risk, nuclear risk, financial risk etc.). Stochastic tools for risk analysis must be able to provide a means of determining both the intensity and probability of occurrence of damaging events such as e.g. extreme floods, earthquakes or avalanches. It is important to be able to develop effective methodologies to prevent natural hazards, including e.g. the construction of barrages.

6.6.2. Results

Different risk measures have been proposed in the one-dimensional framework . The most classical ones are the return level (equivalent to the Value at Risk in finance), or the mean excess function (equivalent to the Conditional Tail Expectation CTE). However, most of time there are multiple risk factors, whose dependence structure has to be taken into account when designing suitable risk estimators. Relatively recent regulation (such as Basel II for banks or Solvency II for insurance) has been a strong driver for the development of realistic spatio-temporal dependence models, as well as for the development of multivariate risk measurements that effectively account for these dependencies. We refer to [82] for a review of recent extensions of the notion of return level to the multivariate framework. In the context of environmental risk, [99] proposed a generalization of the concept of return period in dimension greater than or equal to two. Michele et al. proposed in a recent study [83] to take into account the duration and not only the intensity of an event for designing what they call the dynamic return period. However, few studies address the issues of statistical inference in the multivariate context. In [9], [86], we proposed non parametric estimators of a multivariate extension of the CTE. As might be expected, the properties of these estimators deteriorate when considering extreme risk levels. In collaboration with Elena Di Bernardino (CNAM, Paris), Clémentine Prieur is working on the extrapolation of the above results to extreme risk levels.

Elena Di Bernardino, Véronique Maume-Deschamps (Univ. Lyon 1) and Clémentine Prieur also derived an estimator for bivariate tail [10]. The study of tail behavior is of great importance to assess risk.

With Anne-Catherine Favre (LTHE, Grenoble), Clémentine Prieur supervises the PhD thesis of Patricia Tencaliec. We are working on risk assessment, concerning flood data for the Durance drainage basin (France). The PhD thesis started in October.

6.7. Non-parametric estimation for kinetic diffusions

Participant: Clémentine Prieur.

This research is the subject of a collaboration with Venezuela (Professor Jose R. Leon, Caracas Central University) and is partly funded by an ECOS Nord project.

We are focusing our attention on models derived from the linear Fokker-Planck equation. From a probabilistic viewpoint, these models have received particular attention in recent years, since they are a basic example for hypercoercivity. In fact, even though completely degenerated, these models are hypoelliptic and still verify some properties of coercivitiy, in a broad sense of the word. Such models often appear in the fields of mechanics, finance and even biology. For such models we believe it appropriate to build statistical non-parametric estimation tools. Initial results have been obtained for the estimation of invariant density, in conditions guaranteeing its existence and unicity [8] and when only partial observational data are available. A paper on the non parametric estimation of the drift has been submitted recently [62] (see Samson et al., 2012, for results for parametric models). As far as the estimation of the diffusion term is concerned, we obtained promising results, in collaboration with J.R. León (Caracas, Venezuela) and P. Cattiaux (Toulouse). These results should be submitted shortly.

6.8. CO_2 Storage

Participant: Céline Helbert.

In collaboration with Bernard Guy (EMSE, Saint-Etienne) and more specifically in the context the PhD of Joharivola Raveloson (EMSE, Saint-Etienne), we are interested in the study of the water-rock interactions in the case of CO_2 storage in geological environment. This work is following the study of Franck Diedro in the same subject [87]. In this study we focus on the scale of observation of geochemical phenomena while taking into account the heterogeneity of the reservoir. This heterogeneity at small and large scale helps to maintain a local variability of the chemical composition and influence reaction rates at the pore as well at the reservoir scale. To connect the parameters at both scale (pore and reservoir) we use deterministic and stochastic simulations of a reactive transport code developed by IFPEN.

6.9. Land Use and Transport models calibration

Participants: Thomas Capelle, Laurent Gilquin, Clémentine Prieur, Nicolas Papadakis, Arthur Vidard.

Given the complexity of modern urban areas, designing sustainable policies calls for more than sheer expert knowledge. This is especially true of transport or land use policies, because of the strong interplay between the land use and the transportation systems. Land use and transport integrated (LUTI) modelling offers invaluable analysis tools for planners working on transportation and urban projects. Yet, very few local authorities in charge of planning make use of these strategic models. The explanation lies first in the difficulty to calibrate these models, second in the lack of confidence in their results, which itself stems from the absence of any well-defined validation procedure. Our expertise in such matters will probably be valuable for improving the reliability of these models. To that purpose we participated to the building up of the ANR project CITiES lead by the STEEP EPI. This project has just started early 2013 and Two PhD about sensitivity analysis and calibration were launched this fall.

6.10. Mathematical modelling for CFD-environment coupled systems

Participants: Antoine Rousseau, Maëlle Nodet.

6.10.1. Minimal-time bioremediation of natural water resources

The objective of this work is to provide efficient strategies for the bioremediation of natural water resources. The originality of the approach is to couple minimal time strategies that are determined on a simplified model with a faithful numerical model for the hydrodynamics. Based on a previous paper that deals with an implicit representation of the spatial inhomogenity of the resource with a small number of homogeneous compartments (with a system of ODEs), we implement a coupled ODE-PDE system that accounts for the spatial non-homogeneity of pollution in natural resources. The main idea is to implement a Navier-Stokes model in the resource (such as a lake), with boundary conditions that correspond to the output feedback that

has been determined to be optimal for the simple ODEs model of a (small) bioreactor. A first mathematical model has been introduced and numerical simulations have been performed in academic situations. We built a reduced model that approximates the reference PDE model thanks to a set of ODEs with parameters. Numerical optimization is performed on these parameters in order to better fit the reference model. In addition, bioremediation algorithms proposed by the authors have been sent to Inria Technology Transfert Services for a patent registration. Two publications (ready for submission) will be sent as soon as the patent submission process is complete.
Finally, A. Rousseau spent 2 weeks in Santiago (April 2013) upon Inria Chile's invitation in order to work on the bioremediation of natural resources. AR and Inria Chile made a common answer to a chilean funding program (by COPEC) that was not chosen.

6.10.2. Mathematical modelling for the confinement of lagoons

This work deals with the concept of confinement of paralic ecosystems. It is based on a recent paper by E. Frénod that presents a modelling procedure in order to compute the confinement field of a lagoon.
A. Rousseau and E. Frénod improve in 2012 the existing model in order to account for tide oscillations in any kind of geometry such as a non-rectangular lagoons with a non-flat bottom. The new model, that relies on PDEs rather than ODEs, is then implemented thanks to the finite element method. Numerical results confirm the feasibility of confinement studies thanks to the introduced model. During the internship of J.-P. Bernard, we implemented the proposed method in a realistic situation, namely the Etang de Thau in Languedoc-Roussillon, France (see Figure 6). This led to two publications in 2013 [5] and [13], plus one accepted paper in 2014 [6].

../../../../projets/moise/IMG/Thau.png

Figure 6. Confinement map in the Thau Lagoon (France).

MOKAPLAN Exploratory Action

6. New Results

6.1. Monge-Ampère solver for the Mass Transportation problem and extensions

- **Benamou, Froese (Univ. of Texas at Austin) -** We design a scheme for Aleksandrov solution of Optimal Mass Transportation between atomic measure and continuous densities. The idea is to couple the notion of viscosity solution with an adapted sub gradient discretization at dirac points where the notion of Aleksandrov solution is relevant. This would offer a "PDE" alternative to the classical gradient methods based on costly computational geometry tools [61].
- **Benamou, Collino, Mirebeau (Univ. Paris IX,CNRS) -** A new variational formulation of the determinant of a semi-definite positive matrix has been proposed based on the ideas developed in [60]. This leads to a monotone discretisation of the Monge-Ampère operator. A Newton method preserving convexity is currently being tested. The new scheme is more accurate than the wide stencil, currently the state of the art of monotone scheme for the Monge-Ampère equation.
- **Benamou, Froese (Univ. of Texas at Austin), Oberman (Univ. Mc Gill) -** When the Optimal Mass Transportation data is not balanced, i.e. the densities do not have equal mass. A natural extension of the optimal transport has been proposed by McCann and Caffareli [30] and revisited by Figalli [41]. It is formulated as an obstacle problem which automatically select the portion of mass corresponding to Optimal Mass Transportation. The numerical resolution of this problem is open and we believe ideas linked the state constraint reformulation contained in paper [6] may be applied to obtain a tractable reformulation.

6.2. Variational problems under divergence constraint - Alg2

- **Benamou, Bonne, Carlier -** Dynamic problems: we have extended the Augmented Lagrangian method used for the CFD formulation of the Optimal Mass Transportation to Mean Field Games that is for the optimal control of the continuity equation. A freefem Code has been implemented.
- **Benamou, Carlier -** Static problems with a divergence constraint. We have also extended the Augmented Lagrangian method to static problem where a space divergence constraint appears. This includes the delicate case of the original Monge Optimal Mass Transportation cost (cost=distance) and also Wardrop equilibria in congested transport and related degenerate elliptic equations, like the p-Laplacian operator. A freefem Code has been implemented.

6.3. Multi-marginal problems

- **[Carlier, Oberman (Univ. Mc Gill), Oudet (Univ. of Grenoble) -** New numerical methods for the Wasserstein barycenter and related multi-marginals problems were investigated [49]. A first method uses linear programming, in an implementation that was more efficient than expected. A second method takes advantage of the quadratic structure and leads to an efficient algorithm that can be used in texture synthesis problems arising in image processing.
- **Benamou, Carlier, Nenna** Extension of the CFD formulation and the ALG2 algorithm to the multi marginal problem with quadratic cost (Barycenter).

6.4. JKO gradient flow numerics

../../../../projets/mokaplan/IMG/monge.png

Figure 2. Monge transport flow between sinks and sources.

../../../../projets/mokaplan/IMG/ward.png

Figure 3. Congested transport flows between sink and source.

../../../../projets/mokaplan/IMG/texture_8_1.png

../../../../projets/mokaplan/IMG/texture_8_2.png

../../../../projets/mokaplan/IMG/texture_8_3.png

../../../../projets/mokaplan/IMG/texture_8_mean.

Figure 4. Texture mixing with Wasserstein barycenters, from top to bottom three densities and their barycenter.

- **Benamou, Carlier, Merigot (Univ. of Grenoble, CNRS) , Oudet (Univ. of Grenoble)**q

  A large class of non-linear continuity equations with confinement and/or possibly non local interaction potential can be considered as semi discrete gradient flows with respect to the Euclidean Wassertein distance. The numerical resolution of such problem in dimension 2 and higher is open. Our approach is based on two remarks : the reformulation of the optimization problem in terms of Brenier potential seems to behave better. This introduces a Monge-Ampère operator in the cost functional which needs a monotone discretization in order to preserve the convexity at the discrete level. The first numerical results are very encouraging.

../../../../projets/mokaplan/IMG/heat_onestep.png

Figure 5. One step of Wasserstein JKO gradient flow for the classical entropy (our numerical method) compared to traditional Finite Difference of the heat equation. Left the initial heat profile, right the heat profile after one time step for both methods.

- **Benamou, Carlier, Agueh (Univ. of Victoria)** Splitting methods for kinetic equations, we try to use one JKO step to deal with the non-linear velocity advection part of kinetic equations [31]. This seems to be relevant to granular media equation [16], and also may offer a completely new method for Liouville equations arising from Geometrical Optics [19].

MORPHEME Project-Team

4. New Results

4.1. 3D reconstruction in fluorescence imaging

Participants: Emmanuel Soubies, Laure Blanc-Féraud, Sébastien Schaub.

This work was made in collaboration with Gilles Aubert, Laboratoire J.A. Dieudonné (CNRS,UNS).

We propose a new model for the reconstruction of biological structures using Multiple-Angle Total Internal Reflection Fluorescence Microscopy (MA-TIRFM). This recent microscopy technique allows the visualization of sub-cellular structures around the plasma membrane which is of fundamental importance in the comprehension of exchanges mechanisms of the cell. We present a 3D reconstruction method based on a shape prior information on the observed structures and robust to shot noise and background fluorescence. A novelty with respect to the state of the art is to propose a method allowing the recovery of multiple objects aligned along the axial axis.

TIRFM principle is based on the total internal reflection phenomenon of a light beam at the interface between two mediums of refractive indices n_i (incident) and n_t (transmitted) which produces an evanescent wave capable of exciting fluorophores that are near the coverslip surface. Excited fluorophores emit photons that are then collected by a CCD camera to produce a resulting 2D image (radial dimension). The 2D image formation is formulated as follows [29]:

$$S(x,y,\alpha) = I_0(\alpha) \int_0^{\infty} R(x,y,z) \exp\left(-\frac{z}{d(\alpha,\lambda)}\right) \, dz \tag{18}$$

where $S(x,y,\alpha)$ is the recorded image for the incident angle α, $R(x,y,z)$ denote the 3D unknown fluorophore density, $I_0(\alpha)$ is the intensity at the interface $d(\alpha,\lambda)$ is the penetration depth (theoretically known) and λ is the incident light wavelength. The problem is then to determine R in (1) from acquisitions S_α with different incident angles.

In order to solve this ill-posed inverse problem, we model the 3D unknown fluorophore density by a collection of parametrized objects defined on a state space $\mathcal{X} = P \times M$ by their location $\rho \in P$ and their marks (i.e geometric attributes $\omega \in M$). The optimization problem can be formulated as a minimization problem where both the number of objects in the model and their parameters have to be estimated. This difficult combinatorial optimization problem is tackled by using a Marked Point Process approach [36] which allows modelling interactions between the objects in order to regularize the inverse problem.

Figure 1 right shows the Root Mean Square Errors (RMSE) of each estimated parameter for different noise levels on simulated data. We obtain a hight accuracy reconstruction with an RMSE less than 10 nm for the radial position (x,y) and the radius. A larger RMSE (between 80 and 125 nm, depending on the noise level) is found on the axial position vesicles estimation. As we can see on figure 1 left, the error on the axial position estimation is due to the deepest objects ($>$ 300 nm), objects close to the glass interface are well estimated. Figure 1 right shows also the robustness of the model with respect to shot noise and background fluorescence since the errors remain almost constant with the increasing noise level. The proposed method have also been tested on a real sample of beads of known diameters in order to quantify the quality of the reconstruction. The obtained results are promising for feature estimation of predefined shape structures [17].

4.2. Depth-variant blind restoration for confocal microscopy

Participants: Saima Ben Hadj, Laure Blanc-Féraud.

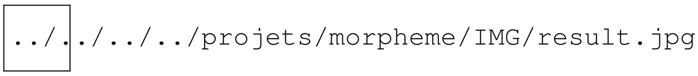


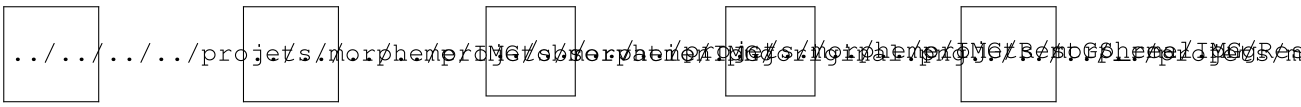
Figure 1. Left : Reconstructions for different noise levels (colors represent z positions of objects). (a) Simulated sample, (b) $\rightarrow$ (c) $\rightarrow$ (d) : Reconstruction for an increasing level of noise. Right : RMSE for different noise levels.

3D images of confocal microscopy basically suffer from two types of distortions: a depth-variant (DV) blur due to the variation of the refractive index between the different mediums composing the system and the imaged specimen, and a Poisson noise due to photon counting process at the sensor.

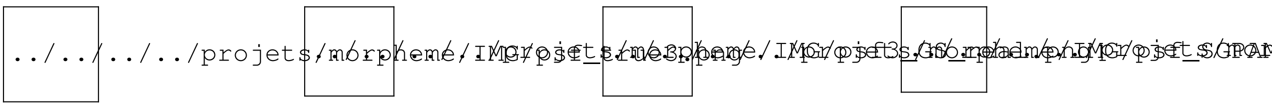
The Point Spread Function (PSF) is depth-variant and its knowledge is crucial for the restoration of these images. Nevertheless, the PSF is inaccessible in practice since it depends on the optical characteristics of the biological specimen and thus needs to be estimated for each different specimen.

In our previous work [5], [4], we developed a method for the joint estimation of the specimen function (the sharp and clean image) and the 3D DV PSF by minimizing a criterion arising from the maximum a posteriori approach. The DV PSF is approximated by a convex combination of a set of space-invariant PSFs taken at different depths.

Recently, we proposed to consider additional constraints on the PSF coming from the optical system modeling [21], [6]. In fact, the confocal microscopy PSF is related to the magnitude of a complex function known as *complex valued-amplitude PSF* whose shape and support are given in the Fourier domain by the numerical aperture of the optical system [30], [35]. This latter is known as it is given by the system manufacturer. We incorporate this constraint in the joint PSF and image estimation algorithm [5] by using the Gerchberg-Saxton algorithm (GS) [31] since it allows to alternate constraints in the spatial and frequency domains. Numerical tests on a simulated image of a bead shell are encouraging (cf. figures 2 (a), (b), (c), and (d) presenting z-slices of the original image, simulated and reconstructed images). In particular, the added constraint allows to better estimate the PSF shape compared to the previous method [5] (cf. figures 2 (e), (f), and (g)).



(a) (b) (c) (d)



(e) (f) (g)

Figure 2. (Y, Z) slices of the simulated observation (a), true image (b) and PSF (e), estimated image (c) and PSF (f) with the additional PSF constraints, estimated image (d) and PSF (g) without the additional PSF constraints.

4.3. Head Tracking and Flagellum Tracing for Sperm Motility Analysis

Participants: Huei Fang Yang, Xavier Descombes, Grégoire Malandain, Sylvain Prigent.

This work is supported in part by ANR MOTIMO project.

Sperm quality assessment plays an important role in human fertility and animal breeding. One of the most important attributes for evaluating semen quality is sperm motility, according to the World Health Organization (WHO) report. When performed manually, semen analysis based on sperm motility is labor-intensive and subject to intra- and inter-observer variability. Computer-assisted sperm analysis (CASA) systems, in contrast, provide rapid and objective semen fertility assessment. In addition, they also offer a means of statistical analysis that may not be achieved by visual assessment. Hence, automated sperm motility analysis systems are highly desirable.

We present a computational framework designed to track the heads and trace the tails for quantitative analysis of sperm motility, which is illustrated in Figure 3 . Our framework includes 3 modules: head detection, head registration, and flagellum tracing. These modules are performed sequentially to obtain the head trajectories and flagellar beat patterns. First, the head detection module detects the sperm heads in the first image of the image data using a Multiple Birth and Cut (MBC) algorithm. The detections are the inputs to the head registration module for obtaining the head trajectories and angles of head rotation. We use a block matching method to register the heads in the subsequent images with respect to the positions and angles of those detected in the first image. This is different from other tracking methods that consider only the head positions. Finally, we propose a flagellum tracing algorithm, based on a Markov chain Monte Carlo (MCMC) sampling method, to obtain the flagellar beat patterns.

We validate our framework using two microscopy image sequences of ram semen samples that were imaged at two different conditions, at which the sperms behave differently. The results show the effectiveness of our framework [19].

4.4. Tree-like Shapes Distance Using the Elastic Shape Analysis Framework

Participants: Alejandro Mottini, Xavier Descombes, Florence Besse.

The analysis and comparison of tree-like shapes is of great importance since many structures in nature can be described by them. In the field of biomedical imaging, trees have been used to describe structures such as neurons, blood vessels and lung airways. Since it is known that axon morphology provides information on their functioning and allows the characterization of pathological states, it is of paramount importance to develop methods to analyze their shape and to quantify differences in structures

We have developed a new method for comparing tree-like shapes that takes into account both topological and geometrical information [14], [15]. Our metric combines the Elastic Shape Analysis Framework originally designed for comparing shapes of 3D closed curves in Euclidean spaces with a matching process between branches. Moreover, the method is able to compute the mean shape of a population of trees.

As a first application, we used our method for the comparison of axon morphology. The performance was tested on a group of 61 (20 normal, 24 type one mutant and 17 type two mutant) 3D images, each containing one axonal tree. We have calculated inter and intra class distances between them and implemented a classification scheme. We have compared our results with the ones obtained by three other methods. Results showed that the proposed method better distinguishes between the two populations than the other methods.

4.5. 3D Modeling of developing organisms

Participants: Gaël Michelin, Grégoire Malandain, Léo Guignard [Virtual Plants], Christophe Godin [Virtual Plants].

This work is made in collaboration with Patrick Lemaire (CRBM).

Image-based studies of developing organs or embryos produce a huge quantity of data. To handle such high-throughput experimental protocols, automated computer-assisted methods are highly desirable. We aim at designing an efficient cell segmentation method from microscopic images. Similary to another work [32], the proposed approach is twofold: first, cell membranes are enhanced or extracted by the means of structure-based filters, and then perceptual grouping (i.e. tensor voting) allows to correct for segmentation gaps (see figure 6). We assessed different structure-based filters as well as different perceptual grouping strategies to identify the most efficient combination, in term of result quality and computational cost [13].

../../../../projets/morpheme/IMG/sperm_analysis_workflow.png

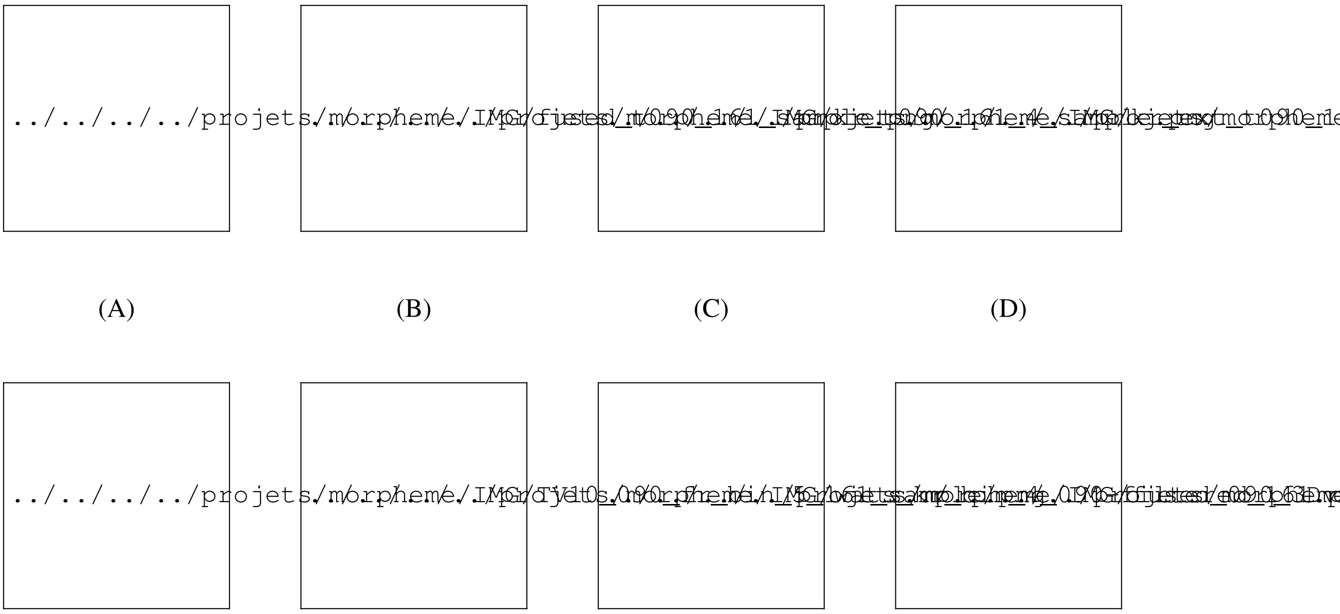
Figure 3. Overview of the proposed framework. The input to our framework is an image sequence. The pre-processing step is to remove the inhomogeneous background and noise. The three main modules in our framework are head detection, head registration, and flagellum tracing. These three modules perform sequentially to obtain the head trajectories and flagellar beat patterns for sperm motility analysis. Note that the output of the head registration module is image sequences for each individual sperms in which the heads are registered. Here, we show the minimum intensity projection (MinIP) of the image sequence.

../../../../projets/morpheme/IMG/normal_tracing.jp

Figure 4. Original confocal microscopy image of an axonal tree (left) and its tracing (right) (maximum intensity projections).

../../../../projets/morpheme/IMG/means_normal_mut.j

Figure 5. Mean normal (left) and mutant (right) axonal trees (2D projections).



(A) (B) (C) (D)

(E) (F) (G) (H)

Figure 6. Illustrations of the different steps of the algorithm: (A) a 2D slice of original image, (B) the resulting surface detector response, (C) the directional extrema of the response image, (D) the deduced binarisation of the cell membranes, (E) the result of Tensor Voting applied to binarised image, (F) the cells segmentation computed from (E), (G) a 3D view of original image, (H) a 3D view of the cells segmentation.

4.6. Spatio-temporal registration of embryo images

Participants: Grégoire Malandain, Léo Guignard [Virtual Plants], Christophe Godin [Virtual Plants].

This work is made in collaboration with Patrick Lemaire (CRBM).

Current imaging techniques can capture temporal sequences of 3D images with very high time resolution over several hours. Comparing sequences covering the same time period opens the way to the study of developmental variability. Stitching together sequences captured from different embryos may help producing a sequence covering the whole development of the animal of interest. For this, it is necessary to align two sequences in both time and space.

We developed a method to align two 3D+t time series, based on the detection and pairing of 3D+t landmarks. These landmarks, which correspond to periods of fast morphogenetic change, are deduced from the analysis of the non-linear transformations that allow to co-register pairs of consecutive 3D images in each sequence (see figure 7). [12].

4.7. Characterizing cell membrane properties

Participants: Sylvain Prigent, Xavier Descombes, Grégoire Malandain, Hélène Barelli [IPMC].

Some mammalian cells show striking differences in the acyl chain composition of their membrane phospholipids. In most cases, the majority of phospholipids bear one saturated and one monounsaturated acyl chains at positions 1 and 2 or the glycerol, respectively. However, some cells and notably neurons contain large amounts of phospholipids with a polyunsaturated fatty acyl chain, generally at position 2. The aim of this work is to compare the impact of the phospholipid polyunsaturation vs monounsaturation on the mechanical and functional properties of the plasma membrane.

For this task, we currently investigate how phospholipid insaturation affects the ability of specialized protein machineries involved in transport vesicle formation, by first detecting vesicles in 2D+t sequences of microscopic images of individual cells, and then tracking detected vesicles through the temporal sequences (see figure8) [23].

4.8. Tracking growing axons in 3D+t fluorescent two-photon microscopy images

Participants: Sylvain Prigent, Xavier Descombes [contact], Florence Besse, Caroline Medioni.

During the maturation of the nervous system, neuronal cells emit cellular extensions (dendrites, axons) allowing them to connect to other neurons, and thus, establish a network in which information is transmitted and/or stored. The formation of axonal extensions and directed migration of these extensions are two key processes controlling the morphology of neuronal cells, and then the number and nature of partners in a given network within a neuron. These two processes are controlled by both external factors to neuronal cell (guidance molecules, neurotrophic signals, ...) and internal factors (transcription factors, post-transcriptional regulators, regulators of the actin cytoskeleton or microtubules, ...). The goal of this work is to automatically extract axonal trajectories from images to then be able to model the processes controlling the morphology of neuronal cells.

The images we use are 3D+t images of growing drosophila brains obtained with a bi-photon confocal microscope. A single movie is about 200 3D frames that correspond to an acquisition every 5 minutes.

The developed method to extract axonal trajectories from 3D+t images is made of 3 main steps. The first step is to detect axonal tips on each 3D frame of the movie. This detection is performed using Marked Point Process. We designed a dedicated model based on an ellipse shape, a prior of no-overlapping between detected ellipses, and a data term calculated by:

$$d = \min(d_B(R_0, R_1), d_B(R_0, R_2), \max(d_B(R_0, R_3), d_B(R_0, R_4)))$$

../../../../projets/morpheme/IMG/final_registration_bis_2.png

Figure 7. Spatio-temporal registration of two time-series of embryo. Enlarged renderings indicate the registered timepoints. Notice that the temporal registration is not linear since the interval length between two registered time is different from one embryo to the next.

../../../../projets/morpheme/IMG/Clat… ../../../../projets/morpheme/IMG/Res…_1_fi…

Figure 8. Left: detection of individual vesicles in one image of the sequence. Right: resulting paths of tracked vesicles through the 2D+t sequence.

where d_B denote the Bhattacharyya distance, R_0 the set of pixels inside the ellipse and R_1, R_2, R_3, R_4 four sets of pixels obtained by partitioning the ellipse contour into 4 regions around the ellipse vertex (see figure 9).

../../../../projets/morpheme/IMG/axons_tip_zoom.png

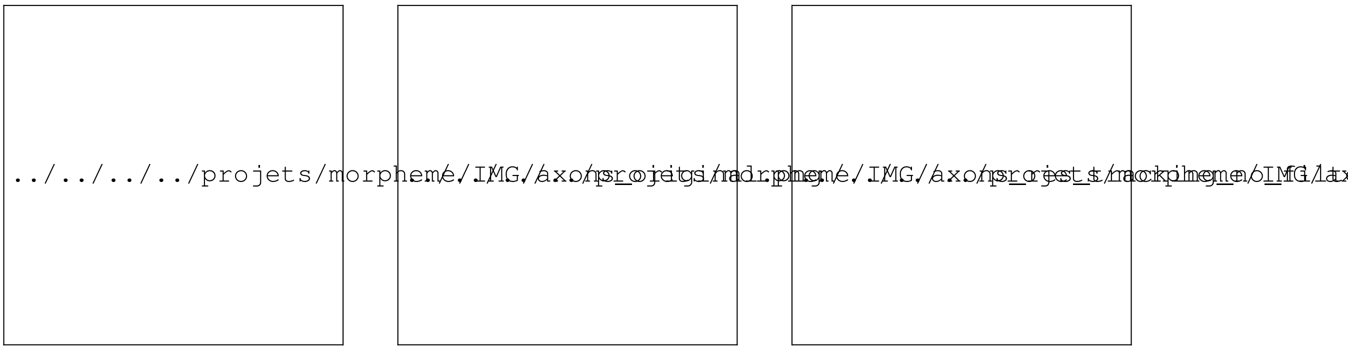
../../../../projets/morpheme/IMG/axons_data_t…

(a) (b)

Figure 9. (a) Zoom on an axon tip. (b) Proposed ellipse model to detect an axon tip.

The second step consists to track the axons along the time frames by linking the tips detections. We designed an association tracking algorithm that builds a graph by connecting spatially close detections in neighboring frames. Negative costs have been introduced to favor long tracks. Then we ran sequentially the shortest path algorithm on this graph to obtain axons trajectories (see figure 10 (b)).

As the axonal tips size is close to the image resolution, the proposed method as the drawback to give false alarms trajectories. We then added a last processing step that aim to analyze the trajectories and remove those that do not correspond to axonal trajectories. This filtering removes three types of trajectories: 1) the trajectories that follow static regions of the image, 2) the short trajectories (less than 5 frames) and 3) random walk trajectories. An example of final automatic tracking is shown in figure 10 (c).



(a) (b) (c)

Figure 10. (a) 2D projection of the last frame. (b) Obtained tracks using the association tracking method. (c) Obtained tracks after false alarms removing.

4.9. A Hierarchical, Graph-cut-based Approach for Extending a Binary Classifier to Multiclass – Illustration with Support Vector Machines

Participants: Alexis Zubiolo, Eric Debreuve, Grégoire Malandain.

The problem of automatic data classification is to build a procedure that maps a datum to a class, or category, among a number of predefined classes. The building of such a procedure is the learning step. Using this procedure to map data to classes is referred to as classification or prediction. The procedure is therefore a classification, or prediction, rule. A datum (text document, sound, image, video, 3-dimensional mesh...) is usually converted to a vector of real values, possibly living in a high-dimensional space, also called signature. Offline, supervised learning relies on a learning set and a learning algorithm. A learning set is a set of signatures that have been tagged with their respective class by an expert. The learning algorithm input is formed by this set together with some parameters, its output being a prediction rule. Some learning algorithm, or method, apply only to the 2-class case. Yet, adapting such a binary classifier to a multiclass context might be preferred to using intrinsically multiclass algorithms, for example if it has strong theoretical grounds and/or nice properties; if free, fast and reliable implementations are available...The most common multiclass extensions of a binary classifier are the one-versus-all (OVA) (or one-versus-rest) and one-versus-one (OVO) approaches. In any extension, several binary classifiers are first learned between pairs of groups of classes. Then, all or some of these classifiers are called when predicting the classes of new samples. When the number of classes increases, the number of classifiers involved in the learning and the prediction steps becomes computationally prohibitive. Hierarchical combinations of classifiers can limit the prediction complexity to a logarithmic law in the number of classes (at best). Combinatorial approaches can be found in the literature. Because of their high learning complexity, these approaches are often disregarded in favor of an approximation trading optimality for computational feasibility. In our work, the high combinatorial complexity is overcome by formulating the hierarchical splitting problems as optimal graph partitionings solved with a minimal cut algorithm. In fact, as this algorithm performs only few additions and comparisons, its impact on the whole procedure is not significant. A modified minimal cut algorithm is also proposed in order to encourage balanced hierarchical decompositions (see Fig. 11). The proposed method is illustrated with the Support Vector Machine (SVM) as the binary classifier. Experimentally, it is shown to perform similarly to well-known multiclass extensions while having a learning complexity only slightly higher than OVO and a prediction complexity ranging from logarithmic to linear. This work has been accepted to the International Conference on Computer Vision Theory and Application (VISAPP 2014) [20].

../../../../projets/morpheme/IMG/svmtree.png

Figure 11. Type of tree that the proposed method builds during the learning stage (illustration with 5 classes). An example of classification of a new image signature is also illustrated by showing the visited nodes in boldface (read from root to leaf).

4.10. Classification of neurons to study Parkinson's disease

Participants: Alexis Zubiolo, Eric Debreuve, Xavier Descombes.

This work has been made in collaboration with Michèle Studer's team at iBV

In this project, the goal is to perform unsupervised classification of rat neurons in order to study the Parkinson's disease. The Institut de Biologie Valrose (iBV) provided us with 3-D images of rat cortices obtained by confocal microscopy. The discriminant features between normal and pathological neurons include the number of dendrites, the length and diameter of the apical dendrite, the shape and size of the soma$\cdots$ For each neuron, these features have to be computed automatically from the images. The specificity of this problem is that, for each rat cortex, we are given several images:

- one low resolution (LR) image which shows an overall view of the cortex and allows to compute the features related to the apical dendrite;
- some high resolution (HR) images (typically between 4 and 6) which provide close-ups of the somas of the neurons and allow to compute the other features.

This work consists in (1) extracting the neurons from the images (see Fig. 12), (2) matching the corresponding neurons in the HR and the LR images, (3) computing the features for each neuron, and (4) classifying the neurons, for example using a kernel Support Vector Machine (SVM).

../../../../projets/morpheme/IMG/initial_image.jpg

../../../../projets/morpheme/IMG/segmented_image

(a) (b)

Figure 12. Extraction of the neurons from the high-resolution image.

4.11. Curve and graph classification using a specific metric and kernel Support Vector Machines

Participants: Vladimir Gutov, Eric Debreuve, Xavier Descombes.

The analysis and comparison of trees are of great importance since many natural structures can be described using such models. In biology, lung airways, neurons, blood vessels··· can be represented by trees (or, more generally, by graphs). Starting from a biological problem (automatically classifying neurons as wild or mutant), we studied the question of using Support Vector Machines (SVM) to classify continuous data such as curves, trees and graphs. Indeed, SVMs are designed for discrete data, looking for an optimal separation hyperplane or manifold in a discrete normed space. Manifolds are found when the original, linear SVM formulation is extended using the so-called kernel trick. The Gaussian kernel is the most popular one. By definition, the isotropic Gaussian kernel involves the two sample data to be compared through the distance between them. This opens the application of Gaussian-kernelized SVMs to any normed space. When dealing with (continuous) curves, the Fréchet distance can be used. We also tested a metric based on shape analysis [34]. Finally, a (meta-)distance between trees proved to be efficient in comparing axons [33] (see Fig. 13). The “meta” qualifier means that this distance builds upon a metric between curves and is valid for any such metric. It was tested using the shape-based metric [34]. We adapted an open-source SVM implementation to be able to use the three aforementioned metrics (two between curves, one between trees) and we validated the classification approach on synthetic data and on a small database of 20 wild-type neurons and 24 mutants provided by biologists.

4.12. Random forests for zooplankton classification

Participant: Eric Debreuve.

This work has been made in collaboration with Florent Baronian (Engineering student), Luc Deneire (I3S) and Marc Picheral (LOV)

../../../../projets/morpheme/IMG/axon.png

Figure 13. Left: Axonal tree imaged with a confocal microscope (maximum intensity projection of the 3-D acquisition). Right: Manual extraction of the axon.

An UVP embedded system (Underwater Vision Profiler) is a device mainly composed of a digital camera with a fixed focal distance, and a flash system designed to illuminate only the focal plane. The device is attached to a boat by a cable and it is let going deep to take pictures at various depths. The purpose of such acquisition campaigns is to analyze the population of zooplankton organisms in different oceanic regions (see Fig. 14). The fifth version of the UVP developed at the Laboratoire d'Océanographie de Villefranche (LOV) only stores the pictures. All the processing is done offline: zooplankton organisms are segmented, features are extracted and a classification into types of organisms is performed. With the upcoming sixth version, the goal is to make the device smaller, lighter, and autonomous (for some time) in order to be placed it in appliances designed to drift or navigate autonomously for weeks or months. This imposes to perform all the image processing tasks aboard, which limits the available processing power. Our work is to propose a classification method taking into account the constraints given by the teams in charge of the hardware design. We implemented a Random forest-based classifier which combines both good performances and low computational requirements. Since the images contain a lot of spurious objects called aggregates, we proposed a two-stage approach: the first stage is either a binary classifier or model checker tailored to eliminate the aggregates, while the second stage actually classifies the zooplankton organisms. We tested a combination of a one-class SVM (model checker) and a Random forest, and a combination of two Random forests, the first one being restricted to a binary classifier. Results were encouraging.

Figure 14. Left: Version 5 of the UVP embedded system (Underwater Vision Profiler) used to take pictures of the zooplankton at different depths. Right: Some organisms composing the zooplankton (Lars Stemmann, LOV). Images acquired by the UVP have a much worse resolution.

4.13. Detection of Hedgehog protein using confocal imaging

Participants: Sylvain Prigent, Xavier Descombes.

This work was made in collaboration with P. Therond's group at iBV

P. Therond's lab is focusing on the understanding of how the secreted Hedgehog (Hh) morphogen, a dually lipidated highly hydrophobic molecule bound to membranes, is secreted, released and transported from the place of production in Drosophila. High resolution microscopy developed to identify and visualize such processes was successfully applied to address this question, but dynamics of membrane transport is still poorly understood mainly due to the lack of a reliable model or the need of fixation.

To statistically quantify the position of a population of Hedgehog proteins inside a cell, we need an automatic image processing method that detect each individual Hedgehog proteins in a 3D confocal image. The 3D images are obtained either by scanning Z slices (see figure 15), or by scanning XY slices (see figure 16). For both types of images, we used Marked Point Process (MPP) to detect 2D objects independently on each frame since proteins appear only in one slice. Using the Z slices, we observed that proteins appear as vertical rectangles whose size is close to the image resolution. We then design an MPP model to search for a configuration of vertically oriented rectangles that do not overlap one to each other. To define a data term for a given rectangle in the image, we calculate the Bhattacharyya distance between the population of pixels inside the rectangle and the population of pixels in the border of the rectangle. For the XY slices, we defined an MPP model to search for circles that do not overlap one to each other, and the data term is calculated with the Bhattacharyya distance as for the rectangle model. Examples of obtained results are shown in figures 15 and 16 .

(a)

(b)

Figure 15. Example of proteins detection on a single z image slice (a) Original z slice. (b) Obtained proteins detection.

(a) (b)

Figure 16. Example of proteins detection on a single xy image slice (a) Original xy slice. (b) Obtained proteins detection.

MORPHEO Team

6. New Results

6.1. Robust human body shape and pose tracking

This work considers markerless human performance capture from multiple camera videos and, in particular, the recovery of both shape and parametric motion information, as often required in applications that produce and manipulate animated 3D contents using multiple videos. To this aim, an approach is proposed that jointly estimates skeleton joint positions and surface deformations by fitting a reference surface model to 3D point reconstructions. The approach is based on a probabilistic deformable surface registration framework coupled with a bone binding energy. The former makes soft assignments between the model and the observations while the latter guides the skeleton fitting. The main benefit of this strategy lies in its ability to handle outliers and erroneous observations frequently present in multi view data. For the same purpose, we also introduce a learning based method that partitions the point cloud observations into different rigid body parts that further discriminate input data into classes in addition to reducing the complexity of the association between the model and the observations. We argue that such combination of a learning based matching and of a probabilistic fitting framework efficiently handle unreliable observations with fake geometries or missing data and hence, it reduces the need for tedious manual interventions. The work was presented at the 3DV conference [7] where it received the best paper runner up award.

Figure 4. Human pose recovery with 3 different standard datasets.

6.2. Inverse dynamics on rock climbing with and without measurement of contact forces

Rock climbing involves complex interactions of the body with the environment (Figure 5). It represents an interesting problem in biomechanics as multiple contacts in the locomotion task make it an underconstrained problem. In this study we are interested in evaluating how a climber transfers weight through the holds. The motivation of this study is also technical as we are developing an inverse dynamics method that automatically estimates in 3D, not only the usual torques at joint angles, but also the wrenches at contacts [9].

../../../../projets/morpheo/IMG/climbInvDyn.png

Figure 5. Inverse dynamics on rock climbing with and without measurement of contact forces.

6.3. Video-based methodology for markerless human motion analysis

This study presents a video-based experiment for the study of markerless human motion. Silhouettes are extracted from a multi-camera video system to reconstruct a 3D mesh for each frame using a reconstruction method based on visual hull. For comparison with traditional motion analysis results, we set up an experiment integrating video recordings from 8 video cameras and a Vicon™ marker-based motion capture system (Figure 6). Our preliminary data provided distances between the 3D trajectories from the Vicon system and the 3D mesh extracted from the video cameras. In the long term, the main ambition of this method is to provide measurement of skeleton motion for human motion analyses while eliminating markers [8].

../../../../projets/morpheo/IMG/ACAPS2013.png

Figure 6. Video-based methodology for markerless human motion analysis.

6.4. 3D shape cropping

We introduce shape cropping as the segmentation of a bounding geometry of an object as observed by sensors with different modalities. Segmenting a bounding volume is a preliminary step in many multi-view vision applications that consider or require the recovery of 3D information, in particular in multi-camera environments. Recent vision systems used to acquire such information often combine sensors of different types, usually color and depth sensors. Given depth and color images we present an efficient geometric algorithm to compute a polyhedral bounding surface that delimits the region in space where the object lies. The resulting cropped geometry eliminates unwanted space regions and enables the initialization of further processes including surface refinements. Our approach exploits the fact that such a region can be defined as the intersection of 3D regions identified as non empty in color or depth images. To this purpose, we propose a novel polyhedron combination algorithm that overcomes computational and robustness issues exhibited by traditional intersection tools in our context. We show the correction and effectiveness of the approach on various combination of inputs. This work was presented at the Vision Modeling and Visualization workshop 2013 [6].

6.5. Multi-view object segmentation in space and time

In this work, we address the problem of object segmentation in multiple views or videos when two or more viewpoints of the same scene are available. We propose a new approach that propagates segmentation coherence information in both space and time, hence allowing evidences in one image to be shared over the complete set. To this aim the segmentation is cast as a single efficient labeling problem over space and time with graph cuts. In contrast to most existing multi-view segmentation methods that rely on some form of dense reconstruction, ours only requires a sparse 3D sampling to propagate information between viewpoints. The approach is thoroughly evaluated on standard multi-view datasets, as well as on videos. With static views, results compete with state of the art methods but they are achieved with significantly fewer viewpoints. With multiple videos, we report results that demonstrate the benefit of segmentation propagation through temporal cues, in ICCV 2013 [5].

6.6. Segmentation of temporal mesh sequences into rigidly moving components

This work considers the segmentation of meshes into rigid components given temporal sequences of deforming meshes (Figure 9). We have proposed a fully automatic approach that identifies model parts that consistently move rigidly over time. This approach can handle meshes independently reconstructed at each time instant. It allows therefore for sequences of meshes with varying connectivities as well as varying topology. It incrementally adapts, merges and splits segments along a sequence based on the coherence of motion information within each segment. In order to provide tools for the evaluation of the approach, we also introduce new criteria to quantify a mesh segmentation. Results on both synthetic and real data as well as comparisons are provided in the paper [1].

6.7. Segmentation of plant point cloud models into elementary units

High-resolution terrestrial Light Detection And Ranging (tLiDAR), a 3-D remote sensing technique, has recently been applied for measuring the 3-D characteristics of vegetation from grass to forest plant species. The resulting data are known as a point cloud which shows the 3-D position of all the hits by the laser beam giving a raw sketch of the spatial distribution of plant elements in 3-D, but without explicit information on their geometry and connectivity. In this study we propose a new approach based on a delineation algorithm that clusters a point cloud into elementary plant units. The algorithm creates a graph (points + edges) to recover plausible neighbouring relationships between the points and embed this graph in a spectral space in order to segment the point-cloud into meaningful elementary plant units. Our approach is robust to inherent geometric outliers and/or noisy points and only considers the x, y, z coordinate tLiDAR data as an input. It has been presented at the FSPM conference [4].

../../../../projets/morpheo/IMG/teaser.png

Figure 7. Result of shape cropping using three input depth maps for polyhedral reconstruction.

../../../../projets/morpheo/IMG/teaser2.png

Figure 8. Multi-view object segmentation using our method with the 3 wide-baseline views shown only, with no photo-consistency hypothesis and no user interaction.

../../../../projets/morpheo/IMG/these_romain.png

Figure 9. Segmentation of temporal mesh sequences into rigidly moving components.

../../../../projets/morpheo/IMG/fspm.jpg

Figure 10. Segmentation of a plant point cloud model into elementary units.

MUTANT Project-Team

6. New Results

6.1. Operational Timed Semantics

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

One common use-case of real-time musical interactions between musicians and computers is *Automatic Accompaniment* where the system is comprised of a real-time machine listening system that in reaction to recognition of events in a score from a human performer, launches necessary actions for the accompaniment section. While the real-time detection of score events out of live musicians' performance has been widely addressed in the literature, score accompaniment (or the reactive part of the process) has been rarely discussed. In [13], we are trying to deal with this missing component in the literature from a programming language perspective. We show how language considerations would enable better authoring of time and interaction during programming/composing and how it addresses critical aspects of a musical performance (such as errors) in real-time. We sketch the real-time features required by automatic musical accompaniment seen as a reactive system and formalize the timing strategies for musical events taking into account the various temporal scales used in music. Various strategies for the handling of synchronization constraints and the handling of errors are presented.

The behavior of the system *Antescofo* have been formally modeled as a *network of parametric timed automata*. 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 enables better authoring of time and interaction during programing/composing, permitting to use state of the art software verification tools for the static analysis of Antescofo scores. It also provides means to address critical aspects of musical performances in real-time.

6.2. Timed Static Analysis of Interactive Music Scores

Participant: Florent Jacquemard.

It is well known that every musician performance of the same work will differ from another. It is therefore a challenging task to be able to predict the behavior of interactive music systems like *Antescofo* in response to any possible performance, and prevent unwanted outcomes. With Léa Fanchon, we have been working on a module for timing analysis of augmented scores that complements the real-time score authoring and performance in *Antescofo*, with the aim of exploring possible behavior of authored scores with respect to possible deviations in human musician performance.

For this purpose we have studied [24] the application of formal models and methods from the literature of real-time systems verification to the static analysis of interactive music systems. We have considered in particular the good parameters problem, which consists in synthesizing a set of timing parameter valuations (representing performances here) guarantying a good behavior of the system analyzed. The methods presented in [24] have been applied to *Antescofo*, providing the following input to users:

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

This study is one of the first of this kind in computer music literature, and the methods presented are general enough to apply to the verification of other interactive multimedia applications.

6.3. Automating the Generation of Test Suites for Antescofo

Participants: Florent Jacquemard, Clément Poncelet.

Clément Poncelet has started to develop during his Master thesis [35] a framework for black box conformance testing of *Antescofo*. This work is pursued in a PhD supported by DGA and Inria. The most important task in this context is the generation of relevant test data for the system, given an augmented score in *Antescofo* language. This data includes input, containing musical events (notes, chords etc) together with their timings. In a sense, the input data simulates a musical execution of the score. The input data must then be passed to *Antescofo* for black-box execution, in order to observe the system's reactions and compare them the expected output. For the latter comparison task, we need to be able to define the expected output, hence to have a formal model of the expected behavior of the system on the given score. For this purpose, we are using models of the system made of timed automata, which are computed automatically from given music scores. Then, we use tools from the UPPAAL suite [40] in order to generate testing data, based on relevant covering criteria and a formal model of the environment (i.e. the musician). This work has been presented at the poster session of MSR 2013 (national colloquium on modeling reactive systems) and a journal paper is in preparation.

6.4. Synchronous Embedding of Antescofo DSL

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

Antescofo can been seen as the coupling of a listening machine and a real-time reactive system. Therefore, it faces some of the same major challenges as embedded systems. We have been working with Guillaume Baudart, Louis Mandel, and Marc Pouzet (EPI Parkas, ENS) in strengthening the ties between the reactive aspects of *Antescofo* and that of synchronous languages, in particular ReactiveML [44]. In [17], we present a synchronous semantics for the core language of Antescofo and an alternative implementation, based on an embedding inside the synchronous language ReactiveML [44]. The semantics reduces to a few rules, is mathematically precise and leads to an interpretor of a few hundred lines whose efficiency compares well with that of the current implementation. On all musical pieces we have tested, response times have been less than the reaction time of the human ear. Moreover, this embedding permitted the prototyping of several new programming constructs. Some examples are available, together with the ReactiveML source code at http://reactiveml.org/emsoft13/.

6.5. Tree Structured Representation of Symbolic Temporal Data

Participant: Florent Jacquemard.

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).

Following this idea, we have been working on using several tree automata techniques for the challenging and long-standing problem of automatic transcription of rhythm (in traditional music notations) from symbolic input data (symbolic traces with timestamps in ms, like e.g. in MIDI format). To summarize, the main problem in rhythm transcription is to find an acceptable balance between timing precision (the goal is to minimize the loss obtained by transformation of ms timing values into fractions of beats) and the complexity of the notation obtained. The relative importance of these two measures may vary largely according to the user (composer), his workflow, and the musical style considered. It is therefore important to be able to control this balance during the transcription process, in order to adapt to the case of users. In traditional approaches, the transcription is done by an alignement of the input trace on a grid, and the two measures (precision of the grid and complexity) are either defined by parameters fixed a priori or hardcoded e.g. for a precise musical style and composition workflow. During two internships co-supervised by Jean Bresson (Ircam, main developer of OpenMusic) and Florent Jacquemard, we have been studying more flexible new approaches, based on computations on the tree representation of rhythms.

Pierre Donat-Bouillud (L3 ENS Rennes) [29] has worked on an approach by transformation of trees following some rewrite rules. The general idea is to start with a complex tree representing timings very close to the input data, and to simplify it by rewriting until an acceptable level of complexity is reached. The rewrite rules are either generic (defining an equational theory of rhythm notation) or user defined (defining approximations). This approach has been implemented in an OpenMusic library.

Adrien Maire (M1 ENS Cachan) has studied another very promising approach based on stochastic tree automata learning in an interactive authoring scenario. The generated automaton is supposed to represent (by the weighted tree langage it defines) the expected complexity of rhythm notations (i.e. the user's "style").

Moreover, we have following other work on several classes of tree recognizers and tree transformations which could be of interest in this context. With Luis Barguñó, Carlos Creus, Guillem Godoy, and Camille Vacher, [11] we define a class of ranked tree automata called TABG generalizing both the tree automata with local brother tests of Bogaert and Tison [37] and with global equality and disequality constraints (TAGED) of Filiot et al. [39]. TABG can test for equality and disequality modulo a given flat equational theory between brother subterms and between subterms whose positions are defined by the states reached during a computation. In particular, TABG can check that all the subterms reaching a given state are distinct. This constraint is related to monadic key constraints for XML documents, meaning that every two distinct positions of a given type have different values. We have proven decidability of the emptiness problem for TABG. This solves, in particular, the open question of decidability of emptiness for TAGED. We further extended our result by allowing global arithmetic constraints for counting the number of occurrences of some state or the number of different equivalence classes of subterms (modulo a given flat equational theory) reaching some state during a computation. We also adapt the model to unranked ordered terms. As a consequence of our results for TABG, we prove the decidability of a fragment of the monadic second order logic on trees extended with predicates for equality and disequality between subtrees, and cardinality.

With Michaël Rusinowitch (EPI Cassis), we have introduced in [25] an extension of unranked tree automata called bi-dimensional context-free hedge automata. The languages they define are context free in two dimensions: in the the sequence of successors of a node and also along paths. This formalism is useful for the static type-checking of tree transformations such as XML updates defined in the W3C XQuery Update Facility. We have developed with the same author in the past years a general framework for the verification of unranked (XML) tree transformations based on tree automata techniques. It has been presented this year in an invited keynote [16]. We have also presented with Emmanuel Filiot and Sophie Tison a survey on tree automata with constraints [33] during a Dagstuhl Seminar (number 13192) on tree transducers and formal methods.

6.6. Online Automatic Structure Discovery of Audio Signals

Participants: Arshia Cont, Vincent Lostanlen [MS Internship].

Following recent team findings in [12] and the framework introduced in [4], we pursued the problem of automatic discovery of audio signals using methods of information geometry through a Masters Thesis undertaken by Vincent Lostanlen (MS ATIAM) [34]. This work introduces a novel way of representing and calculating *Similarity Matrices* for continuous multimedia signals and in real-time. In this approach, the signal is first segmented into homogeneous chunks using the change detection algorithm proposed by the team in [12], and proposes a method for constituting similarity relations between segments using *Bregman Information Geometry* and exploiting intersections between information balls.

Compared to traditional approaches to similarity matrix computing, the approach proposed in [34] is strictly on-line (thus suitable for real-time computing) and provides a sparse view of audio structures. We will pursue this project by increasing its robustness and evaluating results on larger databases including other timed-signals such as video.

6.7. Temporal Coherency Criterion for Alignment Inference Algorithms

Participants: Philippe Cuvillier [PhD Student], Arshia Cont.

The question of modeling time and duration is of utmost importance for stability and robustness of real-time alignment algorithms and constitute one of the major success factors for the *Antescofo* listening machine described in [2]. Meanwhile, regular algorithms undergo stability in highly uncertain environments where observations obtained from the signal are highly uninformative and temporal information is of crucial importance.

PhD student Philippe Cuvillier defined *Coherency Criteria* for such applications and attempted to formalize such criteria in terms of probabilistic models and inference algorithms in case of Hidden Semi-Markov Chains. The results show that not all probabilistic families meet such criteria including some commonly used by engineers and designers. Preliminary results are submitted for publications and experimental results are being pursued.

MYRIADS Project-Team

6. New Results

6.1. Dependable Cloud Computing

Participants: Roberto-Gioacchino Cascella, Stefania Costache, Florian Dudouet, Eugen Feller, Filippo Gaudenzi, Yvon Jégou, Ancuta Iordache, David Margery, Christine Morin, Anne-Cécile Orgerie, Guillaume Pierre, Nikos Parlavantzas, Yann Radenac, Matthieu Simonin, Cédric Tedeschi.

6.1.1. Multi-data Center and Multi-cloud

6.1.1.1. Deployment of distributed applications in a multi-provider environment

Participants: Roberto-Gioacchino Cascella, Stefania Costache, Florian Dudouet, Piyush Harsh, Filippo Gaudenzi, Yvon Jégou, Christine Morin.

The move of users and organizations to Cloud computing will become possible when they will be able to exploit their own applications, applications and services provided by cloud providers as well as applications from third party providers in a trustful way on different cloud infrastructures. In the framework of the Contrail European project [39] [50], we have designed and implemented the Virtual Execution Platform (VEP) service in charge of managing the whole life cycle of OVF distributed applications under Service Level Agreement rules on different infrastructure providers [51]. In 2013, we designed the CIMI inspired REST-API for VEP 2.0 with support for Constrained Execution Environment (CEE), advance reservation and scheduling service, and support for SLAs [55], [54] [56]. We integrated support for delegated certificates and developed test scripts to integrate the Virtual Infrastructure Network (VIN) service. VEP 1.1 was slightly modified to integrate the usage control (Policy Enforcement Point (PEP)) solution developed by CNR. The CEE management interface was developed during 2013 and is available through the graphical API as well as through the RESTful API.

6.1.1.2. Towards a distributed cloud inside the backbone

Participants: Anne-Cécile Orgerie, Cédric Tedeschi.

The DISCOVERY proposal currently in phase of construction and lead by Adrien Lèbre from ASCOLA team, and currently on leave at Inria aims at designing a distributed cloud, leveraging the resources we can find in the network's backbone. [3]

In this context, and in collaboration with ASCOLA and ASAP teams, we started the design of an overlay network whose purpose is to be able, with a limited cost, to locate geographically-close nodes from any point of the network. The basis for this overlay is described as part of a recent research report [44].

6.1.1.3. Multi-cloud application deployment in ConPaaS

Participants: Guillaume Pierre, Yann Radenac.

We extended ConPaaS to support application deployment over multiple clouds. There are two main reasons for this: first, it is a necessary mechanism to allow application migration from one cloud to another, without any service interruption. Second, for some applications it may be useful to execute over multiple clouds on a permanent basis, for reliability reasons for example. The main challenges to address were ensuring full network connectivity between resources acquired in multiple clouds. We addressed these issues by integrating the IPOP virtual network in ConPaaS. Second, we designed protocols to ensure application and data migration without any service interruption during the migration.

6.1.2. Scalability of Snooze Self-healing Cloud Management System

Participants: Eugen Feller, Yvon Jégou, David Margery, Christine Morin, Anne-Cécile Orgerie, Matthieu Simonin.

[3] The DISCOVERY website: http://beyondtheclouds.github.io

We evaluated the scalability and resilience of Snooze IaaS management system [26]. Unlike existing systems, for scalability, ease of configuration, and high availability, Snooze is based on a self-organizing and self-healing hierarchical architecture of system services [36], [27], [27]. In Snooze hierarchy, each compute server is managed by a local controller that interacts with one of the group managers to which it is dynamically assigned and the set of group managers is coordinated by a group leader elected among them. We performed an extensive scalability study of Snooze across over 500 servers of the Grid'5000 experimentation testbed. We evaluated the Snooze self-organizing and self-healing hierarchy with thousands of system services. The results show that the resource consumption of the Snooze system services is bounded both during the hierarchy construction and system operation. We also show that Snooze prototype implementation is robust enough to manage thousands of servers and hundreds of VMs. Moreover, its autonomic behavior allows to achieve high availability in the presence of a large number of simultaneous system services failures. Indeed, as long as at least two group managers remain operational the system remains alive. We also demonstrated the application deployment scalability across hundreds VMs on the example of a Hadoop MapReduce application. We participated in the Scale Challenge organized in the framework of the ACM/IEEE CC-Grid 2013 conference [26] and won the second prize.

6.1.3. Application Performance Modeling in Heterogeneous Cloud Environments

Participants: Ancuta Iordache, Guillaume Pierre.

Heterogeneous cloud platforms offer many possibilities for applications for make fine-grained choice over the types of resources they execute on. This opens for example opportunities for fine-grain control of the tradeoff between expensive resources likely to deliver high levels of performance, and slower resources likely to cost less. We designed a methodology for automatically exploring this performance vs. cost tradeoff when an arbitrary application is submitted to the platform. Thereafter, the system can automatically select the set of resources which is likely to implement the tradeoff specified by the user. A publication on this topic is currently in preparation.

6.1.4. Flexible SLA & SLO Management

Participants: Stefania Costache, Christine Morin, Nikos Parlavantzas.

Merkat is a market-based, SLO-driven, PaaS system for private clouds. Merkat dynamically shares resources between competing applications to ensure a fair resource utilization in terms of application priority and actual resource needs. Resources are allocated through a proportional-share auction while autonomous controllers apply elasticity rules to scale application demand according to resource availability and user priority. Merkat provides users the flexibility to adapt controllers to their application types, and it can support diverse application types and performance goals. Merkat is implemented in Python and uses OpenNebula for virtual machine operations.

We evaluated Merkat in simulation and we analyzed the behavior of the system for multiple user types [23]. Furthermore, we deployed Merkat on Grid'5000 and EDF's tested and tested it with applications representative to EDF [22]. Results showed that: (i) the system provides flexible support for different application types (static and malleable) and different SLOs (deadline and performance); (ii) the system provides good user satisfaction achieving acceptable performance degradation, compared to existing centralized solutions. Furthermore, we extended Merkat to manage different clusters and run MPI applications on them. We also submitted a survey on evolution of resource management systems for shared virtualized computing infrastructures to an international journal. This work was carried out in the framework of Stefania Costache's PhD thesis [11].

6.2. Heterogeneous Resource Management

Participants: Eliya Buyukkaya, Djawida Dib, Eugen Feller, Tran Ngoc Minh, Christine Morin, Nikos Parlavantzas, Guillaume Pierre.

6.2.1. Cross-resource scheduling in heterogeneous cloud environments

Participants: Eliya Buyukkaya, Tran Ngoc Minh, Guillaume Pierre.

Allocating resources to applications in a heterogeneous cloud environment is harder than in a homogeneous environment. In a heterogeneous cloud some rare resources are more precious than others, and should be treated carefully to maximize their utilization. Similarly, applications may request groups of resources that exhibit certain inter-resource properties such as the available bandwidth between the assigned resources. We are currently investigating scheduling algorithms for handling such scenarios.

6.2.2. Maximizing private cloud provider profit in cloud bursting scenarios

Participants: Christine Morin, Djawida Dib, Nikos Parlavantzas.

Current PaaS offerings either provide no support for SLA guarantees or provide limited support targeting a restricted set of application types. To overcome this limitation, we are developing an open, SLA-driven PaaS system, called Meryn, that aims at providing SLA guarantees to diverse application types while maximizing the PaaS provider profit. Meryn supports cloud bursting and applies a decentralized protocol for selecting cloud resources, trying to minimize the cost of running applications without affecting their agreed quality properties. We have performed a preliminary evaluation of Meryn [24] and worked on optimising the system and performing further experiments on the Grid5000 testbed. This work is part of Djawida Dib's PhD thesis.

6.2.3. Data life-cycle management in clouds

Participants: Eugen Feller, Christine Morin.

Infrastructure as a Service (IaaS) clouds provide a flexible environment where users can choose and control various aspects of the machines of interest. However, the flexibility of IaaS clouds presents unique challenges for storage and data management in these environments. Users use manual and/or ad-hoc methods to manage storage and data in these environments. FRIEDA is a Flexible Robust Intelligent Elastic Data Management framework that employs a range of data management strategies approaches in elastic environments. In the context of the DALHIS associate team [4], we evaluated the importance of this framework on multiple cloud testbeds. Our evaluation showed that storage planning needs to be performed in coordination with compute planning and the specific configuration of virtual machine had a strong impact on the application (e.g., some applications performed better on small instances than large instances) [40].

6.3. Energy-efficient Resource Infrastructures

Participants: Alexandra Carpen-Amarie, Bogdan Florin Cornea, Ismael Cuadrado Cordero, Djawida Dib, Eugen Feller, Yunbo Li, Christine Morin, Anne-Cécile Orgerie, Guillaume Pierre.

6.3.1. Energy-efficient IaaS clouds

Participants: Alexandra Carpen-Amarie, Christine Morin, Anne-Cécile Orgerie.

Energy consumption has always been a major concern in the design and cost of data centers. The wide adoption of virtualization and cloud computing has added another layer of complexity to enabling an energy-efficient use of computing power in large-scale settings. Among the many aspects that influence the energy consumption of a cloud system, the hardware-component level is one of the most intensively studied. However, higher-level factors such as virtual machine properties, their placement policies or application workloads may play an essential role in defining the power consumption profile of a given cloud system. In this work, we explored the energy consumption patterns of Infrastructure-as-a-Service (IaaS) cloud environments under various synthetic and real application workloads. For each scenario, we investigated the power overhead triggered by different types of virtual machines, the impact of the virtual cluster size on the energy-efficiency of the hosting infrastructure and the tradeoff between performance and energy consumption of MapReduce virtual clusters through typical cloud applications [21].

6.3.2. Energy-aware IaaS-PaaS co-design

Participants: Alexandra Carpen-Amarie, Djawida Dib, Guillaume Pierre, Anne-Cécile Orgerie.

[4] http://project.inria.fr/dalhis

The wide adoption of the cloud computing paradigm plays a crucial role in the ever-increasing demand for energy-efficient data centers. Driven by this requirement, cloud providers resort to a variety of techniques to improve energy usage at each level of the cloud computing stack. However, prior studies mostly consider resource-level energy optimizations in IaaS clouds, overlooking the workload-related information locked at higher levels, such as PaaS clouds. We argue that cross-layer cooperation in clouds is a key to achieving an optimized resource management, both performance and energy-wise. To this end, we claim there is a need for a cooperation API between IaaS and PaaS clouds, enabling each layer to share specific information and to trigger correlated decisions. We identified the drawbacks raised by such co-design objectives and discuss opportunities for energy usage optimizations, and plan to start the research to address these issues in 2014.

6.3.3. Performance and energy-efficiency evaluation of Hadoop deployment models

Participants: Eugen Feller, Christine Morin.

The exponential growth of scientific and business data has resulted in the evolution of the cloud computing and the MapReduce parallel programming model. Cloud computing emphasizes increased utilization and power savings through consolidation while MapReduce enables large scale data analysis. The Hadoop framework is the most popular open source software implementing the MapReduce model. In our work, we evaluated Hadoop performance in two modes – the traditional model of collocated data and compute services and separated mode where the services are deployed on separate services. The separation of data and compute services provides more flexibility in environments where data locality might not have a considerable impact such as virtualized environments and clusters with advanced networks. In this work, we also conducted an energy efficiency evaluation of Hadoop on physical and virtual clusters in different configurations. The experiments were performed on the Grid'5000 experimentation testbed. To enable virtual machine management, the Snooze cloud stack developed by the Myriads project-team was used. Our extensive evaluation shows that: (1) performance on physical clusters is significantly better than on virtual clusters; (2) performance degradation due to separation of the services depends on the data to compute ratio; (3) application completion progress correlates with the power consumption and power consumption is heavily application specific [28].

6.3.4. Energy consumption models and predictions for large-scale systems

Participant: Christine Morin.

Responsible, efficient and well-planned power consumption is becoming a necessity for monetary returns and scalability of computing infrastructures. While there is a variety of sources from which power data can be obtained, analyzing this data is an intrinsically hard task. In our work, we described a generic approach to analyze large power consumption datasets collected from computing infrastructures. As a first step, we proposed a data analysis pipeline that can handle the large-scale collection of energy consumption logs, apply sophisticated modeling to enable accurate prediction, and evaluate the efficiency of the analysis approach. We presented the analysis of a power consumption data set collected over a 6-month period from two clusters of the Grid'5000 experimentation platform used in production. We used Hadoop with Pig to handle the large volume of data. Our data processing generated a summary of the data that provides basic statistical aggregations, over different time scales. The aggregate data was then analyzed as a time series using sophisticated modeling methods with R statistical software. We exploited time series to detect outliers and derive hourly and daily power consumption predictive models. We demonstrated the accuracy of the predictive models and the efficiency of the data processing performed on a 55-node cluster at NERSC [34]. Energy models from such large dataset can help in understanding the evolution of consumption patterns, predicting future energy trends, and providing basis for generalizing the energy models to similar large-scale systems.

6.3.5. Simulating Energy Consumption of Wired Networks

Participant: Anne-Cécile Orgerie.

Predicting the performance of applications, in terms of completion time and resource usage for instance, is critical to appropriately dimension resources that will be allocated to these applications. Current applications, such as web servers and Cloud services, require lots of computing and networking resources. Yet, these resource demands are highly fluctuating over time. Thus, adequately and dynamically dimension these resources is challenging and crucial to guarantee performance and cost-effectiveness. In the same manner, estimating the energy consumption of applications deployed over heterogeneous cloud resources is important in order to provision power resources and make use of renewable energies. Concerning the consumption of entire infrastructures, some studies show that computing resources represent the biggest part in Cloud's consumption, while others show that, depending on the studied scenario, the energy cost of the network infrastructure that links the user to the computing resources can be bigger than the energy cost of the servers. In this work, we aim at simulating the energy consumption of wired networks which receive little attention in the Cloud computing community even though they represent key elements of these distributed architectures. To this end, we are contributing to the well-known open-source simulator ns3 by developing an energy consumption module named ECOFEN.

6.3.6. Simulating the impact of DVFS within SimGrid

Participants: Alexandra Carpen-Amarie, Christine Morin, Anne-Cécile Orgerie.

Simulation is a a popular approach for studying the performance of HPC applications in a variety of scenarios. However, simulators do not typically provide insights on the energy consumption of the simulated platforms. Furthermore, studying the impact of application configuration choices on energy is a difficult task, as not many platforms are equipped with the proper power measurement tools. The goal of this work is to enable energy-aware experimentations within the SimGrid simulation toolkit, by introducing a model of application energy consumption and enabling the use of DVFS techniques for the simulated platforms. We provide the methodology used to obtain accurate energy estimations, highlighting the simulator calibration phase. The proposed energy model is validated by means of a large set of experiments featuring several benchmarks and scientific applications. This work is available in the latest SimGrid release.

6.4. Unconventional Models for Large Computations and Platforms

Participants: Marko Obrovac, Christine Morin, Cédric Tedeschi.

6.4.1. Chemical computing at large scale

Participants: Marko Obrovac, Cédric Tedeschi.

One of the commonly cited problem when dealing with chemistry-inspired computing is its lack of experimental validation. The DHT-based runtime developed recently, in the framework of Marko Obrovac's PhD thesis [13], has been deployed over the Grid'5000 platform with promising results. This runtime is now mature enough for being considered as a viable candidate to underlie a distributed workflow engine [32].

6.4.2. Template workflows

Participants: Christine Morin, Cédric Tedeschi.

In the framework of the DALHIS associate team [5], we plan to combine the high-level template workflow language TIGRES [6], developed by our partner team from Lawrence Berkeley National Lab (LBL) with the workflow management system developed in the team [17]. This work started with the development of a parser of TIGRES.

6.5. Experimental Platforms

Participants: Alexandra Carpen-Amarie, Maxence Dunnewind, Nicolas Lebreton, Julien Lefeuvre, David Margery, Eric Poupart.

[5] http://project.inria.fr/dalhis
[6] http://tigres.lbl.gov/home

6.5.1. Energy measurement

Participants: David Margery, Maxence Dunnewind, Nicolas Lebreton.

In the context of the ECO_2Clouds project, the BonFIRE infrastructure was updated. At the hardware level power distribution units that report electricity usage for each outlet were installed. At the software layer, a probe reporting energy sources used was configured. This probe gets its information from RTE, the French Electricity transport network, and allows publication of CO_2 metrics for each machine in the testbed. Moreover, access to these metrics was abstracted through the general API to access BonFIRE.

6.5.2. Deployment of IaaS management system

Participant: Alexandra Carpen-Amarie.

The Grid'5000 platform has become one of the most complete testbeds for designing or evaluating large-scale distributed systems, playing an essential role in enabling experimental research at all levels of the Cloud Computing stack and providing configurable cloud platforms similar to commercially available clouds.

However, the complexity of managing the deployment and tuning of large-scale private clouds emerged as a major drawback. Typically, users study specific cloud components or carry out experiments involving applications running in cloud environments. A key requirement in this context is seamless access to ready-to-use cloud platforms, as well as full control of the deployment settings.

To address these needs, we developed a set of deployment tools for open-source IaaS environments, capable of installing and tuning fully-functional clouds on the Grid'5000 testbed [20]. The deployment tools support four widely-used IaaS clouds, namely OpenNebula, CloudStack, Nimbus and OpenStack.

They rely on the concept of extensible engines for defining experiments. Such engines implement all the stages of an experiment: physical node reservations in Grid'5000, environment deployment, configuration and experiment execution. We designed generic engines for nodes reservation and deployment according to a set of requirements specified in a cloud configuration file. Thus, these engines do not require any prior knowledge of lower-level Grid'5000 tools, allowing the user to easily achieve multi-site Grid'5000 deployments based on multiple environments.

6.5.3. BonFIRE

Participants: Maxence Dunnewind, David Margery, Eric Poupart.

The project was reviewed in December 2013 during CloudCom 2013 in Bristol and rated Excellent. The main achievement this year is the introduction of a reservation system for resources on the BonFIRE platform.

6.5.4. Fed4FIRE

Participants: Julien Lefeuvre, Nicolas Lebreton, David Margery.

In Fed4FIRE, two key technologies have been adopted as common protocols to enable experimenter to interact with testbeds. SFA, to provision resources, and OMF to control them. Here, we contributed to a proposal to secure usage of OMF and to a design to allow using BonFIRE through SFA.

NACHOS Project-Team

6. New Results

6.1. Discontinuous Galerkin methods for Maxwell's equations

6.1.1. DGTD-$\mathbb{P}_p$ method based on hierarchical polynomial interpolation

Participants: Loula Fezoui, Stéphane Lanteri.

The DGTD (Discontinuous Galerkin Time Domain) method originally proposed by the team for the solution of the time domain Maxwell's equations [16] relies on an arbitrary high order polynomial interpolation of the component of the electromagnetic field, and its computer implementation makes use of nodal (Lagrange) basis expansions on simplicial elements. The resulting method is often denoted by DGTD-$\mathbb{P}_p$ where p refers to the interpolation degree that can be defined locally i.e. at the element level. In view of the design of a hp-adaptive DGTD method, i.e. a solution strategy allowing an automatic adaptation of the interpolation degree p and the discretization step h, we now investigate alternative polynomial interpolation and in particular those which lead to hierarchical or/and orthogonal basis expansions. Such basis expansions on simplicial elements have been extensively studied in the context of continuous finite element formulations (e.g. [59]) and have thus been designed with global conformity requirements (i.e. H_1, $H(\mathrm{rot})$ or (div)) whose role in the context of a discontinuous Galerkin formulation has to be clarified. This represents one of the objectives of this study. This year, we have started the development of a new software platform in Fortran 95 implementing DGTD-$\mathbb{P}_p$ able to deal with different polynomial basis expansions on a tetrahedral element, for the solution of the 3D time domain Maxwell equations.

6.1.2. DGTD-$\mathbb{P}_p\mathbb{Q}_k$ method on multi-element meshes

Participants: Clément Durochat, Stéphane Lanteri, Raphael Léger, Claire Scheid, Mark Loriot [Distene, Pôle Teratec, Bruyères-le-Chatel].

In this work, we study a multi-element DGTD method formulated on a hybrid mesh which combines a structured (orthogonal) discretization of the regular zones of the computational domain with an unstructured discretization of the irregularly shaped objects. The general objective is to enhance the flexibility and the efficiency of DGTD methods for large-scale time domain electromagnetic wave propagation problems with regards to the discretization process of complex propagation scenes. With this objective in mind, we have designed and analyzed a DGTD-$\mathbb{P}_p\mathbb{Q}_k$ method formulated on non-conforming hybrid quadrangular/triangular meshes (2D case) or non-conforming hexahedral/tetrahedral meshes (3D case) for the solution of the time domain Maxwell's equations [23]-[22].

6.1.3. DGTD-$\mathbb{P}_p$ method for Debye media and applications to biolectromagnetics

Participants: Claire Scheid, Maciej Klemm [Communication Systems & Networks Laboratory, Centre for Communications Research, University of Bristol, UK], Stéphane Lanteri.

This work is undertaken in the context of a collaboration with the Communication Systems & Networks Laboratory, Centre for Communications Research, University of Bristol (UK). This laboratory is studying imaging modalities based on microwaves with applications to dynamic imaging of the brain activity (Dynamic Microwave Imaging) on one hand, and to cancerology (imaging of breast tumors) on the other hand. The design of imaging systems for these applications is extensively based on computer simulation, in particular to assess the performances of the antenna arrays which are at the heart of these systems. In practice, one has to model the propagation of electromagnetic waves emitted from complex sources and which propagate and interact with biological tissues. In relation with these issues, we study the extension of the DGTD-$\mathbb{P}_p$ method originally proposed in [16] to the numerical treatment of electromagnetic wave propagation in dispersive media. We consider an approach based on an auxiliary differential equation modeling the time evolution of the electric polarization for a dispersive medium of Debye type (other dispersive media will be considered subsequently). The stability and a priori convergence analysis of the resulting DGTD-$\mathbb{P}_p$ method has been studied [25], and its application to the simulation of the propagation in realistic geometrical models of head tissues is underway in the context of our participation to the DEEP-ER FP7 project.

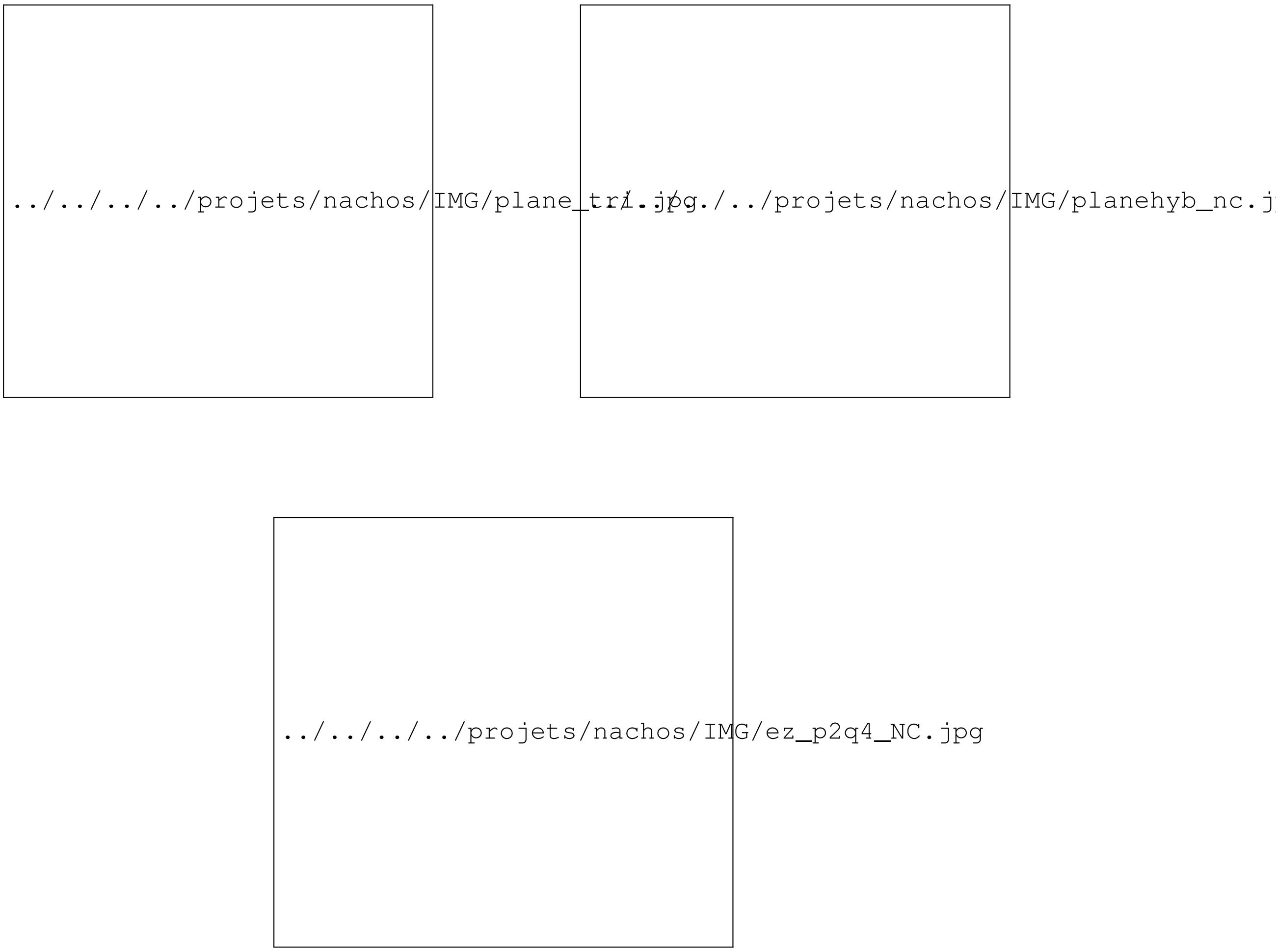


Figure 4. Scattering of a plane wave by a disk. Conforming triangular mesh (top left) and non-conforming quadrangular/triangular mesh (top right). Contour lines of electrical field component E_z from a simulation with a DGTD-$\mathbb{P}_2\mathbb{Q}_4$ method (bottom).

6.1.4. DGTD-$\mathbb{P}_p$ method for nanophotonics

Participants: Claire Scheid, Maciej Klemm [Communication Systems & Networks Laboratory, Centre for Communications Research, University of Bristol, UK], Stéphane Lanteri, Raphael Léger, Jonathan Viquerat.

Modelling and numerical simulation aspects are crucial for a better understanding of nanophotonics. Media that one encounters are complex and the geometries quite involved, so that while a FDTD method failed to be accurate enough, a non conforming discretisation method seems to be well adapted. In this direction, since the end of 2012, we are actively studying the numerical modeling of electromagnetic wave interaction with nanoscale metallic structures. In this context, one has to take into account the dispersive characteristics of metals in the frequency range of interest to nanophotonics. As a first step in this direction, we have considered an auxiliary differential equation approach for the numerical treatment of a Drude, Drude-Lorentz and a generalized dispersion models in the framework of a DGTD-$\mathbb{P}_p$ method [20]-[36]. We performed the corresponding numerical analysis as well as numerical validation tests cases. Some methodological improvements, such as curvilinear elements and higher order time discretization schemes are also underway.

6.1.5. Frequency domain hybridized DGFD-$\mathbb{P}_p$ methods

Participants: Stéphane Lanteri, Liang Li [Faculty Member, School of Mathematical Sciences, Institute of Computational Science, University of Electronic Science and Technology of China Chengdu, China], Ronan Perrussel [Laplace Laboratory, INP/ENSEEIHT/UPS, Toulouse].

For certain types of problems, a time harmonic evolution can be assumed leading to the formulation of the frequency domain Maxwell equations, and solving these equations may be more efficient than considering the time domain variant. We are studying a high order Discontinuous Galerkin Frequency Domain (DGFD-$\mathbb{P}_p$) method formulated on unstructured meshes for solving the 2D and 3D time harmonic Maxwell equations. However, one major drawback of DG methods is their intrinsic cost due to the very large number of globally coupled degrees of freedom as compared to classical high order conforming finite element methods. Different attempts have been made in the recent past to improve this situation and one promising strategy has been recently proposed by Cockburn *et al.* [48] in the form of so-called hybridizable DG formulations. The distinctive feature of these methods is that the only globally coupled degrees of freedom are those of an approximation of the solution defined only on the boundaries of the elements. This work is concerned with the study of such Hybridizable Discontinuous Galerkin (HDG) methods for the solution of the system of Maxwell equations in the time domain when the time integration relies on an implicit scheme, or in the frequency domain. We have been one of the first groups to study HDGFD-$\mathbb{P}_p$ methods based on nodal interpolation methods for the solution of the 2D and 3D frequency domain Maxwell equations [26]-[27].

6.1.6. Exact transparent condition in a DGFD-$\mathbb{P}_p$ method

Participants: Mohamed El Bouajaji, Nabil Gmati [ENIT-LAMSIN, Tunisia], Stéphane Lanteri, Jamil Salhi [ENIT-LAMSIN, Tunisia].

In the numerical treatment of propagation problems theoretically posed in unbounded domains, an artificial boundary is introduced on which an absorbing condition is imposed. For the frequency domain Maxwell equations, one generally use the Silver-Müller condition which is a first order approximation of the exact radiation condition. Then, the accuracy of the numerical treatment greatly depends on the position of the artificial boundary with regards to the scattering object. In this work, we have conducted a preliminary study aiming at improving this situation by using an exact transparent condition in place of the Silver-Müller condition. Promising results have been obtained in the 2D case [30].

6.2. Discontinuous Galerkin methods for the elastodynamic equations

6.2.1. DGTD-$\mathbb{P}_p$ method for viscoelastic media

Participants: Nathalie Glinsky, Stéphane Lanteri, Fabien Peyrusse.

We continue developing high order non-dissipative discontinuous Galerkin methods on simplicial meshes for the numerical solution of the first order hyperbolic linear system of elastodynamic equations. These methods share some ingredients of the DGTD-$\mathbb{P}_p$ methods developed by the team for the time domain Maxwell equations among which, the use of nodal polynomial (Lagrange type) basis functions, a second order leap-frog time integration scheme and a centered scheme for the evaluation of the numerical flux at the interface between neighboring elements. The resulting DGTD-$\mathbb{P}_p$ methods have been validated and evaluated in detail in the context of propagation problems in both homogeneous and heterogeneous media including problems for which analytical solutions can be computed. Particular attention was given to the study of the mathematical properties of these schemes such as stability, convergence and numerical dispersion.

A recent novel contribution is the extension of the DGTD method to include viscoelastic attenuation. For this, the velocity-stress first-order hyperbolic system is completed by additional equations for the anelastic functions including the strain history of the material. These additional equations result from the rheological model of the generalized Maxwell body and permit the incorporation of realistic attenuation properties of viscoelastic material accounting for the behaviour of elastic solids and viscous fluids. In practice, we need solving 3L additional equations in 2D (and 6L in 3D), where L is the number of relaxation mechanisms of the generalized Maxwell body. This method has been implemented in 2D and validated by comparison to results obtained by a finite-difference method, in particular for wave propagation in a realistic basin of the area of Nice (south of France)

6.2.2. DGTD-$\mathbb{P}_p$ method for the assessment of topographic effects

Participants: Etienne Bertrand [CETE Méditerranée], Nathalie Glinsky.

This study addresses the numerical assessment of site effects especially topographic effects. The study of measurements and experimental records proved that seismic waves can be amplified at some particular locations of a topography. Numerical simulations are exploited here to understand further and explain this phenomenon. The DGTD-$\mathbb{P}_p$ method has been applied to a realistic topography of Rognes area (where the Provence earthquake occured in 1909) to model the observed amplification and the associated frequency. Moreover, the results obtained on several homogeneous and heterogeneous configurations prove the influence of the medium in-depth geometry on the amplifications measures at the surface .

6.2.3. DGTD-$\mathbb{P}_p$ method for arbitrary heterogeneous media

Participants: Nathalie Glinsky, Diego Mercerat [CETE Méditerranée].

We have recently devised an extension of the DGTD method for elastic wave propagation in arbitrary heterogeneous media. In realistic geological media (sedimentary basins for example), one has to include strong variations in the material properties. Then, the classical hypothesis that these properties are constant within each element of the mesh can be a severe limitation of the method, since we need to discretize the medium with very fine meshes resulting in very small time steps. For these reasons, we propose an improvement of the DGTD method allowing non-constant material properties within the mesh elements. A change of variables on the stress components allows writing the elastodynamic system in a pseudo-conservative form. Then, the introduction of non-constant material properties inside an element is simply treated by the calculation, via convenient quadrature formulae, of a modified local mass matrix depending on these properties. This new extension has been validated for a smoothly varying medium or a strong jump between two media, which can be accurately approximated by the method, independently of the mesh .

6.2.4. DGFD-$\mathbb{P}_p$ method for frequency domain elastodynamics

Participants: Hélène Barucq [MAGIQUE3D project-team, Inria Bordeaux - Sud-Ouest], Marie Bonnasse, Julien Diaz [MAGIQUE3D project-team, Inria Bordeaux - Sud-Ouest], Stéphane Lanteri.

We have started this year a research direction aiming at the development of high order discontinuous Galerkin methods on unstructured meshes for the simulation of frequency domain elastodynamic and viscelastic wave propagation. This study is part of the Depth Imaging Partnership (DIP) between Inria and TOTAL. The PhD thesis of Marie Bonnasse is at the heart of this study which is funded by TOTAL.

6.3. Multiscale finite element methods for time-domain wave models

Participants: Marie-Helene Lallemand Tenkes, Stéphane Lanteri, Claire Scheid, Frédéric Valentin [LNCC, Petrópolis, Brazil].

Mathematical (partial differential equation) models embedding multiscale features occur in a wide range of natural situations and industrial applications involving wave propagation. This is for instance the case of electromagnetic or seismic wave propagation in heterogenous media. Although the related applications take place at the macro-scale, it is well known that the parameters describing the macro-scale processes are eventually determined by the solution behavior at the micro-sacle. As a result, each stage of the modeling of the underlying problem is driven by distinct sets of PDEs with highly heterogeneous coefficients and embedded high-contrast interfaces. Because of the huge difference in physical scales in heterogenous media it is not computationally feasible to fully resolve the micro-scale features directly. Macroscopic models or upscaling techniques have therefore to be developed that are able to accurately capture the macroscopic behavior while significantly reducing the computational cost. In this context, researchers at LNCC have recently proposed a new family of finite element methods [51]- [50], called Multiscale Hybrid-Mixed methods (MHM), which is particularly adapted to be used in high-contrast or heterogeneous coefficients problems. Particularly, they constructed a family of novel finite element methods sharing the following properties: (i) stable and high-order convergent; (ii) accurate on coarse meshes; (iii) naturally adapted to high-performance parallel computing; (iv) induce a face-based a posteriori error estimator (to drive mesh adaptativity); (v) locally conservative. We have started this year a new reserach direction aiming at the design of similar MHM methods for solving PDE models of time-domain electromagnetic and seismic wave propagation.

6.4. Time integration strategies and resolution algorithms

6.4.1. Hybrid explicit-implicit DGTD-$\mathbb{P}_p$ method

Participants: Stéphane Descombes, Stéphane Lanteri, Ludovic Moya.

Existing numerical methods for the solution of the time domain Maxwell equations often rely on explicit time integration schemes and are therefore constrained by a stability condition that can be very restrictive on highly refined meshes. An implicit time integration scheme is a natural way to obtain a time domain method which is unconditionally stable. Starting from the explicit, non-dissipative, DGTD-$\mathbb{P}_p$ method introduced in [16], we have proposed the use of Crank-Nicolson scheme in place of the explicit leap-frog scheme adopted in this method [5]. As a result, we obtain an unconditionally stable, non-dissipative, implicit DGTD-$\mathbb{P}_p$ method, but at the expense of the inversion of a global linear system at each time step, thus obliterating one of the attractive features of discontinuous Galerkin formulations. A more viable approach for 3D simulations consists in applying an implicit time integration scheme locally i.e in the refined regions of the mesh, while preserving an explicit time scheme in the complementary part, resulting in an hybrid explicit-implicit (or locally implicit) time integration strategy. In [7], we conducted a preliminary numerical study of a hyrbid explicit-implicit DGTD-$\mathbb{P}_p$ method, combining a leap-frog scheme and a Crank-Nicolson scheme, and obtained promising results. More recently, we further investigated two such strategies, both theoretically (especially, convergence in the ODE and PDE senses) [17] and numerically in the 2D case [28]. A last topic is to propose higher order time integration techniques based on the second-order locally implicit method to fully exploit the attractive features of this approach combined with a DG discretisation which allows to easily increase the spatial convergence order. Promising results in 2D reaching high order in time, between $3, 5$ and 4, have been obtained in [29] by applying Richardson extrapolation and composition methods.

6.4.2. Optimized Schwarz algorithms for the frequency domain Maxwell equations

Participants: Victorita Dolean, Martin Gander [Mathematics Section, University of Geneva], Stéphane Lanteri, Ronan Perrussel [Laplace Laboratory, INP/ENSEEIHT/UPS, Toulouse].

../../../../projets/nachos/IMG/Ez2.jpg ../../../../projets/nachos/IMG/nac

Figure 5. Scattering of a plane wave by an airfoil profile. Contour lines of electrical field component E_z (left) and locally refined triangular mesh with partitioning in explicit/implicit zones (right).

Even if they have been introduced for the first time two centuries ago, over the last two decades, classical Schwarz methods have regained a lot of popularity with the developement of parallel computers. First developed for the elliptic problems, they have been recently extended to systems of hyperbolic partial differential equations, and it was observed that the classical Schwartz method can be convergent even without overlap in certain cases. This is in strong contrast to the behavior of classical Schwarz methods applied to elliptic problems, for which overlap is essential for convergence. Over the last decade, optimized versions of Schwarz methods have been developed for elliptic partial differential equations. These methods use more effective transmission conditions between subdomains, and are also convergent without overlap for elliptic problems. The extension of such methods to systems of equations and more precisely to Maxwell's system (time harmonic and time discretized equations) has been studied in [9]. The optimized interface conditions proposed in [9] were devised for the case of non-conducting propagation media. We have recently studied the formulation of such conditions for conducting media [4]. Besides, we have also proposed an appropriate discretization strategy of these optimized Schwarz algorithms in the context of a high order DGFD-$\mathbb{P}_p$ method formulated on unstructured triangular meshes for the solution of the 2D frequency domain Maxwell equations [42].

../../../../projets/nachos/IMG/decomp3b.jpg

../../../../projets/nachos/IMG/DDdifcarre11.jpg

../../../../projets/nachos/IMG/convPk.jpg

Figure 6. Propagation of a plane wave in a multilayered heterogeneous medium. Problem setting and two-subdomain decompositin (top). Contour lines of the real part of the E_z component of the electrical field (bottom left) and asymptotic convergence of the optimized Schwarz algorithms (bottom right).

NANO-D Team

5. New Results

5.1. Adaptively Restrained Particle Simulations for Isobaric-Isothermal Ensemble

Participants: Zofia Trstanova, Stephane Redon.

We continued working on the Adaptively Restrained Particles Simulations (ARPS) approach that was proposed by Svetlana Artemova and Stephane Redon [11] and that was designed to speed up the particles simulations by switching on and off the degrees of freedom based on the kinetic energy of the particle.

It has been shown, for the NVE and the NVT ensemble, that this method has many advantages [11]. We want to extend ARPS for the isobaric-isothermal ensemble (NPT) since this ensemble is very often used in particle simulations, because many chemical reactions happen under constant pressure. An adaptive method for this ensemble with advantages of ARPS might be very useful in many scientific domains (physics, biology, chemistry).

We combined the ARPS method with an existing method that describes the NPT ensemble. We already obtained very promising analytical and numerical results that support the main characteristically advantages of ARPS shown by Svetlana Artemova and Stephane Redon. For instance, Figure 4 shows preservation of the radial distribution function.

5.2. Interactive large-scale deformations of molecular structures

Participants: Jelmer Wolterink, Himani Singhal, Marc Piuzzi, Stephane Redon.

We have developed new interaction methods for large-scale deformation of molecular structures. These new methods allow a user to attach control points to molecules, and use these control points to easily deform the structures while preserving their realism (e.g. local interactions, etc.). The new methods may be applied to any type of molecule (e.g. proteins, carbon nanotubes, etc.), and may be used in combination with interactive simulation.

5.3. Towards parallel adaptive molecular simulations

Participants: Krishna Kant Singh, Benjamin Bouvier, Jean-Francois Mehaut, Stephane Redon.

The adaptive algorithms that we are developing have two main components. The first component determines when and how degrees of freedom can be deactivated and reactivated during a simulation. The second component takes advantage of the frozen degrees of freedom to accelerate the calculation of the potential energy and interatomic forces. Indeed, the potential energy and forces can often be expressed as a (potentially complex) sum of terms which only depend on relative atomic positions. When the relative positions do not change, it is not necessary to update the corresponding terms, which reduces the computation time. We have shown that it is possible to significantly speed up simulations using this approach, while being able to recover static equilibrium statistics [11].

We have now begun to study the possibility of developing adaptive *parallel* simulation algorithms, and have begun to review and benchmark popular simulation packages (GROMACS, NAMD, OpenMM, etc.), depending on the number of atoms, the number of available cores, etc.

5.4. Protein secondary structure prediction for dynamic simulations

Participants: Marc Piuzzi, Sergei Grudinin, Stephane Redon.

../../../../projets/nano-d/IMG/RDF.png

Figure 4. Radial distribution function obtained with different ARPS simulations compared to the full dynamics (blue dash line)

There is a tight link between a protein's function and its molecular structure. Hence, global stability is essential for a protein to keep its role inside the cell. Various chemical interactions helps stabilizing the structure (covalent bonds, hydrogen bonds, etc.) but not all parts of a protein present the same stability. The most stable regions of a protein present numerous hydrogen bonds on backbone atoms composing geometrically distinguishable secondary structures (the primary structure being the amino acid sequence): helices and beta sheets.

These structures have been well studied and although important properties have been defined, there is no absolute definition of what is a helix or a beta sheet. Thus, various methods have been developed to predict the secondary structure of a protein using the amino acid sequence and/or the protein structure using different parameters and structural descriptors.

However, none of these methods have been made in the context of interactive simulation where the shape of the protein is dynamic: here the prediction has to be done at each time step on the whole protein. Moreover, the result is deterministic and returns only the type of structure without any information about the accuracy. We are developing a new approach that is appropriate in an interactive context, where secondary structure assignment has to continuously change during interaction.

5.5. Motion Planning for Quasi-Static Simulation

Participants: Leonard Jaillet, Stephane Redon.

Recently, motion planning methods inspired from Robotics have been applied to the study of biological molecular systems [8]. These approaches rely on compact graph representations that aim to capture large amplitude motions more efficiently than classic simulation techniques, despite their lower resolution.

We developed within the SAMSON's architecture a new motion planning strategy to perform quasi-static simulation at the nano-scale.

The user provides as inputs the initial and final state of the system he or she wants to simulate. Then, the method searches a transition path that follows the low-energy valleys of the conformational landscape (see figure 5).

The adaptation of motion planning approaches to quasi-static simulation at the nano-scale comes with several challenges. First, these approaches must be adapted to tackle the high dimensionality involved in the case of nanosystems, dimensionality that is directly related to the number of atoms considered. Second, these approaches must be extended to face the complexity of the underlying physics that comes from the various types of interactions between atoms.

The method we propose is able to perform simulations involving bonds breaking. This is, up to our knowledge, the first motion planning approach able to simulate chemical reactions.

5.6. Molecular Modeling

5.6.1. Rapid determination of RMSDs corresponding to macromolecular rigid body motions

Participants: Petr Popov, Sergei Grudinin.

Finding the root mean sum of squared deviations (RMSDs) between two coordinate vectors that correspond to the rigid body motion of a macromolecule is an important problem in structural bioinformatics, computational chemistry and molecular modeling. Standard algorithms compute the RMSD with time proportional to the number of atoms in the molecule. We developed *RigidRMSD*, a new algorithm that determines a set of RMSDs corresponding to a set of rigid body motions of a macromolecule in constant time with respect to the number of atoms in a molecule. Our algorithm is particularly useful for rigid body modeling applications such as rigid body docking, and also for high-throughput analysis of rigid body modeling and simulation results. A C++ implementation of our algorithm will be available at http://nano-d.inrialpes.fr/software/RigidRMSD.

../../../../projets/nano-d/IMG/Methane.png

Figure 5. Snapshots of the transition path obtained with our motion planning simulation method. It represents a chemical reaction where two molecules of methanes interact to form a an ethane and a dihydrogen.

To demonstrate the efficiency of the RigidRMSD library, we compared the clustering application implemented with our algorithm to the one from the Hex software. We chose Hex for the comparison because it is a very fast rigid body docking tool and also because it explicitly provides the clustering time. For the comparison, we collected a small benchmark of 23 protein dimers of various size. After, we launched Hex version 6.3 on this benchmark and collected docking solutions before clustering, sizes of clusters, and clustering time. We then also clustered these solutions using the *RigidRMSD* library. Figure 6 shows the clustering time of the HEX clustering algorithm with respect to our clustering using two rotation representations as a function of the number of atoms in the smaller protein (left) and the number of docking solutions before the clustering (right). We can clearly see that our implementation of the clustering algorithm is more than an order of magnitude faster compared to the Hex implementation. Also, the quaternion representation of rotation is on average three times more efficient compared to the matrix representation.

../../../../projets/nano-d/IMG/Graph-01-log.png ../../../../projets/nano-d/IMG/Graph-02-log.png

Figure 6. Left: Time spent on clustering by Hex and RigidRMSD with respect to the number of atoms in the ligand protein. Number of considered solutions and the RMSD threshold was fixed to 10,000 and 10.0 Å, respectively. Right: Average time spent on clustering by Hex and RigidRMSD with respect to the number of docking solutions. For this plot we chose five structures with the number of atoms of about 2,000 and plotted the standard deviation of the running time. For both plots, the RMSD threshold was fixed to 10.0 Å.

5.6.2. Fast fitting atomic structures into a low-resolution density map using 3D orthogonal Hermite functions

Participants: Georgy Derevyanko, Sergei Grudinin.

We developed a new algorithm for fitting protein structures into a low resolution electron density (e.g. cryo-electron microscopy) map. The algorithm uses 3D orthogonal Hermite functions for fast operations on the electron density.

Orthogonal Hermite function of order n is defined as:

$$\psi_n(x;\lambda) = \frac{\sqrt{\lambda}}{\sqrt{2^n n! \sqrt{\pi}}} \exp\left(-\frac{\lambda^2 x^2}{2}\right) H_n(\lambda x), \tag{19}$$

where $H_n(x)$ is the Hermite polynomial and λ is the scaling parameter. In Fig. 7 we show several orthogonal Hermite functions of different orders with different parameters λ. These functions form an orthonormal basis set in $L^2(\mathbb{R})$. A 1D function $f(x)$ decomposed into the set of 1D Hermite functions up to an order N reads

$$f(x) = \sum_{i=0}^{N} \widehat{f}_i \psi_i(x; \lambda) \tag{20}$$

Here, $\widehat{f}_i$ are the decomposition coefficients, which can be determined from the orthogonality of the basis functions $\psi_i(x;\lambda)$. Decomposition in Eq. 2 is called the *band-limited decomposition* with $\psi_i(x;\lambda)$ basis functions. To decompose the electron density map and the protein structures, we employ the 3D Hermite functions:

$$\psi_{n,l,m}(x,y,z;\lambda) = \psi_n(x;\lambda)\psi_l(y;\lambda)\psi_m(z;\lambda), \tag{21}$$

which form an orthonormal basis set in $L^2\left(\mathbb{R}^3\right)$. A function $f(x,y,z)$ represented as a band-limited expansion in this basis reads

$$f(x,y,z) = \sum_{i=0}^{N} \sum_{j=0}^{N-i} \sum_{k=0}^{N-i-j} \widehat{f}_{i,j,k} \psi_{i,j,k}(x,y,z;\lambda) \tag{22}$$

Our algorithm accelerates rotation of the Fourier image of the electron density by using 3D orthogonal Hermite functions. As a part of the new method, we presented an algorithm for the rotation of the density in the Hermite basis and an algorithm for the conversion of the expansion coefficients into the Fourier basis. We implemented the program of fitting a protein structure to a low-resolution electron density map, which uses the cross-correlation or the Laplacian-filtered cross-correlation as the fitting criterion. We demonstrated that in the Hermite basis, the Laplacian filter has a particularly simple form. To assess the quality of density encoding in the Hermite basis, we uses two measures, the R-factor and the cross-correlation factor. Finally, we validated our algorithm using two examples and compare its efficiency with two widely used fitting methods, ADP_EM and *colores* from the Situs package.

5.6.3. Fast Rotational-Translation Matching of Rigid Bodies by Fast Fourier Transform Acceleration of Six Degrees of Freedom

Participants: Alexandre Hoffmann, Sergei Grudinin.

We introduced a new method for rigid molecular fitting. This problem is usually solved by maximizing the Cross Corelation Function (CCF), which is computed using the Fast Fourier Transform (FFT) algorithm. Our method handles six degrees of freedom at once and requires only one computation of the Cross Corelation Function, with the six-dimensional Fast Fourier Transform. Our method only requires a low pre-processing time ($O(N^7)$), which is comparable to the cost of the subsequent 6D FFT ($O(N^6 \log(N^6))$). It also uses a dual Hermite-Fourier representation, which allows to represent a small molecule with a fewer number of coefficients in the Hermite basis.

5.6.4. Prediction of complexes with point group symmetry using spherical polar Fourier docking correlations

Participants: David W. Ritchie, Sergei Grudinin.

Many proteins form symmetric homo-oligomers that perform a certain physiological function. We present the first point group symmetry docking algorithm that generates perfectly symmetrical protein complexes for arbitrary point group symmetry types (C_n, D_n, T, O, and I). We validate the algorithm on proteins from the 3D-Complex database, where it achieves on average the success rate of 55%. The running time of the algorithm is less than a minute on a modern workstation.

../../../../projets/nano-d/IMG/hermitePlotNLambda_01.png

Figure 7. Left: 1D Hermite functions of order 6 for three different scaling parameters λ. Right: 1D Hermite functions of two different orders for the scaling parameter $\lambda = 1$.

Many of the protein complexes in the protein Data bank (PDB) are symmetric homo-oligomers. According to the 3D-Complex database, C_2 homo-dimers comprise the majority of known homo-oligomers. However, many complexes have higher order rotational symmetry (i.e. C_n>2), and a good number have multiple rotational symmetry axes, namely those with dihedral (D_n), tetrahedral (T), octahedral (O), and icosahedral (I) point group symmetries. Although symmetrical complexes are often solved directly by X-ray crystallography, it would still be very useful to be able to predict whether or not a given monomer might self-assemble into a symmetrical structure. We present a new point group symmetry docking algorithm. In the last few years, several protein-protein docking programs have been adapted to predict symmetrical pair-wise docking orientations for C_n and D_n symmetries. However, to our knowledge, there does not yet exist an algorithm which can automatically generate perfectly symmetrical protein complexes for arbitrary point group symmetry types.

We introduce the notion of a "docking equation" in which the notation $A(\underline{x}) \longleftrightarrow B(\underline{x})$ represents an interaction between proteins A and B in 3D space. It is also useful introduce the operators $\widehat{T}(x,y,z)$ and $\widehat{R}(\alpha,\beta,\gamma)$, which represent the actions of translating an object by an amount (x,y,z) and rotating it according to the three Euler rotation angles (α,β,γ). Then, guided by Figure 8 , and assuming that we start with two identical monomers at the origin, we can write down a C_n docking equation for the two monomers as

$$\widehat{T}(0,y,0)\widehat{R}(\alpha,\beta,\gamma)A(\underline{x}) \longleftrightarrow \widehat{R}(0,0,\omega)\widehat{T}(0,y,0)\widehat{R}(\alpha,\beta,\gamma)B(\underline{x}). \tag{23}$$

Then, we perform a series of fast Fourier transform (FFT) correlation searches using the Hex spherical polar Fourier docking algorithm to determine the four parameters (y,α,β,γ). For higher symmetries, D_n, T, O, and I, we introduce two more parameters and perform a series of FFT in a similar way. The calculation for each structure takes less than a minute on a modern workstation.

../../../../projets/nano-d/IMG/D3.png

Figure 8. Illustrations of the C_3 and D_3 point group symmetries.

We validated our method on protein structures from the 3D-Complex database, which contains 17,183 protein complexes with assigned biological unit and symmetry type. It mostly contains cyclic and dihedral proteins, as well as 86 tetrahedral, 47 octahedral, and 6 icosahedral complexes (excluding all viral structures). Starting from the structures of monomers, we generated symmetric biological units based on the symmetry type for each complex provided by 3D-Complex. Figure 9 summarizes the performance of our method on these proteins, where we say that the model is correct if all pair-wise RMSDs are smaller than 10 Ångstroms. On average, we found about 55% of correct predictions ranked first.

../../../../projets/nano-d/IMG/3DSIG-plot2.png

Figure 9. Summary of the correctly predicted complexes found on the first place (blue) and in the top ten solutions (green).

Figure 10 shows correctly predicted examples from each of the symmetry types. Each complex is perfectly symmetrical, although due to the soft docking function in Hex it is possible that some interfaces might contain minor steric clashes.

5.7. Software Engineering

5.7.1. SAMSON User interface

Participants: Jocelyn Gate, Maria Werewka, Stephane Redon.

We have continued the development of SAMSON, our open-architecture platform for modeling and simulation of nanosystems (SAMSON: Software for Adaptive Modeling and Simulation Of Nanosystems):

- We have moved to Qt5 to handle the Graphical User Interface.

../../../../projets/nano-d/IMG/3DSIG-plot3.png

Figure 10. Illustrations of the correctly predicted complexes. For each complex, the group symbol and the PDB code are shown.

- We now compile SAMSON is 64 bits only. This removes limitations of 32 bits applications, in particular concerning memory limits.
- We now have a complete installer mechanism for both users and developers.
- We extended the set of development tools (action generators, UUID generators, etc.)
- We have changed the windowing system to allow windows to move outside SAMSON.
- We have designed a coherent style for icons, windows, menus, etc.
- We have added 3D rendering.

There are now more than 40 modules in SAMSON (parsers, editors, models, apps, etc.).

The current user interface of SAMSON is visible in Figure 11 .

5.7.2. SAMSON Elements

Participants: Svetlana Artemova, Stephane Redon.

We have added new SAMSON Elements (modules).

We have been working on input and output for SAMSON. Precisely, we now have the possibility to download molecules to SAMSON and save them to external files in three possible formats:

- pdb (Protein Data Bank format, containing experimentally determined 3d structures and widely used for applications in biology);
- mol2 (Sybyl chemical modeler input file, containing chemical compounds and small ligands);
- xyz (basic format, containing atoms coordinates).

Basic properties of atoms, residues, and molecules have been determined and structures storing these properties were implemented.

Finally, since energy minimization is crucial for providing physically-correct structures while interactively editing molecules in SAMSON, we have implemented several fast and stable algorithms to perform such energy minimization in SAMSON.

5.7.3. SAMSON Website

Participants: Mohamed Yengui, Jocelyn Gate, Stephane Redon.

We are developing a web application aiming at distributing and valorizing SAMSON and SAMSON Elements (modules). The goal of the website is to develop a community of users and developers in all areas of nanoscience (physics, biology, chemistry, electronics, etc.). The website will:

- allow users and developers to create and manage accounts on the website.
- allow visualizing, searching and downloading SAMSON and SAMSON Elements.
- allow the creation, validation and dissemination of SAMSON Elements.
- provide tracking requests for the arrival of new SAMSON Elements or the modification of an existing SAMSON Element.

To achieve this, we have designed the architecture in a way that speeds development effort for a faster product release, while keeping in mind scalability, security and high reliability.

We have also implemented and tested locally the account validation process. A user can now sign up, confirm the registration from the received email and authenticate with the registered account to download SAMSON and SAMSON elements from the website. We will make the site public when we release SAMSON.

../../../../projets/nano-d/IMG/SAMSON-Current.jpg

Figure 11. The current user interface of SAMSON showing an app to download molecules directly from the Protein Data Bank. The data graph on the left shows the hierarchical structure of the structural model.

NECS Project-Team

6. New Results

6.1. Networked systems and graph analysis

6.1.1. Distributed graph-discovery

Participants: A. Kibangou [Contact person], T.-M. D. Tran, F. Garin, A. de Almeida [UFC Brazil].

The availability of information on the communication topology of a wireless sensor network is essential for the design of the estimation algorithms. In the context of distributed self-organized sensor networks, there is no central unit with the knowledge of the network, and the agents must run some distributed network discovery algorithm.

We have studied the problem of estimating the eigenvalues of the Laplacian matrix associated with a graph modeling the interconnections between the nodes of a given network. Our approach is based on properties resulting from the factorization of the average consensus matrix. Indeed, as recently shown [45], the average consensus matrix can be written as a product of Laplacian based consensus matrices whose stepsizes are given by the inverse of the nonzero Laplacian eigenvalues. By distributively solving the factorization of the average consensus matrix, we have shown that the Laplacian eigenvalues can be computed as the inverse of the stepsizes in each estimated factor, where these factors are constrained to be structured as Laplacian based consensus matrices. A constrained optimization problem was formulated and distributed gradient descent methods have been formulated. As formulated, the problem can be viewed as a consensus problem with equality constraints. In contrast to the state-of-the-art, the proposed algorithm does not require great resources in both computation and storage. This algorithm can also be viewed as a way for decentralizing the design of finite-time average consensus protocol recently proposed in the literature.

Laplacian eigenvalues have several interesting properties that can help to study networks, however they cannot uniquely characterize the topology of the network. Therefore, we have directly studied the problem of topology identification in [20]. The considered set-up concerns a collaborative wireless sensor network where nodes locally exchange coded informative data before transmitting the combined data towards a remote fusion center equipped with an antenna array. For this communication scenario, a new blind estimation algorithm was developed for jointly recovering network transmitted data and connection topology at the fusion center. The proposed algorithm is based on a two-stage approach. The first stage is concerned with the estimation of the channel gains linking the nodes to the fusion center antennas. The second stage performs a joint estimation of network data and connection topology matrices by exploiting a constrained (PARALIND) tensor model for the collected data at the fusion center.

Distributed network-discovery algorithms become even more challenging in the case where the algorithm must be anonymous, namely in the case when the agents do not have or do not want to disclose their identifiers (id.s), either for technological reasons (in time-varying self-organized networks, assigning unique identifiers to agents is a challenge) or for privacy concerns. In anonymous networks, even simple tasks such as counting the number of agents are challenging problems. In [24] we have proposed an algorithm for node-counting in anoymous networks. It is based on a graph-constrained LTI system similar to linear consensus, and on system identification: the idea is that the order of the system is the number of agents, and based on local observations each agent tries to identify the order of the system, testing the rank of the Hankel matrix from the output data.

6.1.2. Observability in consensus networks

Participants: A. Kibangou [Contact person], C. Commault [Grenoble INP].

Studying the observability problem of a system consists in answering the question: is it possible, for a given node, to reconstruct the entire network state just from its own measurements and those of its neighbors? Studying observability for arbitrary graphs is particularly a tough task, therefore, studies are generally restricted to some families of graphs; for instance, recently, observability has been studied for paths and circular graphs and also grids where the study was carried out based on rules on number theory. We have considered families of graphs admitting an association scheme such that strongly regular graphs and distance regular graphs. The regularity properties of these kinds of graphs can particularly be useful for robustifying the network as for cryptographic systems. Based on the so-called Bose-Mesner algebra, we have stated observability conditions on consensus networks modeled with graphs modeled with strongly regular graphs and distance regular graphs; for this purpose, we have introduced the notion of local observability bipartite graph that allows characterizing the observability in consensus networks. We have shown that the observability condition is given by the nullity of the so-called "local bipartite observability graph"; when the nullity of the graph cannot be derived directly from the structure of the local bipartite observability graph, the rank of the associated bi-adjacency matrix allows evaluating the observability. The bi-adjacency matrix of the local bipartite observability graph must be full column rank for guaranteeing observability. From this general necessary and sufficient condition, we have deduced sufficient conditions for strongly regular graphs and distance regular graphs [25].

6.2. Collaborative and distributed algorithms

6.2.1. Finite-time average consensus

Participants: A. Kibangou [Contact person], T.-M. D. Tran.

Nowadays, several distributed estimation algorithms are based on the average consensus concept. Average consensus can be reached using a linear iterations scheme where each node repeatedly updates its value as a weighted linear combination of its own value and those of its neighbors; the main benefit of using a linear iterations scheme is that, at each time-step, each node only has to transmit a single value to each of its neighbors. Based on such a scheme, several algorithms have been proposed in the literature; however, in the most proposed algorithms the weights are chosen so that all the nodes asymptotically converge to the same value. Sometimes, consensus can be embedded as a step of more sophisticated distributed; obviously, asymptotic convergence is not suitable for these kinds of distributed methods, and therefore it is interesting to address the question of exact consensus in finite-time. For time-invariant network topologies and in the perfect information exchange case, i.e., without channel noise nor quantization, we have shown that the finite-time average consensus problem can be solved as a matrix factorization problem with joint diagonalizable matrices depending on the graph Laplacian eigenvalues; moreover, the number of iterations is equal to the number of distinct nonzero eigenvalues of the graph Laplacian matrix. The design of such a protocol requires the knowledge of the Laplacian spectrum, which can be carried out in a distributed way (see Section 6.1.1 . The matrix factorization problem is solved in a distributed way, in particular a learning method was proposed and the optimization problem was solved by means of distributed gradient backpropagation algorithms. The factor matrices are not necessarily symmetric and the number of these factor matrices is exactly equal to the diameter of the graph [30].

6.2.2. Linear consensus in large-scale geometric graphs

Participants: F. Garin [Contact person], E. Lovisari [Lund], S. Zampieri [Padova].

Traditional analysis of linear average-consensus algorithms studies, for a given communication graph, the convergence rate, given by the essential spectral radius of the transition matrix (i.e., the second largest eigenvalues' modulus). For many graph families, such analysis predicts a performance which degrades when the number of agents grows, basically because spreading information across a larger graph requires a longer time; however, when considering other well-known quadratic performance indices (involving all the eigenvalues of the transition matrix), the scaling law with respect to the number of agents can be different. This is consistent with the fact that, in many applications, for example in estimation problems, it is natural to expect that a larger number of cooperating agents has a positive, not a negative effect on performance. It is natural to use a different performance measure when the algorithm is used for different purposes, e.g., within a distributed estimation or control algorithm. We are interested in evaluating the effect of the topology of the communication graph on performance, in particular for large-scale graphs. We have focused on graph families which can describe sensor networks, and hence have geometric constraints, namely nodes can be connected only with nearby nodes in the sense of Euclidean distance [16].

6.2.3. Distributed computation methods for multidimensional data

Participants: A. Kibangou [Contact person], A. de Almeida [UFC Brazil].

In [19], we consider the issue of distributed computation of tensor decompositions. A central unit observing a global data tensor assigns different data sub-tensors to several computing nodes grouped into clusters. The goal is to distribute the computation of a tensor decomposition across the different computing nodes of the network, which is particularly useful when dealing with large-scale data tensors. However, this is only possible when the data sub-tensors assigned to each computing node in a cluster satisfies minimum conditions for uniqueness. By allowing collaboration between computing nodes in a cluster, we show that average consensus based estimation is useful to yield unique estimates of the factor matrices of each data sub-tensor. Moreover, an essentially unique reconstruction of the global factor matrices at the central unit is possible by allowing the subtensors assigned to different clusters to overlap in one or several modes. The proposed approach is useful to a number of distributed tensor-based estimation problems in signal and data processing.

6.2.4. Collaborative source seeking

Participants: C. Canudas de Wit [Contact person], R. Fabbiano, F. Garin, Y. Gaudfrin, J. Dumon.

The problem of source localization consists in finding, with one or several agents possibly cooperating with each other, the point or the spatial region from which a quantity of interest is being emitted. Source-seeking agents can be fixed sensors, that collect and exchange some information about the signal field and try to identify the position of the source (or the smallest region in which it is included), or moving devices equipped with one or more sensors, that physically reach the source in an individual or cooperative way. This research area is attracting a rapidly increasing interest, in particular in applications where the agents have limited or no position information and GPS navigation is not available, as in underwater navigation or in cave exploration: for instance, source localization is relevant to many applications of vapor emitting sources such as explosive detection, drug detection, sensing leakage or hazardous chemicals, pollution sensing and environmental studies. Other fields of interest are sound source localization, heat source localization and vent sources in underwater field. Techniques present in literature either are based on a specific knowledge of the solution of the diffusion process, or make use of an extremum-seeking approach, exciting the system with a periodic signal so as to explore the field and collect enough information to reconstruct the gradient of the quantity of interest. Our approach lies in the computation of derivatives (potentially of any order) from Poisson integrals that, for isotropic diffusive source in steady-state, whose solution satisfies the Laplace equation, allows for a gradient search with a small computation load (derivatives are computed by integrals) and without requiring any knowledge of the closed-form solution, avoiding in the same time extremum-seeking oscillations; this has the additional advantage of an intrinsic high-frequency filtering, that makes the method low sensitive to measurement noise. This work is the topic of the Ph.D. of Ruggero Fabbiano, and is described in papers under review.

Moreover, a hardware implementation of the source-seeking algorithm has been done during the internship of Yvan Gaudfrin, at GIPSA-LAB with the support of Jonathan Dumon. A description of the setup and videos of the source-seeking robot are available online: http://necs.inrialpes.fr/pages/platforms.php

6.3. Sensor networks: estimation and data fusion

6.3.1. Data fusion approaches for motion capture by inertial and magnetic sensors

Participants: H. Fourati [Contact person], A. Makni, A. Kibangou.

We are interested to motion capture (or attitude) by fusing Inertial and Magnetic Sensors. In [15], we present a viable quaternion-based Complementary Observer (CO) which is designed for rigid body attitude estimation. The CO processes data from a small inertial/magnetic sensor module containing tri-axial angular rate sensors, accelerometers, and magnetometers, without resorting to GPS data. The proposed algorithm incorporates a motion kinematic model and adopts a two-layer filter architecture. In the latter, the Levenberg Marquardt Algorithm (LMA) pre-processes acceleration and local magnetic field measurements, to produce what will be called the system's output. The system's output together with the angular rate measurements will become measurement signals for the CO. In this way, the overall CO design is greatly simplified. The efficiency of the CO is experimentally investigated through an industrial robot and a commercial IMU during human segment motion exercises. In a recent work [35], a viable quaternion-based Adaptive Kalman Filter (q-AKF) that is designed for rigid body attitude estimation. This approach is an alternative to overcome the limitations of the classical Kalman filter. The q-AKF processes data from a small inertial/magnetic sensor module containing triaxial gyroscopes, accelerometers, and magnetometers. The proposed approach addresses two challenges. The first one concerns attitude estimation during various dynamic conditions, in which external acceleration occurs. Although external acceleration is one of the main source of loss of performance in attitude estimation methods, this problem has not been sufficiently addressed in the literature. An adaptive algorithm compensating external acceleration from the residual in the accelerometer is proposed. At each step, the covariance matrix associated with the external acceleration is estimated to adaptively tune the filter gain. The second challenge is focused on the energy consumption issue of gyroscopes for long-term battery life of Inertial Measurement Units. We study the way to reduce the gyro measurement acquisition while maintaining acceptable attitude estimation. Through numerical simulations, under external acceleration and parsimonious gyroscope's use, the efficiency of the proposed q-AKF is illustrated.

6.3.2. Pedestrian dead-reckoning navigation

Participant: H. Fourati [Contact person].

We proposes a foot-mounted Zero Velocity Update (ZVU) aided Inertial Measurement Unit (IMU) filtering algorithm for pedestrian tracking in indoor environment [22]. The algorithm outputs are the foot kinematic parameters, which include foot orientation, position, velocity, acceleration, and gait phase. The foot motion filtering algorithm incorporates methods for orientation estimation, gait detection, and position estimation. A novel Complementary Filter (CF) is introduced to better pre-process the sensor data from a foot-mounted IMU containing tri-axial angular rate sensors, accelerometers, and magnetometers and to estimate the foot orientation without resorting to GPS data. A gait detection is accomplished using a simple states detector that transitions between states based on acceleration measurements [32]. Once foot orientation is computed, position estimates are obtained by using integrating acceleration and velocity data, which has been corrected at step stance phase for drift using an implemented ZVU algorithm, leading to a position accuracy improvement. We illustrate our findings experimentally by using of a commercial IMU during regular human walking trial in a typical public building. Experiment results show that the positioning approach achieves approximately a position accuracy less than 1 m and improves the performance regarding a previous work of literature [33].

6.3.3. Sensor placement of unreliable sensors

Participants: F. Garin [Contact person], P. Frasca [Twente].

In this work (see [23]), we consider problems in which sensors have to be deployed in a given environment in such a way to provide good coverage of it. It is clear that sensor failures may deteriorate the performance of the resulting sensor network. Then, it is also natural to ask if taking into account such uncertainties changes the coverage optimization problem and leads to a different optimal solution. For simplicity, we start considering a one-dimensional problem, where sensors are to be placed on a line in such a way to optimize the disk-coverage cost The optimal solution for reliable sensors is simply an equally-spaced configuration of the sensors. If we allow that the sensors may fail to take or communicate their measurements, this solution may instead not be optimal. However, as the number of sensors grows to infinity, the ratio between the cost of equally-spaced configurations and the optimal failure-free cost only grows as the logarithm of the number of sensors. We interpret this result as a confirmation of the intrinsic robustness of sensor networks.

6.4. Control design and co-design

6.4.1. Energy-aware networked control

Participants: C. Canudas de Wit [Contact person], F. Garin, N. Cardoso de Castro, D. Quevedo [Newcastle].

We have considered an event-based approach to energy-efficient management of the radio chip in the sensor node of a wireless networked control system. Indeed the radio is the main energy consumer, and intermittent data transmission allows one to reduce the use of the radio. While the existing literature in the control community on event-based control only addresses policies using two radio modes (transmitting/sleep), our work follows some considerations on the radio chip modes well-known in the communication networks literature, and introduces various radio-modes: different 'idle' non-transmitting modes, where only part of the radio chip is switched off (thus consuming more energy than 'sleep', but allowing for faster transition to transmission), and various transmitting modes, with different power levels. We propose an event-based radio-mode switching policy, which allows to perform a trade-off between energy saving and performance of the control application; to this end, a switched model describes the system, taking into account control and communication. The optimal switching policy is computed using dynamic programming, considering a cost either over an infinite time-horizon (see [31]) or over a finite receding horizon (joint work with D. Quevedo, Univ. Newcastle, Australia, described in a paper in preparation).

6.4.2. Adaptive control strategy based reference model for spacecraft motion trajectory

Participants: H. Fourati [Contact person], Z. Samigulina.

In aerospace field, the economic realization of a spacecraft is one of the main objectives which should be accomplished by conceiving the optimal propulsion system and the best control algorithms. Our work focuses on the development of a viable Adaptive Control Approach (ACA) for Spacecraft Motion Trajectory (SMT), see [39].The proposed strategy involves the nonlinear mathematical model of SMT expressed in the central field, which is linearized by the Taylor expansion, and the second Lyapunov method to offer a high rate and unfailing performance in the functioning. The adaptive control system is composed of the cascade of adaptation loop and feedback control loop. When the spacecraft deviates from its reference trajectory model, the ACA acts on the control system to correct this deviation and follow the optimal reference trajectory. Therefore, when the states of the adjustable model are different from its reference values, then the error signal is provided as an input to the adaptation law, which contains the adaptation algorithm. The output will be the state variable feedback control matrix which will be used to calculate the new control law vector. The efficiencies of the linearization procedure and the control approach are theoretically investigated through some realistic simulations and tests under MATLAB. The steady state errors of control between the reference model and the adjustable model of SMT converge to zero. This work is described in [38].

6.5. Transportation networks and vehicular systems

6.5.1. Traffic estimation and prediction

Participants: C. Canudas de Wit [Contact person], A. Kibangou, L. Leon Ojeda, F. Morbidi.

Reconstructing densities in portions of the road links not equipped with sensors constitutes an important task in traffic estimation, forecasting, and control problems. Among many other approaches, model-based observers is one popular technique to build this information. They can also be understood as *virtual sensors* deployed inside of the cells not equipped with *true sensors*. They are used to better track, in real-time, density variations with a fine degree of granularity in the space, as the *virtual cells* can be selected as small as desired.

In [43], a graph constrained-CTM observer was introduced. It allows reconstructing rather accurately the internal states (densities) of a road portion not equipped with sensors. This strategy for real-time density estimation was applied on Grenoble South Ring. In [27], this observer has been associated with an adaptive Kalman filtering approach for traffic prediction in terms of travel time. The adaptive Kalman filtering approach was also been used for predicting input flows in [26].

6.5.2. Traffic control

Participants: C. Canudas de Wit [Contact person], D. Pisarski.

This work has been conducted in two parallel directions, combining steady state analysis and design of an optimal ramp metering controller.

The first direction was to extend the preliminary results presented in the papers [10] and [46]. The goal was to implement the idea of optimal steady state balancing. A relevant software was built up and tested on the model representing the south ring of Grenoble. The results were published in [28]. A comprehensive study of steady state balancing was submitted as a journal paper, under review.

The second direction was to develop a distributed optimal ramp metering controller. This study is motivated by two following facts. The first one is to decentralize and parallelize computation for optimal freeway traffic control problem, and thus to reduce computational complexity. The second one is to reduce the lengths of the communication channels, in order to eliminate the probability of information delay or packet loss. The proposed new control objective provides a uniformly distributed (or balanced) vehicle density such that the usage of freeway (measured by the Total Travel Distance and Total Input Volume) is maximized. This optimal balancing objective is reached by taking a proper state feedback control structure and optimizing the set of gains. Here we imposed distributed condition for both, the feedback structure and the optimization process. We have focused the efforts to design the controller network architecture that is based on the common elements (ramp meter controllers), executing the same computational procedures and applying the control signals based on the same state-feedback structure. This meets a spirit of 'plug and play' (PnP), and is beneficial for both, architecture assembling and component replacement (in case of failure). In order to define the functionality for each of this PnP controller, the analysis on both system controllability and conditions for optimality were carried on. The preliminary work let us to determine the what type of information and upon which communication topology it is required to be sent in order to solve the posed optimization problem. Firstly, the feedback controls for each of the controllers require state information of the section that is controllable for it. In general, each of the controllers demands the state for its closest surrounding sections (downstream and upstream). Secondly, each of the controllers communicates with its closest active neighbors to exchange the information of optimal solution, namely optimal boundary flow or optimal control. We also observed that in any system mode there might be only one inactive controller (the controller that does not have any controllable section) surrounded by two active ones, and thus the maximum required information comes from the two closest neighbors for each of the directions. We noticed also that inactive controllers may serve to convey the information for the active ones, so the communication can be based on a path (or linear) graph. Part of this research was realized in UC Berkeley during the visit of Dominik Pisarski in PATH laboratory.

Traffic control is based on models of traffic, usually the so-called CTM – Cell transmission Model. Some work in the team aims at developing different models, more suitable for control. One such model is based on cells of variable length, as an alternative way to describe the congestion position. This model, proposed in [42], has been refined in the master thesis of Giulio Bontadini, taking into account mass conservation laws.

6.5.3. Vehicle control for disabled people

Participants: C. Canudas de Wit [Contact person], V. Ciarla, J. Dumon, F. Quaine [UJF], V. Cahouet [UJF].

Disabled people face the effort to turn the steering wheel while driving their vehicle. This study, funded by the VolHand project, focuses on the aspect of the assistance during driving maneuver at low speeds (for instance, parking). On common vehicles for healthy people, the system that improves the driver's steering feel in these situations is the power-steering stage, which is mounted at the basis of the steering column and is based on hydraulic technology; the new generation uses an electric motor instead of the hydraulic pump, with more advantages in terms of fuel consumption, better road-feel feedback to the driver and better return-to-center performances of the steering wheel. This work has developed a general methodology to adapt the current technology for disabled people, by introducing additional blocks that can be implemented via software without altering the hardware of the vehicle. In this way, it can be easily exported without additional costs in terms of design and technology for the industrial partner. The methodology has been studied theoretically, joining control aspects with bio-mechanical ones. Moreover, the theoretical study has been tested in laboratory on the hardware-in-the loop setup, using the experimental platform NeCSCar (see Section 5.2). First, a real steering wheel has been linked to a real-time PC-unit and to an electrical motor. A graphical user interface has been implemented to facilitate the access to the software. Then, the last part of the study has been the experimental validation with a tele-operated real vehicle. The vehicle provided the measure of the friction torque to the PC-unit, simulating a real driving situation.

This work is described in [41] and in the Ph.D. thesis [11].

6.5.4. Control of communicating vehicles in urban environment

Participants: C. Canudas de Wit [Contact person], G. de Nunzio.

For a given vehicle there are different ways to travel on a given distance in a given time, associated to different levels of energy consumption; therefore, it is always possible to find an energy-optimal trajectory. Advising the driver via a suitable interface can reduce the energy consumed during the travel, and thus improve the energy efficiency: this is the principle of eco-driving. In urban areas, the optimal trajectory of the vehicle depends on interactions with other vehicles, also on passive signs (panels, priorities, etc.) and active signs (traffic lights); in each case, constraints are imposed on the command (vehicle speed). From the infrastructure perspective, traffic control in urban areas consists in determining the state of traffic signals in order to solve an optimization problem, for example minimizing average travel time of vehicles in the road network. If all vehicles could communicate with one another and with the active infrastructure (traffic lights), we could imagine benefits for each of the two problems which can be considered as a whole: on the one hand, from the vehicles' point of view, more information is available that can be integrated into the online optimization problem; on the other hand, there are new measures and new commands available to control traffic. Indeed, the estimation of the traffic is no longer necessary, as the position and speed of approaching vehicles is known and shared. More importantly, the traffic manager can send instructions to the vehicles. The aim of the research is to evaluate the potential in terms of energy saving and traffic improvement made possible by communicating vehicles. This work is the topic of the Ph.D. thesis of Giovanni De Nunzio, a CIFRE thesis with IFPEN. The paper [21] considers the scenario where vehicle and infrastructure (traffic lights) can communicate, and describes a suitable optimization algorithm that can be run on-board the vehicle so to optimize its energy consumption by avoiding stops and abrupt changes of speed at traffic lights, thanks to the information on upcoming traffic lights on the same road.

NEUROMATHCOMP Project-Team

5. New Results

5.1. Neural Networks as dynamical systems

5.1.1. Dynamics and spike trains statistics in conductance-based Integrate-and-Fire neural networks with chemical and electric synapses

Participants: Bruno Cessac, Rodrigo Cofré.

We investigate the effect of electric synapses (gap junctions) on collective neuronal dynamics and spike statistics in a conductance-based Integrate-and-Fire neural network, driven by a Brownian noise, where conductances depend upon spike history. We compute explicitly the time evolution operator and show that, given the spike-history of the network and the membrane potentials at a given time, the further dynamical evolution can be written in a closed form. We show that spike train statistics is described by a Gibbs distribution whose potential can be approximated with an explicit formula, when the noise is weak. This potential form encompasses existing models for spike trains statistics analysis such as maximum entropy models or Generalized Linear Models (GLM). We also discuss the different types of correlations: those induced by a shared stimulus and those induced by neurons interactions. This work has been presented in several conferences [40], [39], [39], [15] and published in Chaos, Solitons and Fractals [17].

5.2. Mean field approaches

5.2.1. Asymptotic description of neural networks with correlated synaptic weights

Participants: Olivier Faugeras, James Maclaurin.

We study the asymptotic law of a network of interacting neurons when the number of neurons becomes infinite. Given a completely connected network of neurons in which the synaptic weights are Gaussian correlated random variables, we describe the asymptotic law of the network when the number of neurons goes to infinity. We introduce the process-level empirical measure of the trajectories of the solutions to the equations of the finite network of neurons and the averaged law (with respect to the synaptic weights) of the trajectories of the solutions to the equations of the network of neurons. The main result of this work is that the image law through the empirical measure satisfies a large deviation principle with a good rate function which is shown to have a unique global minimum. Our analysis of the rate function allows us also to characterize the limit measure as the image of a stationary Gaussian measure defined on a transformed set of trajectories. This work is available on ArXiV and is under review for a Journal. A preliminary version has been presented at the CNS meeting [42].

5.2.2. Beyond dynamical mean-field theory of neural networks

Participants: Bruno Cessac, Massimiliano Muratori.

We consider a set of N firing rate neurons with discrete time dynamics and a leak term. The nonlinearity of the sigmoid is controlled by a parameter and each neuron has a firing threshold, Gaussian distributed (thresholds are uncorrelated). The network is fully connected with correlated Gaussian random synaptic weights, with mean zero and covariance matrix. When synaptic weights are uncorrelated the dynamic mean field theory allows us to draw the bifurcation diagram of the model in the thermodynamic limit (N tending to infinity): in particular there is sharp transition from fixed point to chaos characterized by the maximum Lyapunov exponent, which is known analytically in the thermodynamic limit. However, mean-field theory is exact only in the thermodynamic limit and when synaptic weights are uncorrelated. What are the deviations from mean-field theory observed when one departs from these hypotheses ? We have first studied the finite size dynamics.

For finite N the maximal Lyapunov exponent has a plateau at 0 corresponding to a transition to chaos by quasi-periodicity where dynamics is at the edge of chaos. This plateau disappears in the thermodynamic limit. Thus, mean-field theory neglects an important finite-sized effect since neuronal dynamics at the edge of chaos has strong implications on learning performances of the network. We also studied the effect of a weak correlation on dynamics. Even when correlation is small, one detects an important deviation on the maximal Lyapunov exponent. This work has been presented at the CNS conference in Paris, 2013 [43].

5.3. Neural fields theory

5.3.1. Existence of localized solutions

Participants: Pascal Chossat, Grégory Faye, James Rankin.

We have started to tackle the problem of rigorously proving the existence of localized solutions to the neural fields equations. Existence of such solutions had been assumed or guessed from numerical simulations by other researchers. In a series of articles starting with [55] we have used ideas from the theory of ordinary differential equations (existence of homoclinic orbits) [56], [19], and the theory of partial differential equations (Swift-Hohenberg equation) [16] to show the existence of localized solutions for an extended variety of neural fields equations. This is important both theoretically and for neuroscience since these solutions are considered to characterize working (short-term) memory.

5.3.2. A Center Manifold Result for Delayed Neural Fields Equations

Participants: Olivier Faugeras, Romain Veltz.

We have developed a framework for the study of delayed neural fields equations and proved a center manifold theorem for these equations. Specific properties of delayed neural fields equations make it difficult to apply existing methods from the literature concerning center manifold results for functional differential equations. Our approach for the proof of the center manifold theorem uses the original combination of results from Vanderbauwhede et al. [1992] together with a theory of linear functional differential equations in a history space larger than the commonly used set of time-continuous functions. This work has appeared in the SIAM Journal on Mathematical Analysis [24].

5.3.3. Interplay between synaptic delays and propagation delays in neural fields equations

Participant: Romain Veltz.

Neural field equations describe the activity of neural populations at a mesoscopic level. Although the early derivation of these equations introduced space dependent delays coming from the finite speed of signal propagation along axons, there has been few studies concerning their role in shaping the nonlinear dynamics of neural activity. This is mainly due to the lack of analytical tractable models. On the other hand, constant delays have been introduced to model the synaptic transmission and the spike initiation dynamics. By incorporating the two kind of delays in the neural fields equations, we are able to find the Hopf bifurcation curves analytically which produce many Hopf-Hopf interactions. We use normal theory to study two different types of connectivity that reveals a surprisingly rich dynamical portrait. In particular, the shape of the connectivity strongly influences the spatiotemporal dynamics. This work has appeared in SIAM Journal on Applied Dynamical Systems [25].

5.3.4. Stochastic neural field equations: A rigorous footing

Participants: James Inglis, Olivier Faugeras.

We extend the theory of neural fields which has been developed in a deterministic framework by considering the influence spatio-temporal noise. The outstanding problem that we address here is the development of a theory that gives rigorous meaning to stochastic neural field equations, and conditions ensuring that they are well-posed. Previous investigations in the field of computational and mathematical neuroscience have been numerical for the most part. Such questions have been considered for a long time in the theory of stochastic partial differential equations, where at least two different approaches have been developed, each having its advantages and disadvantages. It turns out that both approaches have also been used in computational and mathematical neuroscience, but with much less emphasis on the underlying theory. We present a review of two existing theories and show how they can be used to put the theory of stochastic neural fields on a rigorous footing. We also provide general conditions on the parameters of the stochastic neural field equations under which we guarantee that these equations are well-posed. In so doing, we relate each approach to previous work in computational and mathematical neuroscience. We hope this will provide a reference that will pave the way for future studies (both theoretical and applied) of these equations, where basic questions of existence and uniqueness will no longer be a cause for concern. This work is available on ArXiV and is under review for a Journal.

5.4. Spike trains statistics

5.4.1. Decoding the retina with the first wave of spikes

Participants: John Barrett [Institute of Neuroscience, Medical School, Newcastle University, Newcastle UK], Pierre Kornprobst, Geoffrey Portelli, Evelyne Sernagor [Institute of Neuroscience, Medical School, Newcastle University, Newcastle UK].

Understanding how the retina encodes visual information remains an open question. Using MEAs on salamander retinas [60] showed that the relative latencies between some neuron pairs carry sufficient information to identify the phase of square-wave gratings (Using gratings of varying phase, spatial frequency, and contrast on mouse retinas, we extended this idea by systematically considering the relative order of all spike latencies, i.e. the shape of the first wave of spikes after stimulus onset. The discrimination task was to identify the phase among gratings of identical spatial frequency. We compared the performance (fraction correct predictions) of our approach under classical Bayesian and LDA decoders to spike count and response latency of each recorded neuron. Best results were obtained for the lowest spatial frequency. There, results showed that the spike count discrimination performance was higher than for latency under both the Bayesian (0,95$\pm$0,02 and 0,75$\pm$0,11 respectively) and LDA (0,95$\pm$0,01 and 0,62$\pm$0,03 respectively) decoders. The first wave of spikes decoder is (0,46$\pm$0,06) less efficient than the spike count. Nevertheless, it accounts for 50% of the overall performance. Interestingly, these results tend to confirm the rank order coding hypothesis [59] which we are currently investigating further.

This work has been presented in [45].

5.4.2. Spike train statistics from empirical facts to theory: the case of the retina

Participants: Bruno Cessac, Adrian Palacios [CINV-Centro Interdisciplinario de Neurociencia de Valparaiso, Universidad de Valparaiso].

This work focuses on methods from statistical physics and probability theory allowing the analysis of spike trains in neural networks. Taking as an example the retina we present recent works attempting to understand how retina ganglion cells encode the information transmitted to the visual cortex via the optical nerve, by analyzing their spike train statistics. We compare the maximal entropy models used in the literature of retina spike train analysis to rigorous results establishing the exact form of spike train statistics in conductance-based Integrate-and-Fire neural networks. This work has been published in Mathematical Problems in Computational Biology and Biomedicine, F. Cazals and P. Kornprobst, Springer [29].

5.4.3. Hearing the Maximum Entropy Potential of neuronal networks

Participants: Bruno Cessac, Rodrigo Cofré.

We consider a spike-generating stationary Markov process whose transition probabilities are known. We show that there is a canonical potential whose Gibbs distribution, obtained from the Maximum Entropy Principle (MaxEnt), is the equilibrium distribution of this process. We provide a method to compute explicitly and exactly this potential as a linear combination of spatio-temporal interactions. The method is based on the Hammersley Clifford decomposition and on periodic orbits sampling. As an application, we establish an explicit correspondence between the parameters of the Ising model and the parameters of Markovian models like the Generalized-Linear Model. This work has been presented in several conferences [39], [27], and submitted to Phys. Rev. Letters [41], see also the research report [31].

5.4.4. Spatio-temporal spike trains analysis for large scale networks using maximum entropy principle and Monte-Carlo method

Participants: Bruno Cessac, Olivier Marre [Institut de la Vision, Paris, France], Hassan Nasser.

Understanding the dynamics of neural networks is a major challenge in experimental neuroscience. For that purpose, a modelling of the recorded activity that reproduces the main statistics of the data is required. We present a review on recent results dealing with spike train statistics analysis using maximum entropy models (MaxEnt). Most of these studies have been focusing on modelling synchronous spike patterns, leaving aside the temporal dynamics of the neural activity. However, the maximum entropy principle can be generalized to the temporal case, leading to Markovian models where memory effects and time correlations in the dynamics are properly taken into account. We also present a new method based on Monte-Carlo sampling which is suited for the fitting of large-scale spatio-temporal MaxEnt models. The formalism and the tools presented will be essential to fit MaxEnt spatio-temporal models to large neural ensembles. This work has been presented in several conferences [39], [15], [44] and published in Journal of Statistical Mechanics [20].

5.4.5. Spike train statistics and Gibbs distributions

Participants: Bruno Cessac, Rodrigo Cofré.

We introduce Gibbs distribution in a general setting, including non stationary dynamics, and present then three examples of such Gibbs distributions, in the context of neural networks spike train statistics: (i) Maximum entropy model with spatio-temporal constraints; (ii) Generalized Linear Models; (iii) Conductance based Integrate and Fire model with chemical synapses and gap junctions. This leads us to argue that Gibbs distributions might be canonical models for spike train statistics analysis. This work has published in J. Physiol. Paris [15].

5.4.6. A maximum likelihood estimator of neural network synaptic weights

Participants: Bruno Cessac, Wahiba Taouali.

Given a conductance-based Integrate-and-Fire model where the spike statistics dependence on synaptic weights is known, can one reconstruct this network of synaptic weights from the observation of a raster plot generated by the network ? We have solved this inverse problem using an explicit expression of a maximum likelihood estimator based on the Newton-Raphson method. This estimator uses analytically computed gradients and Hessian of the likelihood function given by the product of conditional probabilities. The explicit form of these conditional probabilities can be found in [49]. Our results show that this method allows to estimate the set of connections weights knowing the input, the noise distribution and the leak function. This work has been presented in the CNS conference in Paris, 2013 [47].

5.5. Synaptic plasticity

5.5.1. Effects of Cellular Homeostatic Intrinsic Plasticity on Dynamical and Computational Properties of Biological Recurrent Neural Networks

Participants: Hugues Berry, Bruno Cessac, Bruno Delord, Jérémie Naudé.

Homeostatic intrinsic plasticity (HIP) is a ubiquitous cellular mechanism regulating neuronal activity, cardinal for the proper functioning of nervous systems. In invertebrates, HIP is critical for orchestrating stereotyped activity patterns. The functional impact of HIP remains more obscure in vertebrate networks, where higher-order cognitive processes rely on complex neural dynamics. The hypothesis has emerged that HIP might control the complexity of activity dynamics in recurrent networks, with important computational consequences. However, conflicting results about the causal relationships between cellular HIP, network dynamics and computational performance have arisen from machine learning studies. in this work, we assess how cellular HIP effects translate into collective dynamics and computational properties in biological recurrent networks. We develop a realistic multi scale model including a generic HIP rule regulating the neuronal threshold with actual molecular signaling pathways kinetics, Dale's principle, sparse connectivity, synaptic balance and Hebbian synaptic plasticity (SP). Dynamic mean-field analysis and simulations unravel that HIP sets a working point at which inputs are transduced by large derivative ranges of the transfer function. This cellular mechanism insures increased network dynamics complexity, robust balance with SP at the edge of chaos, and improved input separability. Although critically dependent upon balanced excitatory and inhibitory drives, these effects display striking robustness to changes in network architecture, learning rates and input features. Thus, the mechanism we unveil might represent a ubiquitous cellular basis for complex dynamics in neural networks. Understanding this robustness is an important challenge to unravel principles underlying self-organization around criticality in biological recurrent neural networks. This work has been published in the Journal of Neuroscience [21].

5.5.2. *Short-term synaptic plasticity in the deterministic Tsodyks-Markram model leads to unpredictable network dynamics*

Participants: Jesus Cortes, Mathieu Desroches, Serafim Rodrigues, Romain Veltz, Miguel Munoz, Terrence Sejnowski.

Short-term synaptic plasticity strongly affects the neural dynamics of cortical networks. The Tsodyks and Markram (TM) model for short-term synaptic plasticity accurately accounts for a wide range of physiological responses at different types of cortical synapses. We report a route to chaotic behavior via a Shilnikov homoclinic bifurcation that dynamically organizes some of the responses in the TM model. In particular, the presence of such a homoclinic bifurcation strongly affects the shape of the trajectories in the phase space and induces highly irregular transient dynamics; indeed, in the vicinity of the Shilnikov homoclinic bifurcation, the number of population spikes and their precise timing are unpredictable and highly sensitive to the initial conditions. Such an irregular deterministic dynamics has its counterpart in stochastic/network versions of the TM model: The existence of the Shilnikov homoclinic bifurcation generates complex and irregular spiking patterns and acting as a sort of springboard facilitates transitions between the down-state and unstable periodic orbits. The interplay between the (deterministic) homoclinic bifurcation and stochastic effects may give rise to some of the complex dynamics observed in neural systems.

This work has been published in the Proceedings of the National Academy of Sciences [52].

5.6. Visual Neuroscience

5.6.1. *Bifurcation Study of a Neural Fields Competition Model with an Application to Perceptual Switching in Motion Integration*

Participants: James Rankin, Andrew Meso [Institut de Neurosciences de la Timone, UMR 6193, CNRS, Marseille, France], Guillaume S. Masson [Institut de Neurosciences de la Timone, UMR 6193, CNRS, Marseille, France], Olivier Faugeras, Pierre Kornprobst.

In this work we have investigated the underlying mechanisms that gate multistable perception, by focusing on the presentation of 1:1 barber pole during long presentations, which is perceived to move in a direction that changes every few seconds. This phenomenon has been studied from the perspective of dynamical systems modeling and human psychophysics: From a modeling point of view, numerical tools from bifurcations analysis were applied to the study of a competition model posed as a feature-only neural field equation (with a continuous feature space) where adaptation and noise are implemented as mechanisms that can drive activity switches. Human psychophysics experiments were jointly done by INT (Institut de Neurosciences de la Timone, Marseille): Human observers were presented a moving grating stimulus over 15s while eye movements and reports of perceptual switches were recorded. Investigating the stimulus contrast, we found that the peak in switching rate observed experimentally occurs close to a bifurcation in the model that separates two mechanistic regimes. By identifying signatures of the switching predicted by the model with the behavioural data at different parts of the transition-contrast curve, we found for the first time, evidence for a dominance of driving mechanisms which shifts from noise dominated at low contrasts to adaptation dominated at higher contrasts.

This work has been published in [22], [23].

5.6.2. A Retinotopic Neural Fields Model of Perceptual Switching in 2D Motion Integration

Participants: Pierre Kornprobst, Guillaume S. Masson [Institut de Neurosciences de la Timone, UMR 6193, CNRS, Marseille, France], Kartheek Medathati, James Rankin.

In perceptual multistability a fixed but ambiguous stimulus can invoke multiple interpretations although only one can be held at a time. Visual motion stimuli are inherently ambiguous, for instance due to the aperture problem, which makes motion perception a complex inference task. The underlying cortical dynamics that select one percept out of multiple competing possibilities are not fully understood. Recent studies by [22] and [68] have tried to address this problem using the neural fields formalism. In [22], a switching behaviour for a classical psychophysics stimulus, the multistable barberpole, was successfully captured in a feature-only, one-layer model of MT with adaptation and noise. However, without a representation of space, only some very specific stimulus could be considered. The work reported in [68] provides a much more general framework for motion integration in a two layer-model, however, it fails to capture the switching behaviour as the mechanisms of adaptation and noise were not considered. Building on the strengths of both studies, we propose a model that takes into account the spatial domain in a two-layer configuration whilst incorporating both adaptation and noise. Interactions between two layers processing local motion (V1 and MT) occurred through recurrent and lateral connections. The input stimuli are represented using direction of motion signals extracted using Reichardt detectors at corresponding 2D spatial locations. We use stimuli such as drifting bars and barberpoles to constrain the model to a suitable operating regime. In terms of computations, since the model is demanding, we implemented it using GPUs, extending the methods of [13]. Based on this implementation, we study dynamics of the model focusing on coherency in plaid motion (plaids and crossed barber pole).

This work has been presented in [28]

NEUROSYS Team

6. New Results

6.1. From the microscopic to the mesoscopic scale

Participants: Axel Hutt, Laure Buhry, Meysam Hashemi, Pedro Garcia Rodriguez, Peter beim Graben, James Wallace Sleigh.

Several previous studies focus on the derivation of neural population models. However most of these studies do not consider explicitly the microscopic properties of neurons, such as synaptic receptor dynamics or ion-channel distributions, although they may be implicit. The resulting models in some previous studies are poorly tractable analytically due to their complexity. Moreover, the complexity of previous models makes it difficult to discover those elements in the model that induce certain dynamical features as observed in experiments. Essentially most of previous studies do not consider the spatial interactions of neurons and, importantly, neglect delays present in biological networks. We aim to improve some previous models and a first step to a new statistical approach has been developed [11], [18], [17], [22] to bridge the scales between the network activity of coupled spiking single neurons and statistical quantities of populations, e.g., the mean membrane potential in the network and the networks population firing rate. Our work considers the specific effect of anaesthetics and takes into account the physiological effects of extra-synaptic GABA$_A$-receptors at single neurons, which are highly sensitive to anaesthetic drugs, such as propofol. We find numerically by simulation of a spiking neural network that propofol on single neuron level diminishes the network oscillation power in the $\alpha-$frequency band and affects strongly the spike coherence in the population. Such effects have been shown in previous experimental data obtained during propofol anaesthesia demonstrating the importance of extra-synaptic receptor dynamics in the understanding of experimental phenomena in anaesthesia.

The neural origin of generation and planning of motor action in humans is still unknown. In this context, psychophysical experiments and the neural modeling of the gained results may lead to further insight. We have participated in an experimental and theoretical study [8] to reveal the effect of temporal attention on non-conscious prime processing. Our stochastic accumulator model improves extensively the standard accumulator model for reaction time by involving additional stochastic neural accumulators, which permits an almost perfect fit to experimental data. The model indicates that motor action, which is generated on a population level, obeys a stochastic accumulation of activity of single neuron activity.

6.2. From the mesoscopic to the behavior scale

Participants: Axel Hutt, Laurent Bougrain, Eric Nichols, Maxime Rio, Carolina Saavedra, Louis Korczkovsky, Alexandre Martin, Pierre-Jean Morieux.

To link neural population activity to behavior, it is necessary to understand well the dynamic properties of population models which we have studied in general models [5], [14], [15], [16], [24]. To this end, we have analyzed a neural mass model [4] describing the neural population activity subject to synaptic anaesthetic action to explain characteristic signal features in measured EEG. The model explains the gain of power in the $\alpha-$ and $\delta-$frequency observed experimentally by a dynamic oscillatory instability (Hopf instability). The model considers a cortical population only and hence the result indicates that the experimental feature observed may originate in the cortex.

An extended population model considers not only the cortex but a feedback-loop to the thalamus. This model involves a delayed interaction. At first, we have studied the dynamics of delayed dynamical systems subject to additive stimuli [23], [7], [6] to learn more about the expected activity. Our first study of a linear thalamo-cortical feedback model [21], [20] reveals the descriptive power of neural mass models to describe EEG under anaesthesia.

In order to learn more about the effect of anaesthetics on neural populations, we have participated in the data analysis of an experimental study on anaesthesia in animals [9]. Moreover we have started developing new data analysis techniques to extract novel features from EEG. In his doctoral thesis, Maxime Rio has developed a new method to detect transient amplitude synchronization in multi-variate time series in a subset of time series [1]. Carolina Saavedra has conducted wavelet analysis in her thesis to improve the denoising in BCI-relevant measured signals [2]. Another study [3] proposes a new recurrence plot-technique based on symbolic dynamics. It extracts spatio-temporally recurrence patterns in a multivariate dataset which reflect underlying neural recurrent dynamics.

The event-related potentials (ERP) in EEG are important markers of cognitive processes in the brain and serve as features to control interfaces in BCI. We have performed advanced studies to improve the detection of ERP [13], [12].

NON-A Project-Team

6. New Results

6.1. Homogeneity theory and analysis of nonlinear systems

Homogeneity is a kind of symmetry, if it is presented in a system model, then it may simplify analysis of stability and performance properties of the system. The new results obtained in 2013 are as follows:

- The notion of geometric homogeneity has been extended for differential inclusions in [44]. This kind of homogeneity provides the most advanced coordinate-free framework for analysis and synthesis of nonlinear discontinuous systems. Theorem of L. Rosier on a homogeneous Lyapunov function existence and an equivalent notion of global asymptotic stability for differential inclusions have been presented. Robustness properties (ISS) of sliding mode systems applying the homogeneity concept have been considered in [46].
- Retraction obstruction for time-varying stabilization on compact manifolds has been revisited in [13].
- Several conditions have been proposed to check different robustness properties (ISS, iISS, IOSS and OSS) for generic nonlinear systems applying the weighted homogeneity concept (global or local) in [14], [45]. The advantages of this result are that, under some mild conditions, the system robustness can be established as a function of the degree of homogeneity.
- A new algorithm for the analysis of strange attractors has been presented in [51]. An application of that results for observability-singularity manifolds in the context of chaos based cryptography has been given in [52].
- Exciting multi-DOF systems by feedback resonance has been considered in [20].
- Some conditions on existence of oscillations in hybrid systems have been established in [23], [57]. An application to a humanoid robot locomotion has been considered.
- Considering two chaotic Rossler systems, the paper [83] presents a study on the forced synchronization of two systems, bidirectionally coupled by transmitting unidirectional signals which explicitly depend on a single state variable (from the emitter) and only affect directly the dynamics corresponding to the transmitted state variable (of the receiver).
- The paper [33] is concerned with the construction of local observers for nonlinear systems without inputs, satisfying an observability rank condition. The aim of this study is, first, to define a homogeneous approximation that keeps the observability property unchanged. This approximation is further used in the synthesis of local observer which is proven to be locally convergent for Lyapunov-stable systems.
- The paper [74] addresses the problem of exact average-consensus reaching in a prescribed time. The communication topology is assumed to be defined by a weighted undirected graph and the agents are represented by integrators. A nonlinear control protocol, which ensures a finite-time convergence, is proposed. With the designed protocol, any prescribed convergence time can be guaranteed regardless of the initial conditions.
- The Implicit Lyapunov Function (ILF) method for finite-time stability analysis has been introduced in [75]. The control algorithm for finite-time stabilization of a chain of integrators has been developed. The scheme of control parameters selection has been presented by LMIs. The robustness of the finite-time control algorithm with respect to system uncertainties and disturbances has been studied. The new high order sliding mode control has been derived as a particular case of the developed finite-time control algorithm. The settling time estimate has been obtained using ILF method. The algorithm of practical implementation of the ILF control scheme has been discussed.

6.2. Model-free control

The model free control techniques form a new and quickly developing area of control theory. It has been established by the team members and nowadays these tools find many practical applications and attract a lot of attention due to their clear advantages for designers: they provide a control law independently in the model knowledge. The achievements obtained in 2013 are as follows:

- A new development of the model-free control theory with application to active magnetic bearing control have been presented in [53].
- ”Model-free control” and the corresponding ”intelligent” PID controllers (iPIDs), which already had many successful concrete applications, have been presented in [27] for the first time in a unified manner, where the new advances have been taken into account.
- In [62], it is shown that the "intelligent" controllers, which are associated to the recently introduced model-free control synthesis, may be easily implemented on cheap and small programmable devices.
- An application of the model-free control for regulation of the water level under several constraints has been reported in [40].

6.3. Algebraic technique for estimation, differentiation and its applications

Elementary techniques from operational calculus, differential algebra, and non-commutative algebra lead to a new algebraic approach for estimation and detection. It is investigated in various areas of applied sciences and engineering. The following lists only some applications:

- Design of a stabilizing feedback based on acceleration measurements and an algebraic state estimation method has been proposed in [54].
- An extension of the algebraic differentiation method to fractional derivatives calculation in continuous and discrete time has been studied in [88] and [89] respectively. Applications to identification and parameter estimation of fractional linear systems have been considered in [67], [68].
- Smoothing noisy data with spline functions is well known in approximation theory. Smoothing splines have been already used to deal with the problem of numerical differentiation. In [43], we extend this method to estimate the fractional derivatives of a smooth signal from its discrete noisy data. We begin with finding a smoothing spline by solving the Tikhonov regularization problem. Then, we propose a fractional order differentiator by calculating the fractional derivative of the obtained smoothing spline.
- In [81], we apply an algebraic method to estimate the amplitudes, phases and frequencies of a biased and noisy sum of complex exponential sinusoidal signals. The obtained estimates are integrals of the noisy measured signal: these integrals act as time-varying filters.

6.4. Observability and observer design for nonlinear systems

Observability analysis and observer design are important issues in the field of control theory. Some recent results are listed below:

- An epistemology of observation theory and its application in the design of software sensor in power electronics have been presented in [42].
- New results on observability and detectability of singular linear systems with unknown inputs have been developed in [12].
- The paper [47] supplies a new algorithm to compute the internal dynamics (or inversion dynamics) of affine MIMO control nonlinear systems.
- The design of observers for nonlinear systems with unknown, time-varying, bounded delays, on both state and input, still constitutes an open problem. In [28], we show how to solve it for a class of nonlinear systems by combining the high gain observer approach with a Lyapunov-Krasovskii functional. Sufficient conditions have been provided to prove the practical stability of the observer.

- An influence of restricted isometry property to the observability under sparse measurements has been analyzed in [65].
- The paper [38], [79] concerns the design of a nonlinear observer through a transformation of a nonlinear system into an observer form that supports a high gain observer. Sufficient geometrical condition has been deduced to guarantee the existence of change of coordinates allowing the transformation of a nonlinear system into the proposed normal form. In [80], the Partial Observability Normal Forms (PONF) of nonlinear dynamical systems have been investigated. Necessary and sufficient conditions for the existence of a diffeomorphism bringing the original nonlinear system into a PONF have been established.

6.5. Sliding mode control and estimation

Sliding mode algorithms are very popular for finite-time estimation and regulation. The recent results obtained by the group are as follows:

- Some constructive approximations and an alternative theoretic characterization of some classes of sliding mode control processes has been presented in [11].
- In [64] we investigate observer design under sparse measurement, i.e. under Nyquist-Shannon frequency. An analysis demonstrates that it is impossible to use only a high order sliding mode observer in the case of sparse measurement. Then it has been shown that a high order sliding mode observer coupled with an impulsive observer is a pertinent solution at least for some particular class of systems.
- Anomaly detection has been an active open problem in the networks community for several years. In [35], we aim at detecting such abnormal signals by control theory techniques. Several classes of sliding mode observers have been proposed for a fluid flow model of the TCP/internet protocol network.
- A sliding mode control has been developed for robust stabilization of fractional-order input-delay linear systems in the presence of uncertainties and external disturbances in [78]. First, a fractional-order state predictor has been used to compensate the delay in the input control. Second, a robust sliding mode control has been proposed in order to stabilize the system and to thwart the effect of model uncertainties and external disturbances. The sliding mode controller has been designed by considering a sliding surface defined by fractional order integral.

6.6. Non-linear, Sampled and Time-delay systems

Nonlinearities, sampling, quantization and time-delays cause serious obstructions for control and observer design in many fields of techniques and engineering (e.g. networked and internet systems, distributed systems etc.). The proposed by the team algebraic approach suits well for estimation and regulation in such a type of systems. The recent results are listed below:

- The work [59] aims at decreasing the number of sampling instants in state feedback control for perturbed linear time invariant systems. The approach is based on linear matrix inequalities obtained thanks to Lyapunov-Razumikhin stability conditions and convexification arguments that guarantee the exponential stability for a chosen decay-rate.
- A novel self-triggered control, which aims at decreasing the number of sampling instants for the state feedback control of perturbed linear time invariant systems, has been proposed in [60]. The approach is based on convex embeddings that allow for designing a state-dependent sampling function guaranteeing the system's exponential stability for a desired decay-rate and norm-bounded perturbations. One of the main contributions of the paper [60] is an LMI based algorithm that optimizes the choice of the Lyapunov function so as to enlarge the lower-bound of the sampling function while taking into account both the perturbations and the decay-rate.
- In [63], we consider the issue of stabilizing a class of linear systems using irregular sampled output measurements.

- The paper [73] is dedicated to the stability analysis of nonlinear sampled-data systems, which are affine in the input. Assuming that a stabilizing continuous-time controller exists and it is implemented digitally, we intend to provide sufficient asymptotic/exponential stability conditions for the sampled-data system. This allows to find an estimate of the upper bound on the asynchronous sampling periods. The stability analysis problem is formulated both globally and locally. The main idea of the paper is to address the stability problem in the framework of dissipativity theory. Furthermore, the result is particularized for the class of polynomial input-affine sampled-data systems, where stability may be tested numerically using sum of squares decomposition and semidefinite programming.
- The problem of output control design for linear system with unknown and time-varying input delay, bounded exogenous disturbances and bounded deterministic measurement noises has been considered in [77]. The prediction technique has been combined with Luenberger-like observer design in order to provide the stabilizing output feedback. The scheme of parameters tuning for reduction of measurement noises effect and exogenous disturbances effects has been developed using the Attractive Ellipsoids Method.
- Using the theory of non-commutative rings, the paper [39] studies the delay identification of nonlinear time-delay systems with unknown inputs. A sufficient condition has been given in order to deduce an output delay equation from the studied system. Then necessary and sufficient conditions have been proposed to judge whether the deduced output delay equation can be used to identify the delay, which is involved in this equation.

6.7. Interval control and estimation

In many cases due to parametric and/or signal uncertainties presented in a plant model it is not possible to design a conventional observer, which provides a point-wise estimate of state in a finite time or asymptotically. In this case it is still frequently possible to design interval observers, which generate an estimate on the interval of the admissible values of the state at the current instant of time. The recent new results in this field are listed below:

- The work [49] is devoted to interval observer design for Linear Parameter-Varying (LPV) systems under assumption that the vector of scheduling parameters is not available for measurements. Stability conditions are expressed in terms of matrix inequalities, which can be solved using standard numerical solvers. Robustness and estimation accuracy with respect to model uncertainty is analyzed. Two solutions are proposed for nonnegative systems and for a generic case. The efficiency of the proposed approach is demonstrated through computer simulations.
- Development of interval observers for time invariant [55] and time-varying [21] discrete-time systems has been presented by the members of the team.
- Interval estimation for uncertain systems with time-varying delays has been considered in [22], [56]. A reduced-order interval observer has been designed, stability and robustness conditions have been obtained.
- The paper [24] is devoted to design of interval observers for Linear Time Varying (LTV) systems and a class of nonlinear time-varying systems in the output canonical form. An interval observer design is feasible if it is possible to calculate the observer gains making the estimation error dynamics cooperative and stable. It has been shown in [24] that under some mild conditions the cooperativity of an LTV system can be ensured by a static linear transformation of coordinates.
- The problem of output stabilization of a class of nonlinear systems subject to parametric and signal uncertainties has been studied in [25]. First, an interval observer has been designed estimating the set of admissible values for the state. Next, it has been proposed to design a control algorithm for the interval observer providing convergence of interval variables to zero, that implies a similar convergence of the state for the original nonlinear system. An application of the proposed technique shows that a robust stabilization can be performed for linear time-varying and LPV systems without assumption that the vector of scheduling parameters is available for measurements.

- The paper [26] deals with the problem of joint state and parameter estimation based on a set adaptive observer design. The problem is formulated and solved for an LPV system. The resolution methodology avoids the exponential complexity obstruction usually encountered in the set-membership parameter estimation.
- The output stabilization problem for a linear system with an unknown bounded time-varying input delay has been considered in [34], [76]. The interval observation technique has been applied in order to obtain guaranteed interval estimate of the system state. The procedure of the interval observer synthesis uses lower and upper estimates of the unknown delay and requires to solve a special Silvester's equation. The interval predictor has been introduced in order to design a linear stabilizing feedback. The control design procedure is based on LMIs.
- The paper [37] describes a robust set-membership-based Fault Detection and Isolation (FDI) technique for a particular class of nonlinear systems, the so-called flat systems. The proposed strategy consists in checking if the expected input value belongs to an estimated feasible set computed using the system model and the derivatives of the measured output vector. The output derivatives are computed using a numerical differentiator. The set-membership estimator design for the input vector takes into account the measurement noise thereby making the consistency test robust.
- The objective of the work [82] is to develop some design methods of interval observers for a class of nonlinear continuous-time systems. It has been assumed that the estimated system can be represented as a superposition of the nominal subsystem (belonged to the class of uniformly observable systems) and a Lipschitz nonlinear perturbation vanishing at the origin. Then it has been shown that there exists an interval observer for the system that estimates the set of admissible values for the state consistent with the output measurements.

6.8. Networked robots

The mobile robots constitute an important area of practical development for the team:

- The paper [71] presents a path planning algorithm for autonomous navigation of non-holonomic mobile robots in complex environment. The irregular contour of obstacles is represented by segments. The goal of the robot is to move towards a known target while avoiding obstacles. The velocity constraints, kinematic robot model and non-holonomic constraint are considered in the problem. The optimal path planning problem is formulated as a constrained receding horizon planning problem and the trajectory is obtained by solving an optimal control problem with constraints. Local minima are avoided by choosing intermediate objectives based on the real time environment.
- The paper [69] presents a cooperative path planning approach for the navigation of non-holonomic mobile robots in environment with obstacles. Shared information can be obtained by sharing the local information between robots, thus the trajectories can be more optimized. Visibility graph approach is used to generate a series of intermediate objectives which guarantee the robots to reach the final objective without local minima. Then the reach of intermediate objectives is ensured by the optimal path planning algorithm. The velocity constraints, kinematic constraints and non-holonomic constraints of the mobile robot are considered in the problem.
- The paper [70] presents the real-time identification of different types of non-holonomic mobile robot systems. Since the robot type is a priori unknown, the robot systems are formulated as a switched singular nonlinear system, and the problem becomes the real-time identification of the switching signal, and then the existence of the input-output functions and the distinguishability of the systems are studied.
- An intelligent PID controller (i-PID controller) has been applied to control the non-holonomic mobile robot with measurement disturbance in [72]. Because of the particularity of the non-holonomic systems, this paper proposes to use a switching parameter α in the i-PID controller.

6.9. Applications

As it was mentioned, Non-A is a kind of "method-driven" project, which deals with different aspects of finite-time estimation and control. Thus different applications are possible, ones touched this year are as follows (skipping the networked robots considered in the previous section):

- A sensorless speed control for a DC series motor has been presented in [41] based on sliding-mode control and estimation algorithms.
- The paper [48] presents a feasibility study, which aims to demonstrate the applicability of the CNC automation philosophy for the process of AFM probe-based nano machining conducted on commercial AFM instruments.
- An oscillatory failure case detection for aircrafts using non-homogeneous sliding-mode differentiator in noisy environment has been considered in [50].
- Sensorless fault tolerant control for induction motors has been developed in [18].
- The problem of an actuator fault detection in aircraft systems has been considered in [19]. A particular attention has been paid to the oscillatory failure case study.
- In [58], we consider a vehicle equipped with active front steer and rear torque vectoring. While the former adds an incremental steer angle to the driver's input, the latter imposes a torque by means of the rear axle. The active front steer control is actuated through the front tires, while the rear torque vectoring can be actuated through the rear tires. A nonlinear controller using the super-twisting algorithm has been designed in order to track in a finite time the lateral and yaw angular velocity references.
- Systematic and multifactor risk models have been revisited via algebraic methods, which were already successfully developed in signal processing and in automatic control, in [61].
- In [84], we address the problem of approximating scattered data points by C1-smooth polynomial spline curves and surfaces using L1-norm optimization. The use of this norm helped us to preserve the shape of the data even near to abrupt changes.
- As capacitor voltages are necessary for the three-cell DC-DC chopper control, the estimation of such voltages by an observer is attractive solution in terms of cost. However, due to the hybrid behaviour of this structure, the capacitor voltages may be partially or even not observable for a given switching configuration. In other words, the observability matrix associated to the capacitor voltages never has a full rank. In order to make the observer conceivable, the paper [29] proposes a new design by establishing sufficient conditions under which the capacitor voltages can be reconstructed within appropriate specific switching sequence and not necessarily instantly.
- The problem of converters coordination of a fuel cell system involving a hydrogen fuel cell with supercapacitors for applications with high instantaneous dynamic power has been addressed in [32]. The problem is solved by using a non-linear controller based on passivity.
- The paper [66] is devoted to development of control algorithms for nonlinear parametrically uncertain systems. Original system dynamics is approximated by a set of local NARX models combined by a special mixing rule. Algorithm for local models' parameters estimation and structure adjustment has been developed. The developed technique has been applied to the problem of regulation of spark ignition engines.
- The paper [36] is dedicated to the problem of pneumatic cylinder control without pressure measurement. Based on the theory of homogeneous, finite time stable, ordinary differential equations, a state feedback nonlinear controller has been proposed. The closed loop system stability has been proven and an attraction domain of the controller has been given.

NUMED Project-Team (section vide)

OAK Project-Team

6. New Results

6.1. Scalable and Expressive Techniques for the Semantic Web

The team has continued developing expressive models and scalable algorithms for exploiting Semantic Web data, in particular RDF graphs, as well as rich corpora consisting of Web documents with semantic annotations.

We have studied efficient algorithms for answering RDF queries in the presence of schema (or semantic) constraints such as described through the RDF Schema language. The difficulty here consists of efficiently taking into account the data that is implicitly present in the RDF database due to semantic constraints, and which needs to be reflected in query results. We have identified the expressive database fragment of RDF, which extends previously identified fragments of the RDF specification by allowing more expressive schema and queries, and provided novel efficient algorithms for answering Basic Graph Pattern queries (a popular dialect of the standard SPARQL query language) over RDF graphs pertaining to the RDF Database Fragment. Our query answering algorithms take advantage of the processing power of a relational database management system while also reflecting RDF semantics [25].

The ability to exploit large corpora of heterogeneous RDF data requires tools for analyzing RDF content through the lenses of a specific user perspective, or user need. Such tools are commonplace in the context of relational data management, where data warehousing is a well-developed area, but lack completely in the realm of RDF. We have proposed a novel framework for building and exploiting all-RDF data warehouses [33] and have implemented this framework in a proof-of-concept platform [32]. A main contribution of this work is to preserve RDF graph structure, heterogeneity, and rich semantics from the base data to the analytical schema and analytical schema instance. Thus, our proposal is the first to allow the analysis of rich Semantic Web (RDF) data while preserving its rich content and semantics. For more information on this project, see https://team.inria.fr/oak/warg/.

We have investigated the usage of semantics as a way to enrich, interconnect, and interpret rich corpora of Web data. In particular, within the XR project, we had proposed in prior work the XR (XML+RDF) data model which integrates XML documents and RDF triples treating both as first-class citizens. One particular use of XR is to annotate nodes in XML documents, by RDF triples which may for instance describe their properties or state how nodes are semantically related to some concept or to each other. In [18] we describe the data model and core query language, make a comprehensive analysis of query evaluation algorithms, and describe extensive experiments carried within a fully implemented platform, as part of the PhD thesis of J. Leblay [12]. The XR platform was put to task in an application context related to digital journalism, where an XR content warehouse is continuously enriched through document analysis and annotation. This scenario has lead to a software demonstration [24], [35] and a keynote tutorial [38]. In collaboration with A. Deutsch, we have extended the XR query language and provided query-view composition algorithms in [41].

6.2. Massively Distributed Data Management Systems

Our work on the AMADA platform has shown how the different sub-systems of a popular cloud platform (namely, Amazon Web Services, or AWS in short) can be harnessed to build scalable stores and query evaluation engines for XML and RDF data. In [23], we propose and compare several storage and indexing strategies within AWS, and show that they help reduce not only query evaluation time but also the monetary costs associated to the exploitation of the AWS-based store, since the index helps direct queries only to the subsets of the data likely to have results for the query. Thus, the total effort (and the costs charged by AWS) in relation to the processing of a given query are reduced. A similar study focused mostly on RDF data management appears as a book chapter [40]. More information can be found at http://cloak.saclay.inria.fr/research/amada/.

Semantic Web data collections, that is, RDF graphs, may be very voluminous since RDF natively enables connections between different RDF databases (which may have been produced independently and in ignorance of each other) through the usage of common URIs (resource identifiers) in two or more databases. To scale up to such large volumes, we have developed CliqueSquare, a novel platform for storing and querying RDF graphs in a MapReduce-based architecture such as Hadoop. We have described the storage and query algorithm in [34]. Our analysis of existing frameworks and algorithms for managing large RDF graphs in a highly distributed environment has lead to the tutorial [27].

Large-scale distributed processing of complex data was considered from a different perspective in our Delta project. Here, we considered the setting where one data source publishes new data items at a very high rate, and numerous clients subscribe to some of the updates by means of queries that must be matched by the published items. In this setting, the source may quickly become the bottleneck due to limitations in its capacity to match the published item against the subscription and/or to send the matching updates. We propose a fully automated approach for distributing the data dissemination effort across the network of subscribers, by identifying some which act as secondary data sources for others, in a peer-to-peer fashion. This distributed dissemination network is chosen so as to optimize a combination of overall dissemination costs and data propagation latency; since the space of options has daunting complexity, approximate algorithms involving Binary Integer Programming techniques were proposed in [20], [37], [42], and concluded in the PhD thesis of A. Katsifodimos [11].

6.3. Advanced Algorithms for Efficient XML processing

In 2013, several research works of the team focusing on advanced algorithms for processing XML data have been finalized and concluded through prestigious journal publications.

A first line of work concerned the usage of materialized views to speed up the evaluation of complex XML queries. In our previous work we had demonstrated that such views may bring up very significant speed-up factors of several orders of magnitude. However, materialized views need to be kept up to date when the underlying database changes. In [14] we have described efficient algorithms for updating materialized views expressed in a rich dialect of XQuery, the standard query language for XML.

A second class of work was concerned with XML static type analysis, in particular with the crucial problem of deciding XML type inclusion, that is: whether any XML tree of type τ_1 is also of type τ_2 where τ_1, τ_2 are XML types with interleaving and counting (currently adopted by main stream schema languages). For these types, inclusion is EXPSPACE-complete. We have defined and formally studied a quadratic subtype-checking algorithm for the case where the right-hand side type τ_2 meets some restrictions on symbol occurrences and the use of counting. These restrictions are often met by human-designed types, so our technique perfectly fits the needs of typical XML type-checking algorithms, which frequently require to check for inclusion a machine-generated subtype τ_1 against a human-defined supertype τ_2. Our approach has been validated by extensive experimental results [16]. In addition, we have devised and formally studied an alternative algorithm, still for the asymmetric case where τ_2 is restricted, based on structural, top-down analysis of types expression. This algorithm is almost linear: it has a linear-time backbone, and resorts to the above quadratic approach for some specific parts of the compared types. Our experiments show that this new algorithm is much faster than the quadratic one and that it typically runs in linear time, hence it can be used as a building block for a practical type-checking compiler for XML programs and queries [15].

Third, we have completed and concluded our work on type-based document projection for efficient XML data management. The idea here is to restrict XML documents, prior to evaluating a query over them, to only those parts of the document that the query actually needs to consult. We provide algorithms for determining such document parts and experimentally demonstrate the benefits of such techniques, in [13].

Finally, we have devised a system that is able to process both queries and updates on very large XML documents [22]. As observed in recent works, such very large documents are generated and processed in several contexts, in particular in those involving scientific data and logs. Our system supports a large fragment of XQuery and XUF (XQuery Update Facility). The system exploits dynamic and static partitioning to distribute the processing load among the machines of a MapReduce cluster. The proposed technique applies when queries and updates are iterative, i.e., they iterate the same query/update operations on a sequence of subtrees of the input document. From our experience many real world queries and updates actually meet this property. Our partitioning technique is schema-less, as the presence of a user-supplied schema is not required; indeed, this technique only relies on path information extracted from the input query/update. Experiments conducted on a 8-machine Hadoop cluster have demonstrated that the system is able run both iterative queries and updates on quite large documents.

6.4. Data Transformation Management

With the increasing complexity of data processing queries, for instance in applications such as relational data analysis or integration of Web data (e.g., XML or RDF) comes the need to better manage complex data transformations. This includes systematically verifying, maintaining, and testing the transformations an application relies on. In this context, Oak has focused on verifying the semantic correctness of a declarative program that specifies a data transformation query, e.g., an SQL query. To this end, we have investigated how to leverage data provenance (the information of the origin of data and the query operators) for query debugging. More specifically, we developed and implemented novel algorithms to explain why data is missing from the result of a relational query. As opposed to our previous work, which produced explanations based on the available source data, our new algorithms return explanations based on query operators [31] or both [26].

6.5. Social Data Management

We considered top-k query answering in social tagging systems, also known as folksonomies, a problem that requires a significant departure from existing, socially agnostic techniques. In a network-aware context, one can and should exploit the social links, which can indicate how users relate to the seeker and how much weight their tagging actions should have in the result build-up. Our solutions addressed the main drawbacks of previous approaches. With respect to applicability and scalability, we avoid expensive and hardly updatable pre-computations of proximity values. With respect to efficiency, we show that our algorithm is instance optimal in the existing techniques. Our main results in this direction have been presented recently in [29], [28], [21].

OASIS Project-Team

6. New Results

6.1. Programming and Composition Models for Large-Scale Distributed Computing

6.1.1. Multi-active Objects

Participants: Ludovic Henrio, Fabrice Huet, Justine Rochas.

The active object programming model is particularly adapted to easily program distributed objects: it separates objects into several *activities*, each manipulated by a single thread, preventing data races. However, this programming model has its limitations in terms of expressiveness – risk of deadlocks – and of efficiency on multicore machines. We proposed to extend active objects with *local multi-threading*. We rely on declarative *annotations* for expressing potential concurrency between requests, allowing easy and high-level expression of concurrency. This year we realized the following:

- publication of the multiactive object programming model in COORDINATION 2013 [19]
- extension of the annotations to support the specification of:
  - thread management. This aims at specifying (i) thread reservation and (ii) thread limitation in order to control more finely the allocation of threads in a multiactive object.
  - priority of requests. The progammer can now specify a priority graph to have an influence on the order of execution of requests in a multiactive object

  This extension was initially explored in a master thesis [34] and led to a publication in SAC 2014 [21].
- extensive use of multiactive objects in our CAN P2P network and implementation of usecases.

We plan to continue to improve the model, especially about compile-time checking of annotations and about fault tolerance of multiactive objects.

6.1.2. Algorithmic skeletons

Participant: Ludovic Henrio.

In the context of the SCADA associated team, we worked on the algorithmic skeleton programming model. The structured parallelism approach (skeletons) takes advantage of common patterns used in parallel and distributed applications. The skeleton paradigm separates concerns: the distribution aspect can be considered separately from the functional aspect of an application. In the previous year we designed the possibility for a skeleton to output events, which increases the control and monitoring capabilities. This year we achieved the following objectives:

- Encapsulation of the skandium skeleton runtime in a component in order to allow distributed execution of skeletons: local parallelism is handled by skandium while distributed execution is handled by the GCM component library.
- We applied the event framework for skeletons to design a framework allowing the skeleton execution to adapt autonomically in order to achieve a required qualiti of service. We have first promising results on this aspect and a publication has just been accepted to PMAM 2014.

6.1.3. Behavioural models for Distributed Components

Participants: Eric Madelaine, Nuno Gaspar, Oleksandra Kulankhina, Ludovic Henrio.

In the past [3], we defined the behavioural semantics of active objects and components. This year we extended this work to address group communications. On the practical side, this work contributes to the Vercors platform; the overall picture being to provide tools to the programmer for defining his application, including its behavioural specification. Then some generic properties like absence of deadlocks, but also application specific properties, can be validated on the composed model using an existing model-checker. We mainly use the CADP model-checker, that also supports distributed generation of state-space. This year our main achievements are the following:

- We improved the specification of the behavioural model generation for component systems that we specified last year [36]. A journal version is under submission.
- We extended the formal model of the GCM architecture and included the specification of the non-functional aspects.
- We worked on the design of a bisimulation equivalence relation adapted to pNets; such an equivalence relation would justify some of the verifications and simplifications we do in our verification platform. Bisimulation theory gives tools to prove the equivalent behaviour of two processes, but adapting it to the structural nature and to the parameterized definitions of pNets is a challenging task. We have obtained promising preliminary results on this aspect: we have a good library of examples illustrating the expressiveness of pNets and use it to study bisimulation techniques.
- We additionally have put considerable efforts on the improvement of the Vercors platform (see Section 5.2). We have totally updated the Vercors Components Editor. We have integrated the UML state machines editor from Obeo UML Designer (http://marketplace.obeonetwork.com/module/uml) into Vercors platform. The integrated editor provides the tools for the specification of the components behavior.
- We have started implementing the behavioural semantics of [36] in the Vercors platform. This task consists in generating the behavior of GCM components in the form of pNets from the GCM architecture defined using VCE. This is an important task, involving intricate engineering issues, but also interesting research on methods for reducing the size of the generated models.

This work was done in collaboration with Rabéa Ameur-Boulifa from Télécom-Paristech and Min Zhang from ECNU Shanghai.

In parallel with core developments of the behavioural specification environment, we further collaborated with our industrial partners and enhanced our work around the use of proof assistants for our specification and verification purposes. In particular, this year:

- We made significant improvements on Mefresa, our Mechanized Framework for the Reasoning on Software Architectures. These were published in [17]. Moreover, we obtained preliminary results regarding its integration with GCM/ProActive, our java middleware for parallel and distributed programming.
- We specified, verified and implemented the HyperManager, a GCM distributed application for the management and monitoring of E-Connectware — a solution for the management of distributed RFID infrastructures. This work was published as an industrial case study in [18].

6.1.4. Autonomic Monitoring and Management of Components

Participants: Françoise Baude, Bastien Sauvan.

We have completed the design of a framework for autonomic monitoring and management of component-based applications. We have provided an implementation using GCM/ProActive taking advantage of the possibility of adding components in the membrane. The framework for autonomic computing allows the designer to describe in a separate way each phase of the MAPE autonomic control loop (Monitoring, Analysis, Planning, and Execution), and to plug them or unplug them dynamically.

- This year, we worked on a journal paper presenting our implementation of GCM component model using active objects, and its use to provide autonomic components. The paper is under revision for SPE journal.

6.1.5. Optimization of data transfer in SOA and EDA models

Participants: Amjad Alshabani, Iyad Alshabani, Françoise Baude, Laurent Pellegrino, Bastien Sauvan, Quirino Zagareze.

Traditional client-server interactions rely upon method invocations with copy of the parameters. This can be useless in particular if the receiver does not effectively uses them. On the contrary, copying and transferring parameters lazily, and allowing the receiver to proceed without only some of them is a meaningful idea that we proved to be effective for active objects in the past [38]. This idea wasn't so far realized in the context of the web services technology, the most popular one used today for client-server SOAP-based interactions.

- We contributed to the offloading of objects representing parameters of the web service Java Apache CXF API [46]. It is innovative notably in the way the offloading of parameters for on-demand access can be delegated from services to services, which resembles the concept of first-class futures from ASP.
- Relying upon such an effective approach, we have applied a similar idea of "lazy copying and transfer" to the data parts of events in the context of event-driven architecture applications [26]. The middleware dynamically off-loads data (generally of huge size) attached to an event, according to some user-level policy expressed as annotation in the Java code at the subscriber side. The event itself, without its attachments, gets forwarded into the publish/subscribe brokering system (in our case, the EventCloud middleware, see Section 5.5) and its attachements are transferred to the subscriber on-demand. Compared to some existing propositions geared towards a data centric publish-subscribe pattern (e.g. the DDS OMG standard), ours is more user-friendly as it does not require the user code to explicitly program when to get the data attached to notified events. Also it features very low performance overhead, as additional experiments conducted show it: they are reported in an extended version of the SAC 2013 paper that is under minor revision for a special issue of the Science of Computer Programming journal.

Overall, this work opens the way towards a strong convergence between service oriented and event-driven technologies.

6.1.6. Multi-layer component architectures

Participant: Olivier Dalle.

Since a few years, we have been investigating the decomposition of a simulation application into multiple layers corresponding to the various concerns commonly found in a simulation: in addition to the various modeling domains that may be found in a single simulation application (e.g. telecommunications networks, road-networks, power-grids, and so on), a typical simulation includes various orthogonal concerns such as system modelling, simulation scenario, instrumentation and observation, distribution, and so on. This large number of concerns has put in light some limits of the traditional hierarchical component-based architectures and their associated ADL, as found in the FCM and GCM. In order order to cope with these limitations, we started a new component architecture model called Binding Layers centered on the binding rather than the component, with no hierarchy but advanced layering capabilities, and offering advanced support for dynamic structures. This project is composed of four levels of specification: the two first levels are ready for public realease, but some work is still needed for the development of the validation prototypes.

6.2. Middleware for Grid and Cloud computing

6.2.1. Distributed algorithms for CAN-like P2P networks

Participants: Ludovic Henrio, Fabrice Huet, Justine Rochas.

The nature of some large-scale applications, such as content delivery systems or publish/subscribe systems, built on top of Structured Overlay Networks (SONs), demands application-level dissemination primitives which do not overwhelm the overlay, i.e. which are efficient, and which are also reliable. Building such communication primitives in a reliable manner would increase the confidence regarding their behavior prior to deploying them in real settings. In order to come up with real efficient primitives, we take advantage of the underlying geometric topology of the overlay network and we also model the way peers communicate with each other. Our objective is to design and prove an efficient (in terms of messages and execution time) and reliable broadcast algorithm for CAN-like P2P networks. To this aim, this year, we realized the following:

- publication in FASE 2013 of a formalisation, in Isabelle/HOL, of CAN-like P2P networks [15]. Thank to this work, we proved that there exist a broadcast algorithm that does not produce any duplicated message in those networks. A first naive algorithm was exhibited to prove it.
- design and and publication of an optimal broadcast algorithm for CAN-like P2P networks in OPODIS 2013 [20]. The solution we have proposed is proven to be correct, optimal in terms of number of messages, and also efficient, as it provides a good parallelization during the dissemination.

We are also investigating new algorithms to efficiently build a SON when the peer involved already have data. Most of the work on SONs assume that new peers joining the network will arrive without data and thus get assigned a random position. However, if they already have data, they will have to send them to other peers, depending on the key space they are responsible of. In 2013, we continued on the tracks investigated in 2012:

- We proposed a first version of new join algorithms which try to allocate key sub-spaces to peers so that the amount of data that needs to be moved is minimal. An expected benefit of this work is that it should allow for fast and efficient reconstruction of a SON in case of a crash, without having to use distributed snaphshots.
- We have conducted preliminary experiments which shows a reduction of data transfer between 20% and 90%.

6.2.2. Open Virtual Machines Placement Algorithms

Participants: Fabien Hermenier, Vincent Kherbache, Huynh Tu Dang.

Clients of IaaS providers are looking for dependable infrastructures that can cope with their SLA requirements. To stay attractive, a cloud must then rely on a Virtual Machine (VM) placement algorithm with features matching wrt the SLAs expectations. These constraints are however very specific to each of the tenants but also the infrastructure. They also cover a large range of concerns (reliability, performance, security, energy-efficiency, ...) that are continuously evolving according to new trends and new technologies. To address these issue, we advocate for a flexible VM placement algorithm that can be specialized through plugins to address new concerns.

This year, we first validate our approach with BtrPlace, a composable VM placement algorithm built over Constraint Programming [9]. The usage of Constraint Programming makes placement constraints independent of each other. New constraints can be added without changing the existing implementation. The expressivity of BtrPlace has been verified by implementing more than 20 placement constraints that reproduce, extend but also bring new meaningful restrictions on the VM placement with regards to constraints available in commercial placement algorithm. Each constraint was implemented by an average of 30 lines of Java code. An experienced developer implemented some of the them in half a day, while external developers, without any background in CP, have implemented constraints related to power efficiency [43].

Secondly, we exhibited a lack of reliability in the common approach to address placement constraints in some algorithms. Usually, a constraint controls the VM placement only at the end of the reconfiguration process and ignores the datacenter intermediary states between the beginning and the end of the reconfiguration process. In [11], we advocated that this discrete approach is not sufficient to satisfy the SLAs continuously as an uncontrolled actions schedule may indeed lead to temporary violations. We relied on the flexibility provided by BtrPlace to exhibit these violations and to propose *continuous constraints* to control the quality of service at any moment. We implemented preliminary version of continuous constraints and confirmed they improve the datacenter reliability by removing any temporary violations.

6.2.3. GPU-based High Performance Cloud Computing

Participants: Michael Benguigui, Françoise Baude, Fabrice Huet.

To address HPC, GPU devices are now considered as unavoidable cheap, energy efficient, and very efficient alternative computing units. The barrier to handle such devices is the programming model: it is both very fine grained and synchronous. Our long term goal is to devise some generic solutions in order to incorporate GPU-specific code whenever relevant into a parallel and distributed computation. The first step towards this objective was to gain some insight on how to efficiently program a non trivial but well known algorithm. Our previous work [40] highlights the necessity to target a GPU rather than distributed CPUs to provide the same performance level. By this way we price complex American basket options through the Picazo pricing algorithm, in the same order of time than a CPU cluster implementation on a 64-core cluster. This year, we achieved the following tasks

- We proposed a multi GPU based implementation of this method, allowing pricing time to fall below 1 hour on 18 GPUs, for a 40-assets American option [14].
- We are currently designing a task dispatching model to load balance tasks in a CPU-GPU cluster. This will allow us to drastically lower the overall computation time of a portfolio estimation, and moreover, the computation time of the Monte Carlo value at risk of a portfolio of complex assets.

6.2.4. MapReduce Based Frameworks for Big Data

Participants: Fabrice Huet, Ge Song.

MapReduce is a programming model which allows the processing of vast amounts of data in parallel, on a large number of machines. It is particularly well suited to static or slow changing set of data since the execution time of a job is usually high. However, in practice data-centers collect data at fast rates which makes it very difficult to maintain up-to-date results. To address this challenge, we propose in [25] a generic mechanism for dealing with dynamic data in MapReduce frameworks. Long-standing MapReduce jobs, called *continuous Jobs*, are *automatically* re-executed to process new incoming data at a minimum cost. We present a simple and clean API which integrates nicely with the standard MapReduce model. Furthermore, we describe cHadoop, an implementation of our approach based on Hadoop which does not require modifications to the source code of the original framework. Thus, cHadoop can quickly be ported to any new version of Hadoop. We evaluate our proposal with two standard MapReduce applications (WordCount and WordCount-N-Count), and one real world application (RDF Query) on real datasets. Our evaluations on clusters ranging from 5 to 40 nodes demonstrate the benefit of our approach in terms of execution time and ease of use.

Another important point is the difficulty to predict the performance of a MapReduce job. This is particularly important when using pay-as-you-go resources such a Cloud. We have proposed a simple framework to predict the performance of Hadoop jobs. It is composed of a dynamic light-weight Hadoop job analyzer, and a prediction module using locally weighted regression methods. Our framework makes some theoretical cost models more practical, and also fits well with jobs and clusters diversity. It can also help those users who want to predict the cost when applying for an on- demand cloud service.

6.3. Application Domains

6.3.1. Publish-Subscribe in Distributed Environments

Participants: Françoise Baude, Fabrice Huet, Laurent Pellegrino, Bastien Sauvan, Iyad Alshabani, Maeva Antoine, Amjad Alshabani.

In the context of the FP7 STREP PLAY and French SocEDA ANR research projects we have developed a middleware dubbed EventCloud (Section 5.5). This last aims to store and retrieve Resource Description Framework (RDF) data but also to relay them to interested parties through a publish/subscribe layer that allows the formulation of content-based subscriptions. Content-based subscriptions are automatically deduced from more complex rules deployed onto a Complex Event Processing engine, the aim of these CEP rules being to trigger new (complex) events after detecting interesting situations [24]. The EventCloud architecture relies on a CAN structured P2P overlay network we initially designed and implemented for the former SOA4ALL FP7-IP project [44].

This year we continued to improve the performances of the EventCloud middleware and its usability as a standalone component but also as a component integrated within the previous projects' platform. Concretely, we proposed a new publish/subscribe matching algorithm for RDF events made of several related RDF triples, that was thoroughly presented in [31] and [22]. To further improve performance, we pursue some efforts to finalize the usage of the newest multi-active object library (cf. Section 6.1.1). Also, to handle more efficiently multicast messaging, we replaced our initial and naive solution with the optimal one presented in Section 6.2.1 . Finally, we proposed a solution for managing multiple EventCloud instances on various cloud platforms, especially for the integration of our middleware in the PLAY and SocEDA platforms (whose latest assessment can be found in [32]). Details about EventCloud management are provided in [29].

Since RDF resources have the property to be poorly balanced, we are also investigating new algorithms that decrease load imbalance for events and data.

6.3.2. Large-scale Simulation Platform: Techniques and methodologies

Participants: Olivier Dalle, E. Mancini, Damian Vicino.

In the domain of simulation techniques and methodologies, this year, we conducted research in the two following areas:

Distributed Network Simulation NetStep[16], is a prototype we developped for the distributed simulation of very large scale network simulations, such as the simulation of peer-to-peer applications. We use simulation micro-steps as a means for optimizing the overlap of communications and computations, without changing the original event-driven model. As a consequence, NetStep allows for the reuse of unmodified existing sequential simulators for building large-scale distributed simulations: the overall simulation is divided both in time and space, into a large number of simulation micro-steps, each of which being executed by a legacy sequential simulator. By choosing the time-step smaller than the minimal look-ahead due to communications, we avoid the need for synchronization between logical processes (LPs) during the simulation. Instead, the simulated communications become inputs and outputs of the simulation micro-steps, and are routed in parallel between LPs by a NetStep dedicated entity. Our prototype is based on the SimGrid sequential simulator.

Discrete Time Representation The representation of time in simulations is a long standing issue, for which many solutions and formalisms have been proposed. However, once the formalism is chosen, the implementation of the time representation is still a non trivial problem: Integer values have a limited range and require the selection of a minimal fixed step that does not support well the multi-scale models; Floating Points numbers have numerous limitations and hidden effects such as rounding due to quantization; those issues result in inaccuracies or even timing errors. In collaboration with our partner in the DISSIMINET Associated Team, we have started a new research on this topic. This research will be released in the form of a new Discrete Event Simulation engine library for the DEVS formalism, designed to fully exploit the 2011 C++ standard; it is candidate for inclusion in the BoostC++ Libraries.

OPALE Project-Team

6. New Results

6.1. Mathematical analysis and control of macroscopic traffic flow models

6.1.1. Vehicular traffic

Participants: Alessandra Cabassi, Maria Laura Delle Monache, Paola Goatin, Alexandre Bayen [UC Berkeley, CA, USA], Legesse Lemecha Obsu [Addis Ababa University, Ethiopia].

In collaboration with UC Berkeley, and as part of the Associated Team ORESTE activity (see http://www-sop.inria.fr/members/Paola.Goatin/ORESTE/index.html), we have proposed a new junction model for ramp metering: we introduce a coupled PDE-ODE model, in which the PDE describes the evolution of the cars flow on the main lane and the ODE describes the evolution of the queue length on the on-ramp, modeled by a buffer, which ensures that boundary conditions are satisfied in strong sense. We were able to prove existence and uniqueness of the solution of the corresponding Riemann problem [41]. Relying on the above junction model, we have applied the Discrete Adjoint Method to efficiently compute (locally) optimal ramp-metering parameters to minimize the total travel time on a stretch of highway [80].

In parallel, we have proposed two optimization strategy for instantaneous optimization of total travel times and total waiting times at roundabouts, which give an estimate of the time spent by drivers on the network section. These cost functionals are minimized with respect to the right-of-way parameter of the incoming roads. For each cost functional, the analytical expression is given for each junction, see [72]. This work is part of L.L. Obsu's PhD thesis.

Finally, we designed a new finite volume algorithm to track the trajectory of a bus in the surrounding traffic using a locally non-uniform moving mesh, see [3, 4, 5].

As part of our TRAM3 activity, we also organized the workshop "TRAM2 - Traffic Modeling and Management: Trends and Perspectives", which successfully took place at Inria Sophia Antipolis on March 20-22, 2013 (see https://team.inria.fr/opale/workshop-tram2/).

In the framework of the EIT ITC Labs Multimodal Mobility activity, A. Cabassi's internship was devoted to the calibration and the validation of a first order traffic flow model against processed real data provided by the industrial partners Autoroutes Traffic and VINCI Autoroutes, see [69].

6.1.2. Crowd motion

Participants: Régis Duvigneau, Paola Goatin, Matthias Mimault, Debora Amadori [L'Aquila University, Italy], Christophe Chalons [LJLL, UP7], Massimiliano D. Rosini [ICM, Warsaw University, Poland], Nicolas Seguin [LJLL, UPMC], Monika Twarogowska.

From the analytical point of view, we have been studying the properties of some models in one space dimension. Concerning Hughes' scalar model, we have established a partial existence result in collaboration with D. Amadori and M.D. Rosini (see [75]). M. Mimault's internship in 2012 was devoted to develop a MATLAB code based on wave-front tracking to compute the solutions of Hughes' model with generalized running cost, see [42]. He is currently working on a mixed hyperbolic-elliptic 2x2 system of conservation laws describing two groups of people moving in opposite directions. Finally, in collaboration with C. Chalons and N. Seguin, we generalized previous results on conservation laws with local flux constraints [3], [5] to general flux functions and non-classical solutions arising in pedestrian flow modeling, see [39]. From the numerical point of view, we have implemented some macroscopic models in 2D on unstructured triangular meshes on the Num3sis platform. We provided a comparison between first and second order models in reproducing complex dynamics of crowd motion, such as formation of stop-and-go waves and clogging at bottlenecks. Then, we concentrated on the higher-order model and analyzed the dependence of the behavior of its solutions on some of the parameters of the system. In particular, we produced some examples where placing obstacles in front of the door prevents from blocking and decreases the evacuation time, see [73], [81].

The above researches were partially funded by the ERC Starting Grant "TRAM3 - Traffic management by macroscopic models".

6.2. Optimum design and control in fluid dynamics and its couplings

In computational sciences for physics and engineering, Computational Fluid Dynamics (CFD) are playing one of the major roles in the scientific community to foster innovative developments of numerical methodologies. Very naturally, our expertise in compressible CFD has led us to give our research on numerical strategies for optimum design a particular, but not exclusive focus on fluids.

The framework of our research aims to contribute to numerical strategies for PDE-constrained multi-objective optimization, with a particular emphasis on CPU-demanding computational applications in which the different criteria to be minimized (or reduced) originate from different physical disciplines that share the same set of design variables. These disciplines are often fluids, as a primary focus, coupled with some other disciplines, such as structural mechanics.

Our approach to *competitive optimization* is focused on the two-discipline problem. It is based on a particular construction of *Nash games*, relying on a *split of territory* in the assignment of individual strategies. A methodology has been proposed for the treatment of two-discipline optimization problems in which one discipline, the primary discipline, is preponderant, or fragile. Then, it is recommended to identify, in a first step, the optimum of this discipline alone using the whole set of design variables. Then, an orthogonal basis is constructed based on the evaluation at convergence of the Hessian matrix of the primary criterion and constraint gradients. This basis is used to split the working design space into two supplementary subspaces to be assigned, in a second step, to two virtual players in competition in an adapted Nash game, devised to reduce a secondary criterion while causing the least degradation to the first. The formulation has been proved to potentially provide a set of Nash equilibrium solutions originating from the original single-discipline optimum point by smooth continuation, thus introducing competition gradually [53]. (see also subsectionsubsect:helico).

Our approach to *cooperative optimization*, in theory, is not limited in number of objective functions. It is based on a result of convex analysis established for a general unconstrained mult-iobjective problem in which all the gradients are assumed to be known. The theorem [16] states that in the convex hull of the gradients, there exists a unique vector of minimal norm, ω; if it is nonzero, the vector ω is a descent direction common to all criteria; otherwise, the current design point is Pareto-stationary. This result led us to generalize the classical steepest-descent algorithm by using the vector ω as search direction. We refer to the new algorithm as the multiple-gradient descent algorithm (MGDA). The MGDA yields to a Pareto-stationary point, and actual Pareto-optimality is then tested [54] (see also subsection 6.2.4).

The two approaches have been combined to explore the Pareto front segment-wise as illustrated on Figure 2 .

6.2.1. Multiple-Gradient Descent Algorithm (MGDA)

Participants: Jean-Antoine Désidéri, Régis Duvigneau, Matteo Giacomini, Abderrahmane Habbal, Adrien Zerbinati.

6.2.1.1. Theory and numerical experimentation of the MGDA construction

In multi-objective optimization, the knowledge of the Pareto set provides valuable information on the reachable optimal performance. A number of evolutionary strategies (PAES, NSGA-II, etc), have been proposed in the literature and proved to be successful to identify the Pareto set. However, these derivative-free algorithms are very demanding in terms of computational time. Today, in many areas of computational sciences, codes are developed that include the calculation of the gradient, cautiously validated and calibrated.

The notion of Pareto-stationarity, originally established to be a necessary condition of optimality in differentiable multi-objective optimization of unconstrained problems, has been extended to problems subject to equality constraints. On this basis, we were able to establish that by augmenting, in a classical manner, the objective-functions of a penalty term equal to the square of the constraint violation, and applying the standard MGDA to it, would result in converged solutions that are Pareto-stationary in the extended sense. Numerical experimentation on this is on-going.

../../../../projets/opale/IMG/nash-mgda.png

Figure 2. ***Two-discipline optimization of a generic geometry of a supersonic aircraft, for concurrent drag and sonic-boom reduction*** *(from A. Minelli's doctoral thesis). The wave drag is calculated by the ONERA elsA code in 3D finite-volume Eulerian flow mode over a 6M-node mesh and the sonic boom using a three-layer approach. The Nash-game paths have been devised by appropriate territory splitting in order to be tangent to the Pareto front, and they are interrupted whenever the Pareto-stationarity condition is judged excessively violated. The MGDA paths converge rapidly back to the front. The simulation demonstrates how the two algorithms complement each other and provide a potential for a piecewise description of the Pareto front, evaluated more economically than a stochastic algorithm operating on a large population.*

6.2.1.2. Meta-model-assisted CFD optimization by MGDA

Using MGDA in a multi objective optimization problem requires the evaluation of a large number of points with regard to criteria, and their gradients. In the particular case of a CFD problems, each point evaluation is very costly since it involves a flow computation, possibly the solution of an adjoint-equation. To alleviate this difficulty, we have proposed to construct meta-models of the functionals of interest (lift, drag, etc) and to calculate approximate gradients by local finite differences. These meta-models are updated throughout the convergence process to the evaluation of the new design points by the high-fidelity model, here the 3D compressible Euler equations.

This variant of MGDA has been tested successfully over a problem of external aerodynamic optimum-shape design of an aircraft wing consisting of reducing wave-drag, and augmenting lift. After only a few cycles of database updates, the Pareto front visibly forms, and this result is achieved at a very moderate computational cost [68]. This variant has been extended successfully to an internal flow optimization problem related to an automobile air-conditioning system and governed by the Navier-Stokes equations. This more difficult problem has been proposed by Renault within the OMD2 ANR project. These studies have been reported in A. Zerbinati's doctoral thesis [38].

6.2.1.3. Exact shape gradients

MGDA has successfully been tested over a two-objective optimization problem governed by two-dimensional elasticity. The deformation of a plate is calculated using an isogeometric approximation (see 6.3) and compliance derived from it. The exact parametric shape gradient is calculated, yielding the gradient of the objective function in two antagonistic situations differing by the loading. Pareto-fronts are thus identified.

6.2.1.4. Perspectives

MGDA offers the possibility to handle in a rational way several objective-functions for which gradients are known or approximated concurrently. This potential opens methodological paths to several themes of interest in high-fidelity simulation-based optimization: optimization of complex systems whose performance is evaluated w.r.t. several criteria originating from different, coupled disciplines; optimization under uncertainties, by introducing sensitivities as additional objectives; optimization of time-dependent systems, such as optimization of flow-control devices that generate a periodic flow (see next subsection), by converting the problem into a multi-point problem by time-discretization of the time and parameter-dependent functional; etc.

6.2.2. Flow control

Participants: Régis Duvigneau, Jérémie Labroquère, Emmanuel Guilmineau [Ecole Centrale de Nantes].

Shape optimization methods are not efficient to improve the performance of fluid systems, when the flow is characterized by a strong unsteadiness related to a massive detachment. This is typically the case for the flow around an automotive body or a wing in stall condition. To overcome this difficulty, flow control strategies are developed, that aim at manipulating vortex dynamics by introducing some active actuators, such as periodic blowing/suction jets. In this context, the choice of the control parameters (location, amplitude, frequency) is critical and not straightforward. Therefore, we develop a methodology to determine optimal control parameters by coupling the simulation of unsteady actuated flows with optimization algorithms. Two research axes have been considered :

- the resolution of the unsteady sensitivity equations derived from the state equations, to exhibit the dependency of the flow dynamics with respect to the control ;
- the optimization of control parameters using a statistical metamodel-based strategy.

In this perspective, unsteady Reynolds Averaged Navier-Stokes equations are solved, with some turbulence closures. Different models for synthetic jet have been implemented to simulate the actuation, and then validated for different turbulence closures [70].

Specific developments have be carried out in the metamodel-based optimizer to include a noise term into Gaussian Process model, which is used to filter errors arising from unsteady simulations. A systematic assessment of modeling and numerical errors has been archived [57], for a backward facing step test-case, with the objective of controlling the re-attachment point location.

This activity is conducted in collaboration with the CFD team of Ecole Centrale de Nantes.

6.2.3. Robust design

Participants: Jean-Antoine Désidéri, Régis Duvigneau, Daïgo Maruyama.

This work aims at developing robust design tools for aircraft w.r.t. aerodynamic performance subject to uncertainties, arising from geometrical features and fluctuations of inflow conditions. The robust design process is considered as a multi-objective optimization problem, which consists in minimizing or maximizing statistical moments of the cost function.

In the context of airfoil design, MGDA is used to improve simultaneously the mean and variance of the lift and drag coefficients, yielding a four-objective optimization problem [71].

6.2.4. Sonic boom reduction

Participants: Gérald Carrier [Research Engineer, ONERA/DAAP], Jean-Antoine Désideri, Andrea Minelli, Itham Salah El Din [Research Engineer, ONERA/DAAP].

When an aircraft flies at supersonic speed, it generates at ground level an N-shaped shock structure which can cause serious environmental damage (“sonic boom”). Thus a problem of interest in aerodynamic optimization is to design such an aircraft to reduce the intensity of the sonic boom while maintaining the aerodynamic performance (drag minimization under lift constraint). Andrea Minelli aimed at contributing to this two-discipline optimization problem. In the first part of his work, an inverse problem has been formulated and solved for “shaped sonic boom” and found in excellent agreement with the George-Seebass-Darden theory [82] for the calculation of the Whitham function corresponding to the lowest-boom (axisymmetric) shape. Method and results have been generalized to more general geometries and have been presented internationally in [58].

Besides, aero-acoustic optimizations have been realized successfully by coupling the aerodynamic optimizer (based on Euler calculations by the elsA software) with the sonic-boom computation in a Nash game formulation. These experiments, conducted with our optimization platform FAMOSA, have demonstrated that starting from the shape optimized aerodynamically, one could retrieve smoothly a shape corresponding to nearly-optimal sonic-boom reduction [36]. and [54].

6.2.5. Helicopter rotor blade optimization in both situations of hovering and forward flight

Participants: Michel Costes [Research Engineer, ONERA/DAAP], Jean-Antoine Désideri, Arnaud Le Pape [Research Engineer, ONERA/DAAP], Enric Roca Leon.

E. Roca Leon is conducting a CIFRE thesis supported by EUROCOPTER (Marignane) at ONERA DAAP. This thesis follows the doctoral thesis of A. Dumont in which the adjoint-equation approach was used to optimize a rotor blade in hovering flight. The goal of this new thesis is to solve a two-objective optimization problem in which the hovering-flight criterion is considered preponderant, but a new criterion that takes into account the forward-flight situation is also introduced, concurrently. The second criterion is the power necessary to maintain the forward motion. The first phase of thesis work has been devoted to the set up of a hierarchy of models from low to high fidelity, in order to calibrate appropriate functional criteria. Then, actual two-objective optimizations are conducted via our Nash game approach to competitive optimization with territory splitting based on reduced Hessian diagonalization. A first successful experiment has been realized in which 16 geometrical parameters have been optimized to reduce the power in forward motion while maintaining sub-optimality of the drag in hover. These results have been accepted for presentation at the American Helicopter Society Forum [62], and [53].

6.2.6. Optimum design in naval hydrodynamics

Participants: Régis Duvigneau, Louis Blanchard, Elisa Berini [K-Epsilon company].

Naval hydrodynamics field has recently shown a growing interest for optimum design methods. The computational context is especially complex because it implies unsteady two-phase turbulent flows, with possibly very high Reynolds number (up to 10^9). The use of automated design optimization methods for such problems requires new developments to take into account the large CPU time necessary for each simulation and the specificity of the geometries considered.

Some developments have been initiated on the geometrical modelling of hull shapes by parametric surfaces. The objective was to be able to modify existing hull shapes by controlling a small number of parameters, that are meaningful for naval architects. We have considered as test-case the bow shape for trawler ships. As a second step, an optimum shape procedure has been set up, based on a metamodel-based optimizer, the developed CAD model and the simulation tool for free-surface flows provided by K-Epsilon company. The objective was to reduce the wave drag of a trawler ship by adding a bow, whose parameters are optimized [50].

6.3. Isogeometric analysis and design

Participants: Régis Duvigneau, Bernard Mourrain [Galaad project-team], Alexandros Ginnis [Nat. Tech. Univ. of Athens], Bernd Simeon [Tech. Univ. of Kaiserslautern], Gang Xu [Hangzhou Dianzi Univ.].

Design optimization stands at the crossroad of different scientific fields (and related software): Computer-Aided Design (CAD), Computational Fluid Dynamics (CFD) or Computational Structural Dynamics (CSM), parametric optimization. However, these different fields are usually not based on the same geometrical representations. CAD software relies on Splines or NURBS representations, CFD and CSM software uses grid-based geometric descriptions (structured or unstructured), optimization algorithms handle specific shape parameters. Therefore, in conventional approaches, several information transfers occur during the design phase, yielding approximations that can significantly deteriorate the overall efficiency of the design optimization procedure. Moreover, software coupling is often cumbersome in this context.

The isogeometric approach proposes to definitely overcome this difficulty by using CAD standards as a unique representation for all disciplines. The isogeometric analysis consists in developing methods that use NURBS representations for geometric modeling, computational domain description and solution basis functions. Using such a unique data structure allows to compute the solution on the exact geometry (not a discretized geometry), obtain a more accurate solution (high-order approximation), reduce spurious numerical sources of noise that deteriorate convergence, avoid data transfers between the software. Moreover, NURBS representations are naturally hierarchical and allows to define multi-level algorithms for solvers as well as optimizers.
In this context, some studies on elliptic problems have been conducted in collaboration with the Galaad project-team and Hangzhou Dianzi University, such as the development of methods for adaptive parameterization including an a posteriori error estimate [46], [47], [48]. A collaborative work has also been carried out with the Technical University of Kaiserslautern, concerning the computation of shape gradients for linear elasticity problems, and with the National Technical University of Athens for hull shape optimization [55].

6.4. Optimum design in structural mechanics

6.4.1. *Shape Optimization in Multidisciplinary Non-Linear Mechanics*

Participants: Aalae Benki, Jean-Antoine Désidéri, Abderrahmane Habbal, Gael Mathis [ArcelorMittal, CRAA].

In collaboration with the ArcelorMittal's Center for Research in Automotive and Applications (CRAA), we study the multidisciplinary shape and parameter design of highly non linear mechanical 2D and 3D structures. We have developed methods adapted to the approximation of Pareto Fronts such as Normal Boundary Intersection NBI and Normalized Normal Constraint Method NNCM. Due to the time consuming cost evaluation, the use of cheap to evaluate surrogate models is mandatory. We have studied the consistency of the approach NBI or NNCM plus surrogates, which turned out to be successful for a broad panel of standard mathematical benchmarks. The coupling is successfully applied to a small scale industrial case, namely the shape optimization of a can bottom vis à vis dome reversal pressure and dome growth criteria. We have then defined a Nash game between criteria where the latter are approximated by the RBF metamodels. First, we validated the computation of a Nash equilibrium for mathematical functions, then we computed Nash equilibria for the small scale industrial case of the shape optimization of the can bottom.

Then, we considered the 3D problem of an automotive twist beam. In this 3D case, we aim to Pareto-optimal shapes for two objectives, the first being to minimize the Von-Mises strain to guarantee the formability of the twist beam, and the second being to maximize the stiffness. For solution with higher stiffness than the initial one, we could decrease the thickness to obtain a mass reduction with the same end-user properties.

We also introduced, to our knowledge for the first time in the structural optimization area, the notion of Kalai-Smorodinky equilibria which is aimed at the selection of equilibria among Pareto-optimal solutions. We applied this notion of equilibria to both industrial cases, and compared the results to Nash equilibria. [56] [64]

../../../../projets/opale/IMG/DRPDG2D.png ../../../../projets/opale/IMG/TwistBeam3D.png

Figure 3. Concurrent design in industrial applications. A packaging problem of commercial cans (left). Automotive twist beam (right)

6.4.2. Optimization of Addendum Surfaces in Stamping

Participants: Fatima Zahra Oujebbour, Rachid Ellaia, Abderrahmane Habbal, Ziheng Zhao.

Within the OASIS Consortium (ArcelorMittal, ErDF, Inria, UTC, EURODECISION, ESILV, NECS, Delta-CAD, SCILAB-DIGITEO), the Opale project-team leads the Optimization task. Our aim is to develop decentralized decision-making algorithms dedicated to find efficient solutions (Pareto optimal) in a complex multi-disciplinary framework (forming, stamping, welding non-linear processes, spring-back, vibration, in-function linear processes, crash and fatigue non linear and non differentiable processes) for several (between three and five) criteria. An important difficulty when trying to identify the Pareto Front, even when using adapted methods such the Normal Boundary Intersection, is that the criteria involved (thanks to the high nonlinearity in the mechanical models) exhibit many local optima. So one must use global optimization methods. We have studied the hybrid approach Simulated Annealing with Simultaneous Perturbation SASP for a suite of mathematical test-cases. To envisage the application of our method to the complex CPU time consuming stamping process, we lead an intermediate phase dedicated to the validation of the SASP method for the minimization of the spring-back that follows the stamping of a metal sheet, the design variable being the process parameters (two then four parameters). Then, we considered the more complex shape design of the initial blank. The initial blank design is a critical step in stamping design procedure, therefore it should be optimally designed. Our aim is to find the optimal initial blank shape that avoids or at least minimizes the springback and failure flaws. For this study, the geometry of the blank contour is described by parametric spline curves. Seven control points (P1,...,P7) are used to define the spline curves in order to have a wide variety of geometries. The exact computational evaluation of our criteria, springback and failure, is very expensive (the FE model request around 45 min to predict these two criteria) and the design space is of quite high dimension. Therefore, we considered the recourse to the sparse grid interpolation. Optimization process based on sparse grid interpolation is an optimal alternative in which criteria can be approximated with a suitable interpolation formula that needs significantly less points than the full grid. the obtained metamodel using sparse grid interpolation needs less than 1s to predict springback and failure on the same computation machine. To find the optimal initial blank shape, it was decided to perform the optimization process using the obtained metamodel. The construction of the sparse grid interpolant was based on the Chebyshev Gauss-Lobatto grid type and using the polynomial basis functions. This technique achieves a good accuracy with a competitive number of grid points. The comparison of the obtained fronts shows that we can capture Pareto solutions by NBI and NNCM with fewer points than NSGAII which requires a large number of populations and several generations to obtain the Pareto front. [60] [61] [63] [77]

../../../../projets/opale/IMG/embout1.png ../../../../projets/opale/IMG/embout2.png ../../../../projets/opale/IMG/BlankS

../../../../projets/opale/IMG/NBI.png ../../../../projets/opale/IMG/NNCM.png ../../../../projets/opale/IMG/NSGAII

Figure 4. Multiobjective design of the stamping process of a high performance steel sheet. The design variable is the initial blank shape, and the costs are elastic spring-back and failure. Sparse grid approximation of the costs is used. The Pareto front obtained by NBI and NNCM (lower-left) are compared to a NSGA-II one (lower-right).

6.5. Application of shape and topology design to biology and medicine

6.5.1. Assessing the ability of the 2D Fisher-KPP equation to model cell-sheet wound closure

Participants: Abderrahmane Habbal, Hélène Barelli [Univ. Nice Sophia Antipolis, CNRS, IPMC], Grégoire Malandain [Inria, EPI Morpheme].

We address in this joint collaboration the ability of the widely used Fisher-KPP equations to render some of the dynamical features of epithelial cell-sheets during wound closure.

Our approach is based on nonlinear parameter identification, in a two-dimensional setting, and using advanced 2D image processing of the video acquired sequences. As original contribution, we lead a detailed study of the profiles of the classically used cost functions, and we address the "wound constant speed" assumption, showing that it should be handled with care.

We study five MDCK cell monolayer assays in a reference, activated and inhibited migration conditions. Modulo the inherent variability of biological assays, we show that in the assay where migration is not exogeneously activated or inhibited, the wound velocity is constant. The Fisher-KPP equation is able to accurately predict, until the final closure of the wound, the evolution of the wound area, the mean velocity of the cell front, and the time at which the closure occurred. We also show that for activated as well as for inhibited migration assays, many of the cell-sheet dynamics cannot be well captured by the Fisher-KPP model. Original unexplored utilizations of the model such as wound assays classification based on the calibrated diffusion and proliferation rate parameters is ongoing.[49] [76]

(a) (b) (c)

(d) (e) (f)

Figure 5. ***A regular wound assay*** *(a) Time evolution of wound area (in pixel). (b) Time evolution of the leading-edge length (in pixel). (c) 3D XT view at first and mid-rows. (d) Mean (in time) velocity of pixels located at the leading edge (in pixel/min). (e) Averaged (in space) leading-edge velocity (in pixel/min). (f) 2D XT view at first and mid-rows.*

(a) (b) (c)

Figure 6. ***A regular wound assay. Computational vs experimental wound evolution.*** *(a) Time variation of experimental (blue) versus computed (red) wound area (in pixel). (b) Time variation of the experimental (blue-dot) versus computed (red) migration rate (in pixel/min). (c) 3D XT view at first and mid-rows.*

(a) (b) (c)

Figure 7. ***An accelerating activated wound assay. Computational vs experimental wound evolution.*** *(a) Time variation of experimental (blue) versus computed (red) wound area (in pixel). (b) Time variation of the experimental (blue-dot) versus computed (red) migration rate (in pixel/min). (c) 3D XT view at first and mid-rows.*

ORPAILLEUR Project-Team

6. New Results

6.1. The Mining of Complex Data

Participants: Mehwish Alam, Aleksey Buzmakov, Melisachew Chekol, Victor Codocedo, Adrien Coulet, Elias Egho, Nicolas Jay, Florence Le Ber, Ioanna Lykourentzou, Luis-Felipe Melo, Amedeo Napoli, Chedy Raïssi, Mohsen Sayed, My Thao Tang, Mohsen Sayed, Yannick Toussaint.

Keywords: formal concept analysis, relational concept analysis, pattern structures, frequent itemset, association rule, graph mining, sequence mining, skyline

Formal Concept Analysis, together with itemset search and association rule extraction, are suitable symbolic methods for KDDK, that may be used for real-sized applications. Global improvements are carried on the scope of applicability, the ease of use, the efficiency of the methods, and on the ability to fit evolving situations. Accordingly, the team is extending these symbolic data mining methods for working on biological or chemical data or textual documents, involving objects with multi-valued attributes (e.g. domains or intervals), n-ary relations, sequences, trees and graphs.

6.1.1. FCA and variations: RCA and Pattern Structures

There are a few extensions of FCA for handling contexts involving complex data formats, e.g. graphs or relational data. Among them, Relational Concept Analysis (RCA) is a process for analyzing objects described both by binary and relational attributes [10]. The RCA process takes as input a collection of contexts and of inter-context relations, and yields a set of lattices, one per context, whose concepts are linked by relations. RCA has an important role in KDDK, especially in text mining [86], [85].

Another extension of FCA is based on Pattern Structures (PS) [92], which allows to build a concept lattice from complex data, e.g. nominal, numerical, and interval data. In [100], pattern structures are used for building a concept lattice from interval data. Since then, we worked on a some experiments involving pattern structures, namely sequence mining [41], information retrieval [48] and functional dependencies [38]. one of the next step is the adaptation of pattern structures to graph mining. Moreover, the notion of similarity between objects is also closely related to pattern structures [99]: two objects are similar as soon as they share the same attributes (binary case) or attributes with similar values or the same description (at least in part). Combination of similarity and pattern structures is also under study, in particular for solving information retrieval and annotation problems.

Finally, there is also an on-going work relating FCA and semantic web. This work focuses on the classification within a concept lattice of the answers returned by SPARQL queries [37], [47], [46], [44]. The concept lattice is then used as an index for navigating and ranking the answers w.r.t. their content and interest for a given objective.

6.1.2. Advances in mining complex data: sequences and healthcare trajectories

Sequence data is widely used in many applications. Consequently, mining sequential patterns and other types of knowledge from sequence data has become an important data mining task. The main emphasis has been on developing efficient mining algorithms and effective pattern representation. The most frequent sequences generally provide a trivial information. When analyzing the set of frequent sequences with a low minimum support, the user is overwhelmed by millions of patterns. In our recent work, the general idea is to extract patterns whose characteristic on a given measure such as the support strongly deviates from its expected value under a null model. The frequency of a pattern is considered as a random variable, whose distribution under the null model has to be calculated or approximated. Then, the significance of the pattern is assessed through a statistical test that compares the expected frequency under the null model to the observed frequency. One of the key-points of this family of approaches is to choose an appropriate null model. It will ideally be a trade-off between adjustment to the data and simplicity: the model should capture some characteristics of the data, to integrate prior knowledge, without overfitting, to allow for relevant patterns discovery. We introduced a rigorous and efficient approach to mine statistically significant, unexpected patterns in sequences of itemsets. Experiments on sequences of replays of a video game demonstrated the scalability and the efficiency of the method to discover unexpected game strategies. This work was successfully published as an international conference paper [8].

Other work on sequences is in concern with patient trajectories, i.e. the "path" of a patient during its illness. With the increasing burden of chronic illnesses, administrative health care databases hold valuable information that could be used to monitor and assess the processes shaping the trajectory of care of chronic patients. In this context, temporal data mining methods are promising tools, though lacking flexibility in addressing the complex nature of medical events. In a set of recent works with Elias Egho, a PhD candidate, we present new algorithms to extract patient trajectory patterns with different levels of granularity by relying on external taxonomies [52]. Our algorithms rely on the general FCA framework to formalize the general notion of multidimensional healthcare trajectories. We also continued working on the complex notion of sequences or trajectory similarity measures. We show the interest of our approaches with the analysis of trajectories of care for colorectal cancer using data from the French healthcare information system (see also [41]).

6.1.3. KDDK in Text Mining

Ontologies help software and human agents to communicate by providing shared and common domain knowledge, and by supporting various tasks, e.g. problem-solving and information retrieval. In practice, building an ontology depends on a number of "ontological resources" having different types: thesaurus, dictionaries, texts, databases, and ontologies themselves. We are currently working on the design of a methodology and the implementation of a system for ontology engineering from heterogeneous ontological resources [58]. This methodology is based on both FCA and RCA, and was previously successfully applied in contexts such as astronomy and biology. In the framework of the ANR Hybride project (see 8.2.1.2), an engineer is implementing a robust system based on these previous research results, for preparing the way to new research directions involving trees and graphs.

6.2. KDDK in Life Sciences

Participants: Yasmine Assess, Emmanuel Bresso, Adrien Coulet, Marie-Dominique Devignes, Anisah Ghoorah, Bernard Maigret, Amedeo Napoli, Gabin Personeni, David Ritchie, Mohsen Sayed, Malika Smaïl-Tabbone, My Thao Tang, Mohsen Sayed, Yannick Toussaint.

The Life Sciences constitute a challenging domain for KDDK. Biological data are complex from many points of views, e.g. voluminous, high-dimensional and deeply inter-connected. Analyzing such data is a crucial issue in health care, environment and agronomy. Besides, many bio-ontologies are available and can be used to enhance the knowledge discovery process. Accordingly, the research work of the Orpailleur team in KDDK applied to the Life Sciences is developed in one main direction which is in concern with the use of bio-ontologies to improve KDDK but also information retrieval, access to the so-called "Linked Open Data" and data integration.

6.2.1. Using ILP for the characterization and prediction of drug side-effect profiles

Inductive Logic Programming (ILP) is a learning method which allows expressive representation of the data and produces explicit first-order logic rules [89]. We applied ILP for understanding drug side-effets. Indeed, late appearance of adverse side effets during clinical trials constitute the main reason for stopping the drug development process which is very costly [1]. Improving our ability to understand drug side effects is necessary to reduce this inconvenience. Moreover, it can contribute to design safer drugs and anticipate the appearance of yet unreported side effects of approved drugs. Today, most investigations deal with prediction of single side effects and overlook possible combinations.

In our study, drug annotations are collected from the SIDER and DrugBank databases. Terms describing individual side effects reported in SIDER are clustered with the IntelliGO semantic similarity measure into term clusters (TCs) [83]. Maximal frequent itemsets are extracted from the resulting $drug \times TC$ binary table, leading to the identification of what we call side-effect profiles (SEPs). A SEP is defined as the longest combination of TCs which are shared by a significant number of drugs. Frequent SEPs are explored on the basis of integrated drug and target descriptors using two machine learning methods: decision-trees and ILP. Learning efficiency is evaluated by cross-validation and direct testing with new molecules. Comparison of the two methods shows that the ILP displays a greater sensitivity than decision trees. Although both methods yield explicit models, ILP is able to exploit not only drug properties but also background knowledge, thereby producing rich and expressive rules.

6.2.2. Functional classification of genes

The IntelliGO measure computes semantic similarity between genes in taking into account domain knowledge in Gene Ontology (GO) [83]. IntelliGO is used for functional clustering of a set of genes, i.e. based on functional annotations of these genes. For example, a gene set of interest may include genes showing the same expression profile.

A functional clustering method based on IntelliGO was tested on four benchmarking datasets consisting of biological pathways (KEGG database) and functional domains (Pfam database) [90]. A follow-up of this study was motivated by the fact that the IntelliGO measure, like most of the biological similarity measures, does not verify "triangle inequality" and thus is not a mathematical distance. Interestingly, specific spectral clustering techniques can be used for improving the clustering of the objects for which exists a pairwise (dis-)similarity matrix [115], [125]. Spectral clustering techniques make use of the eigenvalues of this (dis-)similarity matrix to perform dimension reduction before clustering in fewer dimensions. We have conducted a comparative and large-scale gene clustering evaluation using the IntelliGO measure and reference sets. Our results showed an improvement of the clustering quality with "constant-shift spectral clustering" [63].

6.2.3. Analysis of biomedical data annotated with ontologies

Annotating data with concepts of an ontology is a common practice in the biomedical domain. Resulting annotations define links between data and ontologies that are key for data exchange, data integration and data analysis. Since 2011, we collaborate with the National Center for Biomedical Ontologies (NCBO) to develop a large repository of annotations named the NCBO Resource Index [98]. This repository contains annotations of 36 biomedical databases annotated with concepts of more than 200 ontologies of the BioPortal [3]. In 2012, we compared the annotations of a database of biomedical publications (Medline) with two databases of scientific funding (Crisp and ResearchCrossroads) to profile disease research [105]. One main challenge remains to develop a knowledge discovery approach able to mine correlations between annotations based on BioPortal ontologies, i.e. is it possible to discover interesting knowledge units within these annotations?

In 2013, we proposed an adaptation of FCA techniques, namely pattern structures, to explore the annotations of biomedical databases [2]. We considered documents of biomedical databases annotated with sets of ontological concepts as objects in a pattern structure. Corresponding annotations have been classified according to several dimensions, where a dimension is related to a particular aspect of domain knowledge. Then, the pattern structure formalism was applied to classify these annotations, allowing to discover correlations between annotations but also lacks of completion in the annotations that could be fixed afterward. This adaptation of pattern structures opens many perspectives in term of ontology reengineering and knowledge discovery.

In another context, a related work was carried out in the Kolflow project (see 8.2.1.4). We proposed an interactive environment based on Formal Concept Analysis which makes possible a simultaneous enrichment of semantic annotations of medical texts and of the ontology of medical domain [66], [59].

6.2.4. Analysis and interpretation of sequential patterns with Linked Open Data

[3] http://bioportal.bioontology.org/

Linked Data is a set of principles and technologies that rely on the architecture of the Web (URIs and links) to share, model and integrate data. The basic idea is that data objects (e.g., a surgical procedure) are identified by web addresses (URIs), and the information attached to these objects are represented through links to values or other URIs representing other objects.

Considering the potential development and availability of biomedical Linked Data, we investigated it as a source of additional information to support the interpretation of the results of a data mining process, such as sequential pattern discovery. We developed a system using several linked data endpoints to collect descriptive dimensions about the items that constitute sequential patterns. These dimensions are used to automatically classify with Formal Concept Analysis the extracted patterns, thus generating a structure that can support exploration and navigation into the results of the data mining step [55].

6.3. Structural Systems Biology

Participants: Marie-Dominique Devignes, Anisah Ghoorah, Van-Thai Hoang, Bernard Maigret, David Ritchie, Malika Smaïl-Tabbone.

Keywords: bioinformatics, chemistry, docking, knowledge discovery, screening, systems biology

Structural systems biology aims to describe and analyze the many components and interactions within living cells in terms of their three-dimensional (3D) molecular structures. We are currently developing advanced computing techniques for molecular shape representation, protein-protein docking, protein-ligand docking, high-throughput virtual drug screening, and knowledge discovery in databases dedicated to protein-protein interactions.

6.3.1. Accelerating protein docking calculations using graphics processors

We have recently adapted the *Hex* protein docking software [113] to use modern graphics processors (GPUs) to carry out the expensive FFT part of a docking calculation [114]. Compared to using a single conventional central processor (CPU), a high-end GPU gives a speed-up of 45 or more. This software is publicly available at http://hex.loria.fr. A public GPU-powered server has also been created (http://hexserver.loria.fr) [106]. The docking server has performed some 14,000 docking runs during 2013.

Our docking work has facilitated further developments on modeling the assembly of multi-component molecular structures using a particle swarm optimization technique [123], and on modeling protein flexibility during docking [122]. In 2013, in collaboration with the Nano-D team at Inria Grenoble, we developed a new docking algorithm called “DockTrina” [31], which can rapidly model trimers of protein structures by combining multiple pair-wise docking results from *Hex*. We also used *Hex* successfully to model a challenging protein complex containing water molecules at the protein-protein interface [29].

6.3.2. KBDOCK: Protein docking using Knowledge-Based approaches

In order to explore the possibilities of using structural knowledge of protein-protein interactions, Anisah Ghoorah recently developed the KBDOCK system as part of her doctoral thesis project [95]. KBDOCK is available at http://kbdock.loria.fr. KBDOCK combines coordinate data from the Protein Data Bank [87] with the Pfam protein domain family classification [91] in order to describe and analyze all known protein-protein interactions for which the 3D structures are available. We have demonstrated the utility of KBDOCK [94] for template-based docking using 73 complexes from the Protein Docking Benchmark [97]. We recently presented results obtained using KBDOCK at the CAPRI conference on protein docking in Utrecht [21]. In 2013, we updated KBDOCK with the latest data from Pfam and the Protein Data Bank. An article describing the new version of KBDOCK was accepted by the Database Issue of Nucleic Acids Research [6].

6.3.3. Kpax: A new algorithm for protein structure alignment

We have developed a new protein structure alignment approach called Kpax [112]. The approach exploits the fact that each amino acid residue has a carbon atom with a highly predictable tetrahedral geometry. This allows the local environment of each residue to be transformed into a canonical orientation, thus allowing easy comparison between the canonical orientations of residues within pairs of proteins using a novel scoring function based on Gaussian overlaps. The overall approach is two or three orders of magnitude faster than most contemporary protein structure alignment algorithms, while still being almost as accurate as the state-of-the-art TM-Align approach [124]. The Kpax program is available at http://kpax.loria.fr/. The Kpax program is now used heavily behind the scenes in the new KBDOCK web server [6] to find structural templates for docking which might be beyond the reach of sequence-based homology modeling approaches.

6.3.4. gEMpicker and gEMfitter: GPU-accelerated tools for cryo-electron microscopy

Solving the structures of large protein assemblies is a difficult and computationally intensive task. Multiple two-dimensional (2D) images must be processed and classified to identify protein particles in different orientations. These images may then be averaged and stacked to deduce the three-dimensional (3D) structure of a protein. In order to help accelerate the first of these tasks we have recently developed a novel and highly parallel algorithm called “gEMpicker” which uses multiple graphics processors to detecting 2D particles in cryo-electron microscopy images [112]. We have also developed a 3D shape matching algorithm called “gEMfitter” which also exploits graphics processors, and which will provide a useful tool for the final 3D assembly step [112]. Both programs have been made publicly available at http://gem.loria.fr/.

6.3.5. DOVSA: Developing new algorithms for virtual screening

In 2010, Violeta Pérez-Nueno joined the Orpailleur team thanks to a Marie Curie Intra-European Fellowship (IEF) award to develop new virtual screening algorithms (DOVSA). The aim of this project was to advance the state of the art in computational virtual drug screening by developing a novel consensus shape clustering approach based on spherical harmonic (SH) shape representations [110]. As a continuation of this project, and in collaboration with colleagues from the University of Bari in Italy, we recently published a review on drug discovery relating to the GPCR receptor proteins [15]. We also published a book chapter describing the ParaFit program for fast spherical harmonic shape matching [70].

6.4. Around the Taaable research project

Participants: Valmi Dufour-Lussier, Emmanuelle Gaillard, Laura Infante Blanco, Florence Le Ber, Jean Lieber, Amedeo Napoli, Emmanuel Nauer.

Keywords: knowledge representation, description logics, classification-based reasoning, case-based reasoning, belief revision, semantic web

The Taaable project [69] (http://taaable.fr) has been originally created as a challenger of the Computer Cooking Contest (ICCBR Conference). A candidate to this contest is a system whose goal is to solve cooking problems on the basis of a recipe book (common to all candidates), where each recipe is a shallow XML document with an important plain text part. The size of the recipe book (about 1500 recipes) prevents from a manual indexing of recipes: this indexing is performed using semi-automatic techniques.

Beyond its participation to the CCCs, the Taaable project aims at federating various research themes: case-based reasoning (CBR), information retrieval, knowledge acquisition and extraction, knowledge representation, minimal change theory, ontology engineering, semantic wikis, text-mining, etc. CBR is used to perform adaptation of recipe to user constraints. The reasoning process uses a cooking domain ontology (especially hierarchies of classes) and adaptation rules. The knowledge base used by the inference engine is encoded within a semantic wiki, which contains the recipes, the domain ontology, and adaptation rules.

Minimal change theory and belief revision can be used as tools to support adaptation in CBR, i.e. the source case is modified to be consistent with the target problem using a revision operator. Belief revision was applied to Taaable for the adaptation of recipe preparations [3], using one of the engines included in the library Revisor (cf. § 5.4.5).

As acquiring knowledge from experts is costly, a new approach was proposed to allow a CBR system to use partially reliable, non expert, knowledge from the Web for reasoning [68] [5]. This approach is based on a meta-knowledge model to manage knowledge reliability. This model represents notions such as belief, trust, reputation and quality, as well as their relationships and rules to evaluate knowledge reliability. The reliability estimation is used to filter knowledge with high reliability as well as to rank the results produced by the CBR system, ensuring the quality of results.

6.5. Some results in graph theory

Participants: Amedeo Napoli, Chedy Raïssi, Jean-Sébastien Sereni, Mario Valencia.

Keywords: graph theory, extremal graph theory, coloring, clustering

6.5.1. Structural and extremal graph theory

Regarding graph coloring, a conjecture of Gera, Okamoto, Rasmussen and Zhang on set coloring was solved. A *set coloring* of a graph $G=(V,E)$ is a function $c: V\to\{1,...,k\}$ such that whenever u and v are adjacent vertices, it holds that $\{c(x) : x \text{ neighbor of } u\} \neq \{c(x) : x \text{ neighbor of } v\}$. In other words, there must be at least one neighbor of u that has a color not assigned to a neighbor of v, or *vice-versa*. The smallest k such that G admits a set coloring is the *set coloring number* $\chi_s(G)$. We confirmed the conjecture by proving that $\chi_s(G) \geq \lceil \log_2 \chi(G) \rceil + 1$, where $\chi(G)$ is the (usual) chromatic number of G. This bound is tight.

Works have been started on a 12-year-old conjecture by Heckman and Thomas about the fractional chromatic number of graphs with no triangles and maximum degree at most 3. This conjecture is actually a natural generalization of a fact established by Staton in 1979. Heckman and Thomas posits that in every graph with no triangles, maximum degree at most 3 and arbitrary weights on the vertices, there exists an independent set of weight at least $5/14$ times the total weight of the graph.

Regarding extremal graph theory, two results have been obtained. The first one deals with permutation snarks, while the second one reads as follows.
For every 3-coloring of the edges of the complete graph on n vertices, there is a color c and a set X of 4-vertices such that at least $2n/3$ vertices are linked to a vertex in X by an edge of color c.
This theorem is motivated by a conjecture of Erdős, Faudree, Gould, Gyárfás, Rousseau and Schelp from 1989, which asserts that X can be of size 3 only. However, they were only able to prove that X can be of size 22. Recently, Rahil Baber and John Talbot managed to build upon our work in a very nice article: adding a new idea to our argument, they managed to confirm the conjecture.

6.5.2. Graph theory and other fields

Interactions of graph theory with other topics (theoretical computer science, number theory, group theory, sociology and chemistry) have been considered. Most of them are still in progress and some are published. For instance, regarding distributed computing, the purpose of our work was to question the global knowledge each node is assumed to start with in many distributed algorithms (both deterministic and randomized). More precisely, numerous sophisticated local algorithm were suggested in the literature for various fundamental problems. Noticeable examples are the MIS algorithms and the $(\Delta+1)$-coloring algorithms. Unfortunately, most known local algorithms are *non-uniform*, that is, they assume that all nodes know good estimations of one or more global parameters of the network, e.g., the number of nodes n. Our work provides a rather general method for transforming a non-uniform local algorithm into a uniform one. Furthermore, the resulting algorithm enjoys the same asymptotic running time as the original non-uniform algorithm. Our method applies to a wide family of both deterministic and randomized algorithms. Specifically, it applies to almost all of the state of the art non-uniform algorithms regarding MIS and Maximal Matching, as well as to many results concerning the coloring problem.

6.5.3. Other aspects on graph coloring and clustering

Since September 2013, Mario Valencia has obtained a one year invitation (namely Inria "Délégation") for working at Inria Nancy – Grand Est, in the Orpailleur team, on graph theoretical aspects and data clustering. This research work consists in studying the modular decomposition techniques on the threshold graphs issues of the clustering process. More precisely, this study relies on families of graphs having a "good" decomposition as cographs and chordal graphs, and then, and on the analysis of the adaptation of these two families of graphs within a clustering activity.

Other research dimensions are dealing with algorithmic aspects of some variations of the classical graph coloring problem.

- Packing colorings of graphs where we need to color the vertices of a graph in such a way that vertices having a same color c should be at a distance at least equal to $c+1$ in the graph. With P. Torres, a postdoc student, we have obtained some upper bounds for the packing chromatic number of hypercubes graphs of dimension n, denoted by Q_n, and we have computed exactly this parameter for this family of graph for $n = 6, 7, 8$, extending previous results known for $n = 2, 3, 4, 5$ [35].
- (k, i)-coloring of graphs, which is a generalization of a k-tuple coloring of graphs: given positive integers k and i, we want to affect to each vertex a k-set of colors such that the intersection of the k-sets affected to adjacent vertices has cardinality at most equal to i. With F. Bonomo, I. Koch, and G. Duran, we have found a linear time algorithm for this problem on cycles and cacti graphs. Moreover, we have obtained an interesting equivalence between this problem on complete graphs and a problem on weighted binary codes.
- b-coloring of graphs, where we need to color the vertices of a graph in such a way that in each color class j there exists at least one vertex x_j adjacent to at least one vertex in all the other color classes. The goal of this problem is to maximize the number of colors under such a constraint (i.e. the b-chromatic number of a graph). With F. Bonomo, O. Schaudt and M. Stein, we have shown that b-coloring is NP-hard on co-bipartite graphs and polytime solvable on tree-cographs [77].

PANAMA Project-Team

6. New Results

6.1. Recent results on sparse representations

Sparse approximation, high dimension, scalable algorithms, dictionary design, sample complexity

The team has had a substantial activity ranging from theoretical results to algorithmic design and software contributions in the field of sparse representations, which is at the core of the ERC project PLEASE (projections, Learning and Sparsity for Efficient Data Processing, see section 8.2.1).

6.1.1. A new framework for sparse representations: analysis sparse models

Participants: Rémi Gribonval, Nancy Bertin, Srdan Kitic, Cagdas Bilen.

Main collaboration: Mike Davies, Mehrdad Yaghoobi (Univ. Edinburgh), Michael Elad (The Technion).

In the past decade there has been a great interest in a synthesis-based model for signals, based on sparse and redundant representations. Such a model assumes that the signal of interest can be composed as a linear combination of *few* columns from a given matrix (the dictionary). An alternative *analysis-based* model can be envisioned, where an analysis operator multiplies the signal, leading to a *cosparse* outcome. Within the SMALL FET-Open project, we initiated a research programme dedicated to this analysis model, in the context of a generic missing data problem (e.g., compressed sensing, inpainting, source separation, etc.). We obtained a uniqueness result for the solution of this problem, based on properties of the analysis operator and the measurement matrix. We also considered a number of pursuit algorithms for solving the missing data problem, including an ℓ^1-based and a new greedy method called GAP (Greedy Analysis Pursuit). Our simulations demonstrated the appeal of the analysis model, and the success of the pursuit techniques presented.

These results have been published in conferences and in a journal paper [19]. Other algorithms based on iterative cosparse projections [83] as well as extensions of GAP to deal with noise and structure in the cosparse representation have been developed, with applications to toy MRI reconstruction problems and acoustic source localization and reconstruction from few measurements [100].

Successful applications of the cosparse approach to sound source localization, audio declipping and brain imaging have been developed this year. In particular, we compared the performance of several cosparse recovery algorithms in the context of sound source localization [39] and showed its efficiency in situations where usual methods fail [60]. It was also shown to be applicable to the hard declipping problem [61]. Application to EEG brain imaging was also investigated and a paper was submitted to ICASSP'14 (see below).

6.1.2. Theoretical results on sparse representations

Participants: Rémi Gribonval, Anthony Bourrier, Pierre Machart.

Main collaboration: Charles Soussen (Centre de recherche en automatique de Nancy (CRAN)), Jérôme Idier (Institut de Recherche en Communications et en Cybernétique de Nantes (IRCCyN)), Cédric Herzet (Equipe-projet FLUMINANCE (Inria - CEMAGREF, Rennes)), Mehrdad Yaghoobi, Mike Davies (University of Edinburgh), Patrick Perez (Technicolor R&I France), Tomer Peleg (The Technion)

Sparse recovery conditions for Orthogonal Least Squares : We pursued our investigation of conditions on an overcomplete dictionary which guarantee that certain ideal sparse decompositions can be recovered by some specific optimization principles / algorithms. We extended Tropp's analysis of Orthogonal Matching Pursuit (OMP) using the Exact Recovery Condition (ERC) to a first exact recovery analysis of Orthogonal Least Squares (OLS). We showed that when ERC is met, OLS is guaranteed to exactly recover the unknown support. Moreover, we provided a closer look at the analysis of both OMP and OLS when ERC is not fulfilled. We showed that there exist dictionaries for which some subsets are never recovered with OMP. This phenomenon, which also appears with ℓ^1 minimization, does not occur for OLS. Finally, numerical experiments based on our theoretical analysis showed that none of the considered algorithms is uniformly better than the other [21]. More recently, we obtained simpler coherence-based conditions [18] and pursued the analysis of unrecoverable subsets [43].

Fundamental performance limits for ideal decoders in high-dimensional linear inverse problems: The primary challenge in linear inverse problems is to design stable and robust "decoders" to reconstruct high-dimensional vectors from a low-dimensional observation through a linear operator. Sparsity, low-rank, and related assumptions are typically exploited to design decoders which performance is then bounded based on some measure of deviation from the idealized model, typically using a norm. We characterized the fundamental performance limits that can be expected from an ideal decoder given a general model, ie, a general subset of "simple" vectors of interest. First, we extended the so-called notion of instance optimality of a decoder to settings where one only wishes to reconstruct some part of the original high dimensional vector from a low-dimensional observation. This covers practical settings such as medical imaging of a region of interest, or audio source separation when one is only interested in estimating the contribution of a specific instrument to a musical recording. We defined instance optimality relatively to a model much beyond the traditional framework of sparse recovery, and characterized the existence of an instance optimal decoder in terms of joint properties of the model and the considered linear operator [42], [33]. Noiseless and noise-robust settings were both considered [56]. We showed somewhat surprisingly that the existence of noise-aware instance optimal decoders for all noise levels implies the existence of a noise-blind decoder. A consequence of our results is that for models that are rich enough to contain an orthonormal basis, the existence of an L2/L2 instance optimal decoder is only possible when the linear operator is not substantially dimension-reducing. This covers well-known cases (sparse vectors, low-rank matrices) as well as a number of seemingly new situations (structured sparsity and sparse inverse covariance matrices for instance). We exhibit an operator-dependent norm which, under a model-specific generalization of the Restricted Isometry Property (RIP), always yields a feasible instance optimality and implies instance optimality with certain familiar atomic norms such as the ℓ^1 norm.

Connections between sparse approximation and Bayesian estimation: Penalized least squares regression is often used for signal denoising and inverse problems, and is commonly interpreted in a Bayesian framework as a Maximum A Posteriori (MAP) estimator, the penalty function being the negative logarithm of the prior. For example, the widely used quadratic program (with an ℓ^1 penalty) associated to the LASSO / Basis Pursuit Denoising is very often considered as MAP estimation under a Laplacian prior in the context of additive white Gaussian noise (AWGN) reduction.

In 2011 we obtained a result [85] highlighting the fact that, while this is *one* possible Bayesian interpretation, there can be other equally acceptable Bayesian interpretations. Therefore, solving a penalized least squares regression problem with penalty $\phi(x)$ need not be interpreted as assuming a prior $C \cdot \exp(-\phi(x))$ and using the MAP estimator. In particular, we showed that for *any* prior P_X, the minimum mean square error (MMSE) estimator is the solution of a penalized least square problem with some penalty $\phi(x)$, which can be interpreted as the MAP estimator with the prior $C \cdot \exp(-\phi(x))$. Vice-versa, for *certain* penalties $\phi(x)$, the solution of the penalized least squares problem is indeed the MMSE estimator, with a certain prior P_X. In general $dP_X(x) \neq C \cdot \exp(-\phi(x))dx$. This year, we extended this result to general inverse problems [30], [58], [47].

6.1.3. Algorithmic and theoretical results on dictionary learning

Participants: Rémi Gribonval, Nancy Bertin, Cagdas Bilen, Srdan Kitic.

Main collaboration: Rodolphe Jenatton, Francis Bach (Equipe-projet SIERRA (Inria, Paris)), Martin Kleinsteuber, Matthias Seibert (TU-Munich), Mehrdad Yaghoobi, Mike Davies (University of Edinburgh),

Dictionary learning : An important practical problem in sparse modeling is to choose the adequate dictionary to model a class of signals or images of interest. While diverse heuristic techniques have been proposed in the litterature to learn a dictionary from a collection of training samples, there are little existing results which provide an adequate mathematical understanding of the behaviour of these techniques and their ability to recover an ideal dictionary from which the training samples may have been generated.

Beyond our pioneering work [86], [110] [6] on this topic, which concentrated on the noiseless case for non-overcomplete dictionaries, this year we obtained new results showing the relevance of an ℓ^1 penalized cost function for the locally stable identification of overcomplete incoherent dictionaries, in the presence of noise and outliers. Moreover, we established new sample complexity bounds of dictionary learning and other related matrix factorization schemes (including PCA, NMF, structured sparsity ...) [59].

Analysis Operator Learning for Overcomplete Cosparse Representations : Besides standard dictionary learning, we also considered learning in the context of the cosparse model. We consider the problem of learning a low-dimensional signal model from a collection of training samples. The mainstream approach would be to learn an overcomplete dictionary to provide good approximations of the training samples using sparse synthesis coefficients. This famous sparse model has a less well known counterpart, in analysis form, called the cosparse analysis model. In this new model, signals are characterized by their parsimony in a transformed domain using an overcomplete analysis operator.

We considered several approaches to learn an analysis operator from a training corpus [102]. For one of them, which uses a constrained optimization program based on ℓ^1 optimization, we derived a practical learning algorithm, based on projected subgradients, and demonstrated its ability to robustly recover a ground truth analysis operator, provided the training set is of sufficient size. A local optimality condition was derived, providing preliminary theoretical support for the well-posedness of the learning problem under appropriate conditions [24]. Extensions to deal with noisy data have been obtained as well [119].

In more specific situations, when prior information is available on the operator, it is also possible to express the operator on a parametric form, and learn this parameter. For instance, in the sound source localization problem, we showed that unknown speed of sound can be learned jointly in the process of cosparse recovery, under mild conditions. This work was submitted to the iTwist'14 workshop.

6.2. Emerging activities on compressive sensing, learning and inverse problems

Compressive sensing, acoustic wavefields, audio inpainting,

6.2.1. Audio inpainting (SMALL FET-Open project)

Participants: Rémi Gribonval, Nancy Bertin, Corentin Guichaoua, Srdan Kitic.

Inpainting is a particular kind of inverse problems that has been extensively addressed in the recent years in the field of image processing. It consists in reconstructing a set of missing pixels in an image based on the observation of the remaining pixels. Sparse representations have proved to be particularly appropriate to address this problem. However, inpainting audio data has never been defined as such so far.

METISS has initiated a series of works about audio inpainting, from its definition to methods to address it. This research has begun in the framework of the EU Framework 7 FET-Open project FP7-ICT-225913-SMALL (Sparse Models, Algorithms and Learning for Large-Scale data) which began in January 2009. Rémi Gribonval was the coordinator of the project. The research on audio inpainting has been conducted by Valentin Emiya in 2010 and 2011.

The contributions consist of:

- defining audio inpainting as a general scheme where missing audio data must be estimated: it covers a number of existing audio processing tasks that have been addressed separately so far – click removal, declipping, packet loss concealment, unmasking in time-frequency;
- proposing algorithms based on sparse representations for audio inpainting (based on Matching Pursuit and on ℓ^1 minimization);
- addressing the case of audio declipping (*i.e.* desaturation): thanks to the flexibility of our inpainting algorithms, they can be constrained so as to include the structure of signals due to clipping in the objective to optimize. The resulting performance are significantly improved. This work appeared as a journal paper [63].
- addressing the case of audio declipping with the competitive cosparse approach, with promising result especially when the clipping level is low. A contribution was submitted to the iTwist'14 workshop [61].

Current and future works deal with developping advanced sparse decomposition for audio inpainting, including several forms of structured sparsity (*e.g.* temporal and multichannel joint-sparsity), dictionary learning for inpainting, and several applicative scenarios (declipping, time-frequency inpainting).

6.2.2. Blind Calibration of Compressive Sensing systems

Participants: Rémi Gribonval, Cagdas Bilen.

Main collaborations: Gilles Chardon, Laurent Daudet (Institut Langevin), Gilles Puy (EPFL)

We consider the problem of calibrating a compressed sensing measurement system under the assumption that the decalibration consists in unknown gains on each measure. We focus on blind calibration, using measures performed on a few unknown (but sparse) signals. A naive formulation of this blind calibration problem, using ℓ^1 minimization, is reminiscent of blind source separation and dictionary learning, which are known to be highly non-convex and riddled with local minima. In the considered context, when the gains are real valued and non-negative, we showed that in fact this formulation can be exactly expressed as a convex optimization problem, and can be solved using off-the-shelf algorithms. Numerical simulations demonstrated the effectiveness of the approach even for highly uncalibrated measures, when a sufficient number of (unknown, but sparse) calibrating signals is provided. We observed that the success/failure of the approach seems to obey sharp phase transitions [84]. This year, we focused on extending the framework to phase-only decalibration, using techniques revolving around low-rank matrix recovery [27], [26], [34], [52], and to joint phase and gain decalibration [54].

6.2.3. Compressive Gaussian Mixture estimation

Participants: Rémi Gribonval, Anthony Bourrier.

Main collaborations: Patrick Perez (Technicolor R&I France)

When fitting a probability model to voluminous data, memory and computational time can become prohibitive. In this paper, we propose a framework aimed at fitting a mixture of isotropic Gaussians to data vectors by computing a low-dimensional sketch of the data. The sketch represents empirical moments of the underlying probability distribution. Deriving a reconstruction algorithm by analogy with compressive sensing, we experimentally show that it is possible to precisely estimate the mixture parameters provided that the sketch is large enough. Our algorithm provides good reconstruction and scales to higher dimensions than previous probability mixture estimation algorithms, while consuming less memory in the case of numerous data. It also provides a privacy-preserving data analysis tool, since the sketch does not disclose information about individual datum it is based on [38], [40], [29].

6.3. Recent results on tensor decompositions

Multi-linear algebra is defined as the algebra of q-way arrays ($q > 2$), that is, the arrays whose elements are addressed by more than two indices. The first works back as far as Jordan who was interested in simultaneously diagonalizing two matrices at a time [92]. It is noteworthy that such two matrices can be interpreted as both slices of a three-way array and their joint diagonalization can be viewed as Hitchcock's polyadic decomposition [89] of the associated three-way array. Other works followed discussing rank problems related to multi-way structures and properties of multi-way arrays. However, these exercices in multilinear algebra were not linked to real data analysis but stayed within the realm of mathematics. Studying three-way data really started with Tucker's seminal work, which gave birth to the three-mode factor analysis [115]. His model is now often referred to as the Tucker3 model. At the same moment, other authors focused on a particular case of the Tucker3 model, calling it PARAFAC for PARAllel FACtor analysis [88], and on the means to achieve such a decomposition, which will become the famous canonical decomposition [77]. In honor to Hitchcock's pionneer work, we will call it the Canonical Polyadic (CP) decomposition.

Achieving a CP decomposition has been seen first as a mere non-linear least squares problem, with a simple objective criterion. In fact, the objective is a polynomial function of many variables, where some separate. One could think that this kind of objective is easy because smooth, and even infinitely differentiable. But it turns out that things are much more complicated than they may appear to be at first glance. Nevertheless, the Alternating Least Squares (ALS) algorithm has been mostly utilized to address this minimization problem, because of its programming simplicity. This should not hide the inherently complicated theory that lies behind the optimization problem. Moreover, in most of the applications, actual tensors may not exactly satisfy the expected model, so that the problem is eventually an approximation rather than an exact decomposition. This may results in a slow convergence (or lack of convergence) of iterative algorithms such as the ALS one [94]. Consequently, a new class of efficient algorithms able to take into account the properties of tensors to be decomposed is needed.

6.3.1. A novel direct algorithm for CP decompositions

Participant: Laurent Albera.

Main collaborations: Sepideh Hajipour (LTSI & BiSIPL), Isabelle Merlet (LTSI, France), Mohammad Bagher Shamsollahi (BiSIPL, Iran)

Nowadays several techniques are available to solve the CP problem. They can be classified in three main groups [113]: alternating algorithms, which update only a subset of the parameters at each step; derivative-based methods, seeking for an update of all the parameters simultaneously by successive approximations; and direct procedures. The latter algorithms compute the CP decomposition by solving an alternative algebra problem of lower dimensions, but they do not provide a solution in terms of least squares contrarily to the alternating and derivative-based techniques.

We proposed a new direct algorithm to compute the CP decomposition of complex-valued multi-way arrays. The proposed algorithm is based on the Simultaneous Schur Decomposition (SSD) of particular matrices derived from the array to process. We also proposed a new Jacobi-like algorithm to calculate the SSD of several complex-valued matrices. Besides, we analysed our SSD and SSD-based CP techniques in terms of i) identifiability, ii) computational complexity and iii) estimation accuracy through a large number of scenarios including synthetic and real data in the context of CP decomposition. Computer results showed the efficiency of the proposed SSD-based CP method of dealing with some well-known difficult scenarios with swamp-like degeneracies. We also showed that the proposed method outperformed the classical CP algorithms in processing of Paatero multi-way arrays. Finally, the robustness of the proposed algorithm with respect to overfactoring was highlighted. This work was briefly presented at ICASSP'13 [31] while a journal paper for submission to IEEE Transactions on Signal Processing is in preparation.

6.3.2. CP decomposition of semi-symmetric semi-nonnegative three-way arrays

Participant: Laurent Albera.

Main collaboration (line search methods): Julie Coloigner (LTSI, France), Amar Kachenoura (LTSI, France), Lotfi Senhadji (LTSI, France)

Main collaborations (Jacobi-like approaches): Lu Wang (LTSI, France), Amar Kachenoura (LTSI, France), Lotfi Senhadji (LTSI, France), Huazhong Shu (LIST, China)

We proposed new algorithms for the CP decomposition of semi-nonnegative semi-symmetric three-way tensors. In fact, it consists in fitting the CP model for which two of the three loading matrices are nonnegative and equal. Note that such a problem can also be interpreted as a nonnegative Joint Diagonalization by Congruence (JDC) problem.

Line search and trust region strategies

We first circumvented the nonnegativity constraint by means of changes of variable into squares, leading to a (polynomial) unconstrained optimization problem. Two optimization strategies, namely line search and trust region, were then studied. Regarding the former, a global plane search scheme was considered. It consists in computing, for a given direction, one or two optimal stepsizes, depending on whether the same stepsize is used in various updating rules. Moreover, we provided a compact matrix form for the derivatives of the objective function. This allows for a direct implementation of several iterative algorithms such as Conjugate Gradient (CG), Levenberg-Marquardt (LM) and Newton-like methods, in matrix programming environments like MATLAB. Note that the computational complexity issue was taken into account in the design phase of the algorithms, and was evaluated for each algorithm, allowing to fairly compare their performance.

Thus, various scenarios have been considered, aiming at testing the influence of i) an additive noise, which can stand for modeling errors, ii) the collinearity between factors, iii) the array rank and iv) the data size. The comparisons between our CG-like, Newton-like and LM-like methods (where semi-nonnegativity and semi-symmetry constraints are exploited), and classical CP algorithms (where no constraints are considered), showed that a better CP decomposition is obtained when these a priori are exploited, especially in the context of high dimensions and high collinearity. Finally, based on our numerical analysis, the algorithms that seem to yield the best tradeoff between accuracy and complexity are our $\mathrm{CG}_{\mathrm{2steps}}$-like and LM-like algorithms.

This work was accepted for publication with minor revisions to the Elsevier Linear Algebra and Applications journal.

Next, we considered an exponential change of variable leading to a different (non-polynomial) unconstrained optimization problem. Then we proposed novel algorithms based on line search strategy with an analytic global plane search procedure requiring new matrix derivations. Their performance was evaluated in terms of estimation accuracy and computational complexity. The classical ELS-ALS [109] and LM [113] algorithms without symmetry and nonnegativity constraints, and the ACDC algorithm [120] where only the semi-symmetry constraint is imposed, were tested as reference methods. Furthermore, the performance was also compared with our algorithms based on a square change of variable. The comparison studies showed that, among these approaches, the best accuracy/complexity trade off was achieved when an exponential change of variable was used through our ELS-ALS-like algorithm.

This work was submitted to the Elsevier Signal Processing journal.

Jacobi-like approaches

The line search (despite the use of global plane search procedures) and trust region strategies may be sensitive to initialization, and generally require a multi-initialization procedure. In order to circumvent this drawback, we considered in this work Jacobi-like approaches, which are known to be less sensitive to initialization. Note that our line search and trust region approaches can then be used to refine the solution obtained by the latter.

More particularly, we formulated the high-dimensional optimization problem into several sequential polynomial subproblems using i) a square change of variables to impose nonnegativity and ii) LU matrix factorization for parameterization. The two equal nonnegative loading matrices are actually written as the Hadamard product of two equal matrices which can be factorized as the product of elementary lower and upper triangular matrices, each one depending on only one parameter.

The first approach minimizes alternatively the classical least squares objective criterion with respect to each parameter of the two equal nonnegative loading matrices and each column of the third loading matrix. This work was published in the IEEE Signal Processing Letters journal [23]. The second technique reduces the previous optimization problem to the computation of the two equal nonnegative loading matrices. The third loading matrix is algebraically derived from the latter. This requires an appropriate parameterization of the set of matrices whose inverse is nonnegative. This work was briefly presented at EUSIPCO'13 [37] while a journal paper for submission to IEEE Transactions on Signal Processing is in preparation. Numerical experiments on simulated matrices emphasize the advantages of the proposed algorithms over classical CP and JDC techniques, especially in the case of degeneracies.

6.4. Source separation and localization

Source separation, sparse representations, tensor decompositions, semi-nonnegative independent component analysis, probabilistic model, source localization

6.4.1. A general framework for audio source separation

Participants: Frédéric Bimbot, Rémi Gribonval, Nancy Bertin.

Main collaboration: E. Vincent (EPI PAROLE, Inria Nancy); N.Q.K. Duong (Technicolor R&I France)

Source separation is the task of retrieving the source signals underlying a multichannel mixture signal. The state-of-the-art approach consists of representing the signals in the time-frequency domain and estimating the source coefficients by sparse decomposition in that basis. This approach relies on spatial cues, which are often not sufficient to discriminate the sources unambiguously. Recently, we proposed a general probabilistic framework for the joint exploitation of spatial and spectral cues [103], which generalizes a number of existing techniques including our former study on spectral GMMs [66]. This framework makes it possible to quickly design a new model adapted to the data at hand and estimate its parameters via the EM algorithm. As such, it is expected to become the basis for a number of works in the field, including our own.

Since the EM algorithm is sensitive to initialization, we devoted a major part of our work to reducing this sensitivity. One approach is to use some prior knowledge about the source spatial covariance matrices, either via probabilistic priors [82] or via deterministic subspace constraints [91]. The latter approach was the topic of the PhD thesis of Nobutaka Ito [90]. A complementary approach is to initialize the parameters in a suitable way using source localization techniques specifically designed for environments involving multiple sources and possibly background noise [74]. This year, we showed that the approach provides a statistically principled solution to the permutation problem in a semi-infomed scenario where the source positions and certain room characteristics are known [15].

6.4.2. Towards real-world separation and remixing applications

Participants: Nancy Bertin, Frédéric Bimbot, Jules Espiau de Lamaestre, Jérémy Paret, Laurent Simon, Nathan Souviraà-Labastie, Joachim Thiemann.

Shoko Araki, Jonathan Le Roux (NTT Communication Science Laboratories, JP), E. Vincent (EPI PAROLE, Inria Nancy)

Following our founding role in the organization of the Signal Separation Evaluation Campaigns (SiSEC) [65], [101], our invited paper summarized the outcomes of the three first editions of this campaign from 2007 to 2010 [116]. While some challenges remain, this paper highlighted that progress has been made and that audio source separation is closer than ever to successful industrial applications. This is also exemplified by the ongoing i3DMusic project and the contracts with Canon Research Centre France and MAIA Studio.

Our involvement in evaluation campaigns and source separation community was reinforced by the recording and the public release of the DEMAND (Diverse Environments Multi-channel Acoustic Noise Database) database, which provides multichannel real-world indoor and outdoor environment noise [44] under Creative Commons licence.

In order to exploit our know-how for these real-world applications, we investigated issues such as how to implement our algorithms in real time [111], how to adapt EM rules for faster computation in multichannel setting [35], how to reduce artifacts [96], how our techniques compare to beamforming in realistic conditions [36], and (in the context of our collaboration with MAIA studios) how best to exploit extra information or human input. In addition, while the state-of-the-art quality metrics previously developed by METISS remain widely used in the community, we proposed some improvements to the perceptually motivated metrics introduced last year [117].

6.4.3. Exploiting filter sparsity for source localization and/or separation

Participants: Alexis Benichoux, Rémi Gribonval, Frédéric Bimbot.

E. Vincent (EPI PAROLE, Inria Nancy)

Estimating the filters associated to room impulse responses between a source and a microphone is a recurrent problem with applications such as source separation, localization and remixing.

We considered the estimation of multiple room impulse responses from the simultaneous recording of several known sources. Existing techniques were restricted to the case where the number of sources is at most equal to the number of sensors. We relaxed this assumption in the case where the sources are known. To this aim, we proposed statistical models of the filters associated with convex log-likelihoods, and we proposed a convex optimization algorithm to solve the inverse problem with the resulting penalties. We provided a comparison between penalties via a set of experiments which shows that our method allows to speed up the recording process with a controlled quality tradeoff [72], [71]. This was a central part of the Ph.D. thesis of Alexis Benichoux [12] defended this year. A journal paper including extensive experiments with real data has been submitted [69].

We also investigated the filter estimation problem in a blind setting, where the source signals are unknown. On a more theoretical side, we studied the frequency permutation ambiguity traditionnally incurred by blind convolutive source separation methods. We focussed on the filter permutation problem in the absence of scaling, investigating the possible use of the temporal sparsity of the filters as a property enabling permutation correction. The obtained theoretical and experimental results highlight the potential as well as the limits of sparsity as an hypothesis to obtain a well-posed permutation problem. This work has been published in a conference [70] and as a journal paper [14].

Finally, we considered the problem of blind sparse deconvolution, which is common in both image and signal processing. To counter-balance the ill-posedness of the problem, many approaches are based on the minimization of a cost function. A well-known issue is a tendency to converge to an undesirable trivial solution. Besides domain specific explanations (such as the nature of the spectrum of the blurring filter in image processing) a widespread intuition behind this phenomenon is related to scaling issues and the nonconvexity of the optimized cost function. We proved that a fundamental issue lies in fact in the intrinsic properties of the cost function itself: for a large family of shift-invariant cost functions promoting the sparsity of either the filter or the source, the only global minima are trivial. We completed the analysis with an empirical method to verify the existence of more useful local minima [25].

6.4.4. Semi-nonnegative independent component analysis

Participant: Laurent Albera.

Main collaborations: Lu Wang (LTSI, France), Amar Kachenoura (LTSI, France), Lotfi Senhadji (LTSI, France), Huazhong Shu (LIST, China)

Independent Component Analysis (ICA) plays an important role in many areas including biomedical engineering [93], [64], [95], [118], [106], [81], speech and audio [67], [68], [78], [75], radiocommunications [80] and document restoration [114] to cite a few.

For instance in [114], the authors use ICA to restore digital document images in order to improve the text legibility. Indeed, under the statistical independence assumption, authors succeed in separating foreground text and bleed-through/show-through in palimpsest images. Furthermore, authors in [81] use ICA to solve the ambiguity in X-ray images due to multi-object overlappings. They presented a novel object decomposition technique based on multi-energy plane radiographs. This technique selectively enhances an object that is characterized by a specific chemical composition ratio of basis materials while suppressing the other overlapping objects. Besides, in the context of classification of tissues and more particularly of brain tumors [106], ICA is very effective. In fact, it allows for feature extraction from Magnetic Resonance Spectroscopy (MRS) signals, representing them as a linear combination of tissue spectra, which are as independent as possible [112]. Moreover, using the JADE algorithm [76] applied to a mixture of sound waves computed by means of the constant-Q transform (Fourier transform with log-frequency) of a temporal waveform broken up into a set of time segments, the authors of [75] describe trills as a set of note pairs described by their spectra and corresponding time envelopes. In this case, pitch and timing of each note present in the trill can be easily deduced.

All the aforementioned applications show the high efficiency of the ICA and its robustness to the presence of noise. Despite this high efficiency in resolving the proposed applicative problems, authors did not fully exploit properties enjoyed by the mixing matrix such as its nonnegativity. For instance in [81], the thickness of each organ, which stands for the mixing coefficient, is real positive. Furthermore, reflectance indices in [114] for the background, the overwriting and the underwriting, which correspond to the mixing coefficients, are also nonnegative. Regarding tissue classification from MRS data, each observation is a linear combination of independent spectra with positive weights representing concentrations [87]; the mixing matrix is again nonnegative.

By imposing the nonnegativity of the mixing matrix within the ICA process, we shown through computer results that the extraction quality can be improved. Exploiting the nonnegativity property of the mixing matrix during the ICA process gives rise to what we call semi-nonnegative ICA. More particularly, we performed the latter by computing a constrained joint CP decomposition of cumulant arrays of different orders [98] having the nonnegative mixing matrix as loading matrices. After merging the entries of the cumulant arrays in the same third order array, the reformulated problem follows the semi-symmetric semi-nonnegative CP model defined in section 6.3.2 . Hence we use the new methods described in section 6.3.2 to perform semi-nonnegative ICA. Performance results in audio and biomedical engineering were given in the different papers cited in section 6.3.2 .

6.4.5. Brain source localization

Participants: Laurent Albera, Srdan Kitic, Nancy Bertin, Rémi Gribonval.

Main collaborations: Hanna Becker (GIPSA & LTSI, France), Isabelle Merlet (LTSI, France), Fabrice Wendling (LTSI, France), Pierre Comon (GIPSA, France), Christian Benar (La Timone, Marseille), Martine Gavaret (La Timone, Marseille), Gwenaël Birot (FBML, Genève), Martin Haardt (TUI, Germany)

Main collaborations: Hanna Becker (GIPSA & LTSI, France), Pierre Comon (GIPSA, France), Isabelle Merlet (LTSI, France), Fabrice Wendling (LTSI, France)

Tensor-based approaches

The localization of several simultaneously active brain regions having low signal-to-noise ratios is a difficult task. To do this, tensor-based preprocessing can be applied, which consists in constructing a Space-Time-Frequency (STF) or Space-Time-Wave-Vector (STWV) tensor and decomposing it using the CP decomposition. We proposed a new algorithm for the accurate localization of extended sources based on the results of the tensor decomposition. Furthermore, we conducted a detailed study of the tensor-based preprocessing methods, including an analysis of their theoretical foundation, their computational complexity, and their performance for realistic simulated data in comparison to three conventional source localization algorithms, namely sLORETA [105], cortical LORETA (cLORETA) [104], and 4-ExSo-MUSIC [73]. Our objective consisted, on the one hand, in demonstrating the gain in performance that can be achieved by tensor-based preprocessing, and, on the other hand, in pointing out the limits and drawbacks of this method. Finally, we validated the STF and STWV techniques on real epileptic measurements to demonstrate their usefulness for practical applications. This work was recently submitted to the Elesevier NeuroImage journal.

From tensor to sparse models

The brain source imaging problem has been widely studied during the last decades, giving rise to an impressive number of methods using different priors. Nevertheless, a thorough study of the latter, including especially sparse and tensor-based approaches, is still missing. Consequently, we proposed i) a taxonomy of the methods based on a priori assumptions, ii) a detailed description of representative algorithms, iii) a review of identifiability results and convergence properties of different techniques, and iv) a performance comparison of the selected methods on identical data sets. Our aim was to provide a reference study in the biomedical engineering domain which may also be of interest for other areas such as wireless communications, audio source localization, and image processing where ill-posed linear inverse problems are encountered and to identify promising directions for future research in this area. A part of this work was submitted to ICASSP'14 while the whole part was submitted to IEEE Signal Processing Magazine.

A cosparsity-based approach

Cosparse modeling is particularly attractive when the signals of interest satisfy certain physical laws that naturally drive the choice of an analysis operator. We showed how to derive a reduced non-singular analysis operator describing EEG signals from Poisson's equation, Kirchhoff's law and some other physical constraints. As a result, we proposed the CoRE (Cosparse Representation of EEG signals) method to solve the classical brain source imaging problem. Computer simulations demonstrated the numerical performance of the CoRE method in comparison to a dictionary-based sparse approach. This work was submitted to ICASSP'14.

6.5. Audio and speech content processing

Audio segmentation, speech recognition, motif discovery, audio mining

6.5.1. Audio motif discovery

Participants: Frédéric Bimbot, Laurence Catanese.

This work was performed in close collaboration with Guillaume Gravier from the Texmex project-team.

As an alternative to supervised approaches for multimedia content analysis, where predefined concepts are searched for in the data, we investigate content discovery approaches where knowledge emerge from the data. Following this general philosophy, we pursued work on motif discovery in audio contents.

Audio motif discovery is the task of finding out, without any prior knowledge, all pieces of signals that repeat, eventually allowing variability. The developed algorithms allows discovering and collecting occurrences of repeating patterns in the absence of prior acoustic and linguistic knowledge, or training material.

Former work extended the principles of seeded discovery to near duplicate detection and spoken document retrieval from examples [99].

In 2012, the work achieved consisted in consolidating previously obtained results with the motif discovery algorithm and making implementation choices regardless of the structure and the code, in order to minimize the computation time. This has lead to the creation of a software prototype called MODIS.

After the code has been thoroughly optimised, further optimizations to improve the system performances was to change the method used for the search of similarities between patterns. A new functionality has been added to get rid of unrelevant patterns like silence in speech. New versions of dynamic time warping have been implemented, as well as the possibility to downsample the input sequence during the process, which allows a huge gain of computation time.

The principles of the MODIS software has been documented in details [48] and demonstrated during a Show & Tell session at the Interspeech 2013 conference [41].

This work has been carried out in the context of the Quaero Project.

6.5.2. Landmark-driven speech recognition

Participant: Stefan Ziegler.

This work is supervised by Guillaume Gravier and Bogdan Ludusan from the Texmex project-team.

Our previous studies indicate that acoustic-phonetic approaches to ASR, while they cannot achieve state-of-the-art ASR performance by themselves, can prevent HMM-based ASR from degrading, by integrating additional knowledge into the decoding.

In our previous framework we inserted knowledge into the decoding by detecting time frames (referred to as landmarks) which estimate the presence of the active broad phonetic class. This enables the use of a modified version of the viterbi decoding that favours states that are coherent with the detected phonetic knowledge [122].

In 2012 we focused on two major issues. First, we aimed at finding new ways to model and detect phonetic landmarks. Our second focus was on the extension of our landmark detector towards a full acoustic-phonetic framework, to model speech by a variety of articulatory features.

Our new approach for the classification and detection of speech units focuses on developping landmark-models that are different from existing frame-based approaches to landmark detection [121]. In our approach, we use segmentation to model any time-variable speech unit by a fixed-dimensional observation vector. After training any desired classifier, we can estimate the presence of a desired speech unit by searching for each time frame the corresponding segment, that provides the maximum classification score.

We used this segment-based landmark-detection inside a standalone acoustic-phonetic framework that models speech as a stream of articulatory features. In this framework we first search for relevant broad phonetic landmarks, before attaching each landmark with the full set of articulatory features.

Integrating these articulatory feature streams into a standard HMM-based speech recognizer by weighted linear combination improves speech recognition up to 1.5

Additionally, we explored the possibilities of using stressed syllables as an information to guide the viterbi decoding. This work was carried under the leaderhip of Bogdan Ludusan from the team TEXMEX at IRISA [97].

6.5.3. Mobile device for the assistance of users in potentially dangerous situations

Participants: Romain Lebarbenchon, Frédéric Bimbot.

The S-Pod project is a cooperative project between industry and academia aiming at the development of mobile systems for the detection of potentially dangerous situations in the immediate environment of a user, without requiring his/her active intervention.

In this context, the PANAMA research group is involved in the design of algorithms for the analysis and monitoring of the acoustic scene around the user, yielding information which can be fused with other sources of information (physiological, contextual, etc...) in order to trigger an alarm when needed and subsequent appropriate measures.

Currently in its initial phase, work has mainly focused on functional specifications and performance requirements.

6.6. Music Content Processing and Music Information Retrieval

Acoustic modeling, non-negative matrix factorisation, music language modeling, music structure

6.6.1. Music language modeling

Participants: Frédéric Bimbot, Dimitri Moreau, Stanislaw Raczynski.

Main collaboration: S. Fukayama (University of Tokyo, JP), E. Vincent (EPI PAROLE, Inria Nancy), Intern: A. Aras

Music involves several levels of information, from the acoustic signal up to cognitive quantities such as composer style or key, through mid-level quantities such as a musical score or a sequence of chords. The dependencies between mid-level and lower- or higher-level information can be represented through acoustic models and language models, respectively.

We pursued our pioneering work on music language modeling, with a particular focus on the joint modeling of "horizontal" (sequential) and "vertical" (simultaneous) dependencies between notes by log-linear interpolation of the corresponding conditional distributions. We identified the normalization of the resulting distribution as a crucial problem for the performance of the model and proposed an exact solution to this problem [108]. We also applied the log-linear interpolation paradigm to the joint modeling of melody, key and chords, which evolve according to different timelines [107]. In order to synchronize these feature sequences, we explored the use of beat-long templates consisting of several notes as opposed to short time frames containing a fragment of a single note.

The limited availability of multi-feature symbolic music data is currently an issue which prevents the training of the developed models on sufficient amounts of data for the unsupervised probabilistic approach to significantly outperform more conventional approaches based on musicological expertise. We outlined a procedure for the semi-automated collection of large-scale multifeature music corpora by exploiting the wealth of music data available on the web (audio, MIDI, leadsheets, lyrics, etc) together with algorithms for the automatic detection and alignment of matching data. Following this work, we started collecting pointers to data and developing such algorithms.

Effort was dedicated to the investigation of structural models for improving the modeling of chord sequence. Preliminary results obtained during Anwaya Aras' intersnship show that using a matricial structure of time dependencies between successive chords improves the predictability of chord sequences as compared to a purely sequential model.

6.6.2. Music structuring

Participants: Frédéric Bimbot, Anaik Olivero, Gabriel Sargent.

Main collaboration: E. Vincent (EPI PAROLE, Inria Nancy), Intern: E. Deruty

The structure of a music piece is a concept which is often referred to in various areas of music sciences and technologies, but for which there is no commonly agreed definition. This raises a methodological issue in MIR, when designing and evaluating automatic structure inference algorithms. It also strongly limits the possibility to produce consistent large-scale annotation datasets in a cooperative manner.

Last year, our methodology for the *semiotic* annotation of music pieces has developed and concretized into a set of principles, concepts and conventions for locating the boundaries and determining metaphoric labels of music segments. The method relies on a new concept for characterizing the inner organization of music segments called the System & Contrast (S&C) model [2]. The annotation of 383 music pieces has been finalized, documented [28] and released to the MIR scientific community: http://musicdata.gforge.inria.fr/structureAnnotation.html.

For what concerns algorithmic approaches to music structure description [13], we have formulated the segmentation process as the optimization of a cost function which is composed of two terms: the first one corresponds to the characterization of structural segments by means of audio criteria; the second one relies on the regularity of the target structure with respect to a "structural pulsation period". In this context, we have compared several regularity constraints and studied the combination of audio criteria through fusion. We also considered the estimation of structural labels as a probabilistic finite-state automaton selection process : in this scope, we have proposed an auto-adaptive criterion for model selection, applied to a description of the tonal content. We also proposed a labeling method derived from the system-contrast model. We have evaluated and compared several systems for structural segmentation of music based on these approaches in the context of national and international evaluation campaigns (Quaero, MIREX).

As a follow-up to this work on music structure description, we are currently designing new models and algorithms for segmenting and labeling music into structural units. In one approach (Corentin Guichaoua's PhD), music structure is described as a hierarchical tree estimated by a grammar inference process whereas a second approach (Anaik Olivero's Post-doc) addresses music structure description as the estimation of a graph of similarity relationships.

PAREO Project-Team

6. New Results

6.1. Static analysis

Participant: Sergueï Lenglet.

6.1.1. Static analysis for control operators

Control operators allow programs to have access and manipulate their execution context. Abortive control operators, such as *call/cc* in Scheme or SML, capture the entire execution context (also called continuation), while delimited-control operators, such as *shift* and *reset* captures only a part of the continuation (delimited by reset). We want to prove properties (like equivalences between terms or termination) for languages with these operators, using static analysis.

In [9], [16], we study the behavioral theory of a language with delimited control. More precisely, we define environmental bisimilarities for the delimited-control operators shift and reset. We consider two different notions of contextual equivalence: one that does not require the presence of a top-level control delimiter when executing tested terms, and another one, fully compatible with the original CPS semantics of shift and reset, that does. For each of them, we develop sound and complete environmental bisimilarities, and we discuss up-to techniques.

In [8], we present new proofs of termination of evaluation in reduction semantics (i.e., a small-step operational semantics with explicit representation of evaluation contexts) for System F with control operators. We introduce a modified version of Girard's proof method based on reducibility candidates, where the reducibility predicates are defined on values and on evaluation contexts as prescribed by the reduction semantics format. We address both abortive control operators (*callcc*) and delimited-control operators (*shift* and *reset*) for which we introduce novel polymorphic type systems, and we consider both the call-by-value and call-by-name evaluation strategies.

6.1.2. Polymorphism and higher-order functions for XML

In [11], we define a calculus with higher-order polymorphic functions, recursive types with arrow and product type constructors and set-theoretic type connectives (union, intersection, and negation). We study the explicitly-typed version of the calculus in which type instantiation is driven by explicit instantiation annotations. In particular, we define an explicitly-typed λ-calculus with intersection types and an efficient evaluation model for it. In a companion paper [21], we define a local type inference system that allows the programmer to omit explicit instantiation annotations, and a type reconstruction system that allows the programmer to omit explicit type annotations. The work presented in the two articles provides the theoretical foundations and technical machinery needed to design and implement higher-order polymorphic functional languages for semi-structured data.

6.2. Model Transformations

Participants: Jean-Christophe Bach, Pierre-Etienne Moreau.

Model Driven Engineering is a technique that has been applied quite successfully for the design of complex systems. Such systems cannot be released and embedded without complying with the certification required by the application domain: EN 50128 for railways, DO-178C for aeronautics, or ISO 26262 for automotive for instance.

Recently we have developed an extension of *Tom* to support the development of Model Transformations and the generation of traces which are needed to give confidence in the quality of the implemented transformation.

In [12], we present a method, a language and dedicated tooling to ease and to speed up software development based on models transformations. Our approach aims to bridge the gap between general purpose languages and domain specific ones in order to take benefit from both of the two worlds, and to increase software quality. Our approach uses the Tom language which is a shallow extension of general purpose languages. Our proposal allows to write modular transformations whose code is reusable, and which are traceable.

6.3. Property based testing

Participants: Horatiu Cirstea, Pierre-Etienne Moreau, Cosay Topaktas.

Quality is crucial for software systems and several aspects should be taken into account. Formal verification techniques like model checking and automated theorem proving can be used to guarantee the correctness of finite or infinite systems. While these approaches provide a high level of confidence they are sometimes difficult and expensive to apply. Software testing is another approach and although it cannot guarantee correctness it can be very efficient in finding errors.

We have proposed a property based testing framework for the *Tom* language inspired from the ones prosed in the context of functional programming. In the current version relatively simple properties can be already expressed and tested on *Tom* programs. It consists of an exhaustive approach testing all possible input values and guaranteeing that the discovered counter-examples are the smallest ones (the size of the inputs is clearly limited by the execution time) and a random approach where inputs of bigger size could be tested but the minimal counter-example is not guaranteed. A relatively simple shrinking method which searches a smaller counter-example starting from an initial relatively complex one has been also proposed. There is ongoing work on the expressiveness of the property language and the efficiency of the shrinking method. The library is available at http://gforge.inria.fr/projects/tom.

6.4. Nominal Theory

Participant: Christophe Calvès.

Nominal unification is proven to be quadratic in time and space. It was so by two different approaches, both inspired by the Paterson-Wegman linear unification algorithm, but dramatically different in the way nominal and first-order constraints are dealt with.

To handle nominal constraints, Levy and Villaret introduced the notion of replacing while Calvès and Fernández use permutations and sets of atoms. To deal with structural constraints, the former use multi-equation in a way similar to the Martelli-Montanari algorithm while the later mimic Paterson-Wegman.

In [10] we abstract over these two approaches and genralize them into the notion of modality, highlighting the general ideas behind nominal unification. We show that replacings and environments are in fact isomorphic. This isomorphism is of prime importance to prove intricate properties on both sides and a step further to the real complexity of nominal unification.

PARIETAL Project-Team

6. New Results

6.1. Deformable Template estimation for joint anatomical and functional brain images

Participants: Bertrand Thirion [Correspondant], Hao Xu, Stéphanie Allassonnière.

Traditional analyses of Functional Magnetic Resonance Imaging (fMRI) use little anatomical information. The registration of the images to a template is based on the individual anatomy and ignores functional information; subsequently detected activations are not confined to gray matter (GM). In this work, we propose a statistical model to estimate a probabilistic atlas from functional and T1 MRIs that summarizes both anatomical and functional information and the geometric variability of the population. Registration and Segmentation are performed jointly along the atlas estimation and the functional activity is constrained to the GM, increasing the accuracy of the atlas.

More details can be found in [69].

6.2. Randomized parcellation-based inference

Participants: Gaël Varoquaux, Bertrand Thirion, Benoit Da Mota, Virgile Fritsch.

Neuroimaging group analyses are used to relate inter-subject signal differences observed in brain imaging with behavioral or genetic variables and to assess risks factors of brain diseases. The lack of stability and of sensitivity of current voxel-based analysis schemes may however lead to non-reproducible results. We introduce a new approach to overcome the limitations of standard methods, in which active voxels are detected according to a consensus on several random parcellations of the brain images, while a permutation test controls the false positive risk (see Fig. 3). Both on synthetic and real data, this approach shows higher sensitivity, better accuracy and higher reproducibility than state-of-the-art methods. In a neuroimaging-genetic application, we find that it succeeds in detecting a significant association between a genetic variant next to the COMT gene and the BOLD signal in the left thalamus for a functional Magnetic Resonance Imaging contrast associated with incorrect responses of the subjects from a Stop Signal Task protocol.

More details can be found in [55].

6.3. Group-level impacts of within- and between-subject hemodynamic variability in fMRI

Participants: Gaël Varoquaux, Solveig Badillo, Philippe Ciuciu [Correspondant].

Inter-subject fMRI analyses have specific issues regarding the reliability of the results concerning both the detection of brain activation patterns and the estimation of the underlying dynamics. Among these issues lies the variability of the hemodynamic response function (HRF), that is usually accounted for using functional basis sets in the general linear model context. Here, we use the joint detection-estimation approach (JDE) [76], [78], which combines regional nonparametric HRF inference with spatially adaptive regularization of activation clusters to avoid global smoothing of fMRI images (see Fig. 4). We show that the JDE-based inference brings a significant improvement in statistical sensitivity for detecting evoked activity in parietal regions. In contrast, the canonical HRF associated with spatially adaptive regularization is more sensitive in other regions, such as motor cortex. This different regional behavior is shown to reflect a larger discrepancy of HRF with the canonical model. By varying parallel imaging acceleration factor, SNR-specific region-based hemodynamic parameters (activation delay and duration) were extracted from the JDE inference. Complementary analyses highlighted their significant departure from the canonical parameters and the strongest between-subject variability that occurs in the parietal region, irrespective of the SNR value. Finally, statistical evidence that the fluctuation of the HRF shape is responsible for the significant change in activation detection performance is demonstrated using paired t-tests between hemodynamic parameters inferred by GLM and JDE.

../../../../projets/parietal/IMG/method.png

Figure 3. Overview of the randomized parcellation based inference framework on an example with few parcels. The variability of the parcels definition is used to obtain voxel-level statistics.

More details can be found in [49].

../../../../projets/parietal/IMG/solveig.png

Figure 4. General sketch summarizing the HRF computation at the subject and group-levels in activated regions r. Left: Position of the activation peak in r (here left motor cortex) given in mm in the Talairach space. Center: Individual weighted HRF time course extraction. Right: Computation of the group average normalized HRF time course with corresponding error bars ($\pm\sigma$).

6.4. Mapping cognitive ontologies to and from the brain

Participants: Gaël Varoquaux [Correspondant], Bertrand Thirion, Yannick Schwartz.

Imaging neuroscience links brain activation maps to behavior and cognition via correlational studies. Due to the nature of the individual experiments, based on eliciting neural response from a small number of stimuli, this link is incomplete, and unidirectional from the causal point of view. To come to conclusions on the function implied by the activation of brain regions, it is necessary to combine a wide exploration of the various brain functions and some inversion of the statistical inference. Here we introduce a methodology for accumulating knowledge towards a bidirectional link between observed brain activity and the corresponding function. We rely on a large corpus of imaging studies and a predictive engine. Technically, the challenges are to find commonality between the studies without denaturing the richness of the corpus. The key elements that we contribute are labeling the tasks performed with a cognitive ontology, and modeling the long tail of rare paradigms in the corpus. To our knowledge, our approach is the first demonstration of predicting the cognitive content of completely new brain images. To that end, we propose a method that predicts the experimental paradigms across different studies (see Fig. 5).

More details can be found in [63].

6.5. Implications of Inconsistencies between fMRI and dMRI on Multimodal Connectivity Estimation

Participants: Gaël Varoquaux [Correspondant], Bertrand Thirion, Bernard Ng.

There is a recent trend towards integrating resting state functional magnetic resonance imaging (RS-fMRI) and diffusion MRI (dMRI) for brain connectivity estimation, as motivated by how estimates from these modalities are presumably two views reflecting the same underlying brain circuitry. In this work, we show on a cohort of 60 subjects that conventional functional connectivity (FC) estimates based on Pearson's correlation and anatomical connectivity (AC) estimates based on fiber counts are actually not that highly correlated for typical RS-fMRI (7 min) and dMRI (32 gradient directions) data. The FC-AC correlation can be significantly increased by considering sparse partial correlation and modeling fiber endpoint uncertainty, but the resulting FC-AC correlation is still rather low in absolute terms. We further exemplify the inconsistencies between FC and AC estimates by integrating them as priors into activation detection and demonstrating significant differences in their detection sensitivity. Importantly, we illustrate that these inconsistencies can be useful in fMRI-dMRI integration for improving brain connectivity estimation.

More details can be found in [61]. See also [60].

6.6. Extracting brain regions from rest fMRI with Total-Variation constrained dictionary learning

Participants: Gaël Varoquaux [Correspondant], Alexandre Abraham.

Spontaneous brain activity reveals mechanisms of brain function and dysfunction. Its population-level statistical analysis based on functional images often relies on the de nition of brain regions that must summarize e ciently the covariance structure between the multiple brain networks. In this paper, we extend a network-discovery approach, namely dictionary learning, to readily extract brain regions. To do so, we intro duce a new tool drawing from clustering and linear decomposition methods by carefully crafting a penalty. Our approach automatically extracts regions from rest fMRI that better explain the data and are more stable across subjects than reference decomposition or clustering methods (see FIg. 6).

More details can be found in [47].

6.7. Cohort-level brain mapping: learning cognitive atoms to single out specialized regions

Participants: Gaël Varoquaux [Correspondant], Bertrand Thirion, Yannick Schwartz.

../../../../projets/parietal/IMG/forward_regions_stimulus_modality.png

../../../../projets/parietal/IMG/reverse_regions

../../../../projets/parietal/IMG/forward_regions_presentation_modality.png

../../../../projets/parietal/IMG/reverse_regions

../../../../projets/parietal/IMG/forward_regions_instructions.png

../../../../projets/parietal/IMG/reverse_regions

../../../../projets/parietal/IMG/kmeans_hard.png

../../../../projets/parietal/IMG/ward_hard.png

../../../../projets/parietal/IMG/msdl_hard.png

../../../../projets/parietal/IMG/msdl_unreg_hard

../../../../projets/parietal/IMG/smooth_msdl_hard.png

../../../../projets/parietal/IMG/icams_hard.png

Functional Magnetic Resonance Imaging (fMRI) studies map the human brain by testing the response of groups of individuals to carefully-crafted and contrasted tasks in order to delineate specialized brain regions and networks. The number of functional networks extracted is limited by the number of subject-level contrasts and does not grow with the cohort. Here, we introduce a new group-level brain mapping strategy to differentiate many regions reflecting the variety of brain network configurations observed in the population. Based on the principle of functional segregation, our approach singles out functionally-specialized brain regions by learning group-level functional profiles on which the response of brain regions can be represented sparsely. We use a dictionary-learning formulation that can be solved efficiently with on-line algorithms, scaling to arbitrary large datasets. Importantly, we model inter-subject correspondence as structure imposed in the estimated functional profiles, integrating a structure-inducing regularization with no additional computational cost. On a large multi-subject study, our approach extracts a large number of brain networks with meaningful functional profiles (see Fig. 7).

More details can be found in [66].

6.8. Identifying predictive regions from fMRI with TV-$\ell 1$ prior

Participants: Gaël Varoquaux [Correspondant], Bertrand Thirion, Alexandre Gramfort.

Decoding, i.e. predicting stimulus related quantities from functional brain images, is a powerful tool to demonstrate differences between brain activity across conditions. However, unlike standard brain mapping, it offers no guaranties on the localization of this information. Here, we consider decoding as a statistical estimation problem and show that injecting a spatial segmentation prior leads to unmatched performance in recovering predictive regions. Specifically, we use $\ell 1$ penalization to set voxels to zero and Total-Variation (TV) penalization to segment regions. Our contribution is two-fold. On the one hand, we show via extensive experiments that, amongst a large selection of decoding and brain-mapping strategies, TV+$\ell 1$ leads to best region recovery (see Fig. 8). On the other hand, we consider implementation issues related to this estimator. To tackle efficiently this joint prediction-segmentation problem we introduce a fast optimization algorithm based on a primal-dual approach. We also tackle automatic setting of hyper-parameters and fast computation of image operation on the irregular masks that arise in brain imaging.

More details can be found in [59].

6.9. Second order scattering descriptors predict fMRI activity due to visual textures

Participants: Michael Eickenberg, Bertrand Thirion [Correspondant], Alexandre Gramfort.

Second layer scattering descriptors are known to provide good classification performance on natural quasi-stationary processes such as visual textures due to their sensitivity to higher order moments and continuity with respect to small deformations. In a functional Magnetic Resonance Imaging (fMRI) experiment we present visual textures to subjects and evaluate the predictive power of these descriptors with respect to the predictive power of simple contour energy - the first scattering layer. We are able to conclude not only that invariant second layer scattering coefficients better encode voxel activity, but also that well predicted voxels need not necessarily lie in known retinotopic regions (see Fig. 9).

More details can be found in [56].

6.10. Bayesian Joint Detection-Estimation of cerebral vasoreactivity from ASL fMRI data

Participants: Thomas Vincent, Philippe Ciuciu [Correspondant].

../../../../projets/parietal/IMG/atlas.png

Figure 7. (Left) A brain functional atlas can be conceptualized as a parcellation of the brain volume into overlapping networks, where each functional network is characterized by a profile of activation for a set of functional contrasts. (Right) Such an atlas can be learned by applying an adapted dictionary learning to a set of images that display the activation observed in different subjects for a (very large) set of cognitive tasks.

../../../../projets/parietal/IMG/tvl1.png

Figure 8. Results on fMRI data from (from left to right F-test, ElasticNet and TV-ℓ_1). The TV-ℓ_1 regularized model segments neuroscientificly meaningful predictive regions in agreement with univariate statistics while the ElasticNet yields sparse although very scattered non-zero weights.

../../../../projets/parietal/IMG/scattering.png

Figure 9. Some brain regions are better explained by using two scattering layers rather than one (middle). These regions are symetric across hemispheres, and are observed mostly in the dorsal stream of the visual cortex. An atlas of the visual areas (left and right) shows that the mai foci are found in the V1, V2, V3AB and IPS0 regions.

Although the study of cerebral vasoreactivity using fMRI is mainly conducted through the BOLD fMRI modality, owing to its relatively high signal-to-noise ratio (SNR), ASL fMRI provides a more interpretable measure of cerebral vasoreactivity than BOLD fMRI. Still, ASL suffers from a low SNR and is hampered by a large amount of physiological noise. The current contribution aims at improving the recovery of the vasoreactive component from the ASL signal. To this end, a Bayesian hierarchical model is proposed, enabling the recovery of perfusion levels as well as fitting their dynamics. On a single-subject ASL real data set involving perfusion changes induced by hypercapnia, the approach is compared with a classical GLM-based analysis. A better goodness-of-fit is achieved, especially in the transitions between baseline and hypercapnia periods. Also, perfusion levels are recovered with higher sensitivity and show a better contrast between gray- and white matter.

More details can be found in [68].

PARKAS Project-Team

6. New Results

6.1. Reactive Programming

Participants: Guillaume Baudart, 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 published [20]. An extended version is under submission.

We continued the work on a new reactivity analysis which ensures that a process can not prevent the other ones to from executing. This analysis has published in [19]. An English version is under submission.

The runtime of ReactiveML has been cleanup and a multi-threaded implementation has been developed. A paper describing this new implementation will be published in [27].

All these novelties has been described precisely in the PhD thesis of Cédric Pasteur [1].

During the year, ReactiveML has also bee applied to *mixed music*. Mixed music is about live musicians interacting with electronic parts which are controlled by a computer during the performance. It allows composers to use and combine traditional instruments with complex synthesized sounds and other electronic devices. There are several languages dedicated to the writing of mixed music scores. Among them, the Antescofo language coupled with an advanced score follower allows a composer to manage the reactive aspects of musical performances: how electronic parts interact with a musician. However these domain specific languages do not offer the expressiveness of functional programming.

We defined a synchronous semantics for the core language of Antescofo and an alternative implementation based on an embedding inside ReactiveML [9]. The semantics reduces to a few rules, is mathematically precise and leads to an interpretor of only a few hundred lines. The efficiency of this interpretor compares well with that of the actual implementation: on all musical pieces we have tested, response times have been less than the reaction time of the human ear. Moreover, this approach offers to the composer recursion, higher order, inductive types, as well as a simple way to program complex reactive behaviors thanks to the synchronous model of concurrency on which ReactiveML is built [10].

6.2. n-Synchronous Languages

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

Synchronous programming languages in the vein of Lustre were designed for critical real-time systems. They are, however, not that well adapted to embedded applications with more pressing computational needs, since the generated code will usually not contain loops or arrays.

An essential task of a Lustre compiler is to determine whether a program can be executed within bounded memory. This process is called the "clock calculus", and consists in mapping every item of each program stream to a logical date in a global, discrete time scale. For a given stream, the mapping itself is called a "clock", and is a strictly increasing function from stream positions to natural numbers representing ticks: two items cannot be computed at the same time. In practice, this function is represented as an infinite binary stream where the boolean b_i denotes presence (or absence) in the corresponding data stream at the i-th instant.

In recent work, Guatto, Cohen, Mandel and Pouzet considered the extension of the Lustre and Lucid Synchrone clock calculus to allow computing several values instantaneously. This simple idea has a deep impact on all aspects of the language: - its denotational semantics has to account for bursts of values; - the clock calculus now features integers rather than booleans: each integer denotes the size of the burst at the corresponding instant; - causality analysis has to take bursts into account when rejecting self-referential programs; - the code generation process translates bursts to arrays and clocks to counted loops.

A prototype implementation exploiting this idea and generating C code with loops is underway and a paper describing the base of the clock calculus will be published [26].

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

6.3. Mechanization of AODV loop freedom proof

Participant: Timothy Bourke.

The Ad hoc On demand Distance Vector (AODV) routing protocol is described in RFC3561. It allows the nodes in a Mobile Ad hoc Network (MANET) to know where to forward messages so that they eventually reach their destinations. The nodes of such networks are *reactive systems* that cooperate to provide a global service (the sending of messages from node to node) satisfying certain correctness properties (namely 'loop freedom'—that messages are never sent in circles).

We have mechanized an existing formal but pen-and-paper proof of loop freedom of AODV in the interactive theorem prover Isabelle/HOL. While the process algebra model and the fine details of the original proof are quite formal, the structure of the proof is much less so. This necessitated the development of new framework elements and techniques in Isabelle. In particular, we adapted standard theory on inductive assertions to show invariants over individual reactive nodes and introduced machinery for assume/guarantee reasoning to lift these invariants to networks of communicating processes. While the original proof reasoned informally over traces, the mechanized proof is purely based on invariant reasoning, i.e., on reasoning over pairs of reachable states. Our combination of techniques works very well and is likely useful for modelling and verifying similar protocols in an interactive theorem prover.

We are currently finalising a paper describing this work for submission in January.

In collaboration with Peter Hofner (NICTA) and Robert J. van Glabbeek (UNSW/NICTA).

6.4. Hybrid Synchronous Languages

Participants: Timothy Bourke, Jun Inoue, Antoine Madet, Marc Pouzet.

During year 2013, we mainly worked on three directions: (a) the treatment of DAEs; (b) the design and implementation of a causality analysis for hybrid systems modelers; (c) the study of numerical techniques for *non-smooth dynamical systems*.

DAEs As part of our participation in the European project MODRIO and SYS2SOFT projects, we have been developing a prototype for simulating DAE (Differential-Algebraic Equations) systems. DAEs are the basis of the language Modelica and their interaction with discrete features — in particular the novel ones introduced in 2012, like hierarchical automata and clocks — raise difficult semantical and compilation issues. The goal is to precisely define the interaction between synchronous programming constructs and DAEs, in term of semantics and compilation. One strong difficulty at the moment is that existing techniques (index reduction, dymmy derivative) are not modular and force, either to (a) write an interpretor where index reduction is done dynamically every time a mode change occurs or (b) statically enumerate all the modes, performing index reduction for every of those. While the first technique is too slow in practice (and it is not used in the most advanced Modelica compiler), the second one may explode in practice (putting n two-state automata in parallel lead to 2^n states to be enumerated). During year 2013, we have investigated a new approach for index reduction.

Work to-date has focused on implementing standard algorithms from the literature (notably Pantelides, Dummy Derivatives, Dynamic State Selection). Despite the importance of these algorithms to tools like Modelica, we found that important implementation details and "tricks" are not always well documented.

This work is developed hand-in-hand with the interface to the Sundials IDA solver.

Causality Analysis We have designed a causality analysis for a language that mix stream equations, hierarchical automata and ODEs and implemented it in the Zélus compiler. Its purpose is to give a sufficient condition for a hybrid program can be turned into statically scheduled code. Moreover, the analysis ensures that absence of discontinuities outside of declared zero-crossing events. This result is novel and the proof deeply rely on the use of *non standard analysis* introduced in our previous works. This new result has been accepted for publication at HSCC 2014.

Non Smooth Dynamical Systems In parallel, we collaborate with Bernard Brogliato and Vincent Acary (Inria team BIBOP, Grenoble) on non smooth dynamical systems. Beside general-purpose techniques for solving DAEs and implemented in Modelica compilers, there exist dedicated methods for systems with a lot of discontinuities and contacts (in mechanical system, electrical analogous circuits, etc.). They are far more efficient and numerically accurate than general-purpose techniques when the number of contact is important (e.g., transient in electrical circuits, a bag of marbles). They are based on a time stepping execution and do not have to stop at every zero-crossing event. The combination of those techniques with event detection ones (as used in the Simulink tool) is largely unknown. We are currently inverstigating the extension of our previous work to take Brogliato and Acary techniques into account. This is a novel but promising direction of research for the year to come.

In this research activity, we develop the new language Zélus used as a laboratory for experimenting novel programming constructs and compilation techniques. It serves to illustrate our research as Lucid Synchrone did in the past.

In collaboration with Benoit Caillaud and Albert Benveniste of the Inria HYCOMES team.

6.5. Fidelity in Real-Time Programming

Participants: Timothy Bourke, Guillaume Baudart.

We are close to completing a careful analysis of literature related to the quasi-synchronous model for real-time, distributed systems. We have extended existing results by increasing their precision, providing detailed proofs, and simplifying protocol descriptions. The work to-date is documented in a draft document which we expect will eventually become a technical report or journal article.

Quasi-synchronous architectures, sometimes termed Loosely Time-Triggered Architectures (LTTAs), are ubiquitious in the development of distributed, real-time systems. They represent a broad class of systems whose modelling and programming mixes elements of discrete time, physical time, and a notion of approximation. We expect that addressing these elements—in the Zélus programming language—will lead to insights and advances in a broader ambition to program in physical time.

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

Participants: Francesco Zappa Nardelli, Robin Morisset.

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 accepted in [22].

6.7. A verified compiler for relaxed-memory concurrency

Participant: Francesco Zappa Nardelli.

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. In 2013 an article, describing the correctness proof of all the phases of our CompCertTSO compiler (including experimental fence eliminations), appeared in the Journal of the ACM [7].

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

6.8. Language design on top of JavaScript

Participant: Francesco Zappa Nardelli.

This research project aims at improving the design of the JavaScript language. In [23] we present a security infrastructure which allows users and content providers to specify access control policies over subsets of a JavaScript program by leveraging the con- cept of delimited histories with revocation. We implement our proposal in WebKit and evaluate it with three policies on 50 widely used websites with no changes to their JavaScript code and report performance overheads and violations. In [32] we propose a typed extension of JavaScript combining dynamic types, concrete types and like types to let developers pick the level of guarantee that is appropriate for their code. We have implemented our type system and we report on performance and software engineering benefits.

With Gregor Richards and Jan Vitek (Purdue University).

6.9. Tiling for iterated stencils

Participants: Tobias Grosser, Sven Verdoolaege, Albert Cohen.

Time-tiling is necessary for the efficient execution of iterative stencil computations. Classical hyper-rectangular tiles cannot be used due to the combination of backward and forward dependences along space dimensions. Existing techniques trade temporal data reuse for inefficiencies in other areas, such as load im-balance, redundant computations, or increased control flow overhead, therefore making it challenging for use with GPUs.

We proposed a time-tiling method for iterative stencil computations on GPUs. Our method is the first tiling algorithm solving the following constraints simultaneously: it does not involve redundant computations, it favors coalesced global-memory accesses, data reuse in local/shared-memory or cache, avoidance of thread divergence, and concurrency, combining hexagonal tile shapes along the time and one spatial dimension with classical tiling along the other spatial dimensions. Hexagonal tiles expose multi-level parallelism as well as data reuse. Experimental results demonstrate significant performance improvements over existing stencil compilers.

Part of this work also involved our colleagues from the POLYFLOW associate-team at the Indian Institute of Science, Bangalore, India.

6.10. Compilation for scalable on-chip parallelism

Participants: Antoniu Pop, Feng Li, Sven Verdoolaege, Govindarajan Ramaswamy, Albert Cohen.

Task-parallel programming models are getting increasingly popular. Many of them provide expressive mechanisms for inter-task synchronization. For example, OpenMP 4.0 will integrate data-driven execution semantics derived from the StarSs research language. Compared to data-parallel and fork-join models of parallelism, the advanced features being introduced into task-parallel models in turn enable improved scalability through load balancing, memory latency mitigation, mitigation of the pressure on memory bandwidth, and as a side effect, reduced power consumption.

We developed a systematic approach to compile a loop nest into concurrent, dependent tasks. We formulated a partitioning scheme based on the tile-to-tile dependences, represented as affine polyhedra. This scheme ensures at compilation time that tasks belonging to the same class have the same, fully explicit incoming and outgoing dependence patterns. This alleviates the burden of a full-blown dependence resolver to track the readiness of tasks at run time. We evaluated our approach and algorithms in the PPCG compiler, targeting OpenStream, our experimental data-flow task-parallel language with explicit inter-task dependences and a lightweight runtime. Experimental results demonstrate the effectiveness of the approach.

Part of this work also involved our colleagues from the POLYFLOW associate-team at the Indian Institute of Science, Bangalore, India.

6.11. Correct and efficient runtime systems

Participants: Nhat Minh Lê, Robin Morisset, Adrien Guatto, Antoniu Pop, Francesco Zappa Nardelli, Albert Cohen.

User-space scheduling and concurrent first-in first-out queues are two essential building blocks of parallel programming runtimes. They are, however, rarely used together since typical schedulers are oblivious to the ordering constraints introduced by buffered communication.

Chase and Lev's concurrent deque is a key data structure in shared-memory parallel programming and plays an essential role in work-stealing schedulers. We provided the first correctness proof of an optimized implementation of Chase and Lev's deque on top of the POWER and ARM architectures: these provide very relaxed memory models, which we exploit to improve performance but considerably complicate the reasoning. We also studied an optimized x86 and a portable C11 implementation, conducting systematic experiments to evaluate the impact of memory barrier optimizations. Our results demonstrate the benefits of hand tuning the deque code when running on top of relaxed memory models.

Based on this early success, we started working on a more global solution using a new lock-free algorithm for stalling and waking-up tasks in a user-space scheduler according to changes in the state of the corresponding queues. The algorithm is portable and correct, since it is written and proven against the C11 memory model. We showed through experiments that it can serve as a keystone to efficient parallel runtime systems.

These efforts underline the parallelizing compilation research for n-synchronous languages, and the scalable parallel execution of OpenStream.

6.12. Checking Synchronous Compiler Correctness

Participants: Francesco Zappa Nardelli, Guillaume Chelfi, Marc Pouzet.

During year 2013, we have worked on the use of formal verification of compilation steps in the compiler of a Lustre-like synchronous language. Two main directions has been taken:

- The use of SMT-based k-induction techniques to verify the correctness of the successive steps of a synchronous compiler. We used the tool KIND developed by Cesare Tinelli (Iowa state Univ.) and applied it to the Heptagon compiler. The compiler does several source-to-source transformations upto sequential code and KIND was used to verify the equivalence between those successive steps. We came to the conclusion that for most programs, equivalence checking fails unless extra traceability information is added by the compiler.
- The development of a dedicated verification technique to prove the equivalence between a Lustre program and its sequential implementation. We plan to pursue this work during year 2014. Cesare Tinelli will be visiting professor for a month during June 2014.

PAROLE Project-Team

6. New Results

6.1. Speech analysis and synthesis

Participants: Anne Bonneau, Vincent Colotte, Dominique Fohr, Yves Laprie, Joseph Di Martino, Slim Ouni, Agnès Piquard-Kipffer, Emmanuel Vincent, Utpala Musti.

Signal processing, phonetics, health, perception, articulatory models, speech production, learning language, hearing help, speech analysis, acoustic cues, speech synthesis

6.1.1. Acoustic-to-articulatory inversion

The acoustic-to-articulatory inversion from cepstral data has been evaluated on the X-ray database, i.e. X-ray films recorded with the original speech signal. A codebook is used to represent the forward articulatory to acoustic mapping and we designed a loose matching algorithm using spectral peaks to access it. This algorithm, based on dynamic programming, allows some peaks in either synthetic spectra (stored in the codebook) or natural spectra (to be inverted) to be omitted. Quadratic programming is used to improve the acoustic proximity near each good candidate found during codebook exploration. The inversion [40], [10] has been tested on speech signals corresponding to the X-ray films. It achieves a very good geometric precision of 1.5 mm over the whole tongue shape unlike similar works which limit the error evaluation at 3 or 4 points corresponding to sensors located at the front of the tongue.

6.1.1.1. Construction of articulatory models

Articulatory models are intended to approximate the vocal tract geometry with a small number of parameters controlling linear deformation modes. Most of the models have been designed on images of vowels and thus offer a good coverage for vowels but are unable to provide a good approximation for consonants, especially in the region of the constriction. The first problem is related to the nature of contours used to derive linear components. When dealing with vowels there is no contact between the tongue and other fixed articulators (palate, teeth). Factor analysis used to determine linear modes of deformation of the tongue only takes into account the influence of the tongue muscles. This is no longer the case with consonants, since a contact is realized between the tongue and the palate, alveolar ridge or teeth for stops /k, g, t, d/ and the sonorant /l/ in French. The deformation factors thus incorporate the “clipping” effect of the palate. Following the idea of using virtual articulatory targets that lie beyond the positions that can be reached, here the palate, we edited delineated tongue contours presenting a contact with the palate. We chose a conservative solution which consists of keeping the tongue contour up to the contact point and extending it while guaranteeing a “natural shape”. These new contours do not cross the palate for more than 10 mm. As such, this modification alone is not sufficient, because the number of images corresponding to consonants is small even if the corpus used in this work is phonetically balanced. We thus duplicated a number of consonant X-ray images in order to increase the weight of deformation factors corresponding to the tongue tip which is essential for some consonants, /l/ for instance. This approach provides a very good fitting with original tongue contours, i.e. 0.83 mm in average with 6 components over the whole tongue contour and only 0.56 mm in the region of the main place of articulation, which is important with a view of synthesizing speech.

6.1.1.2. Articulatory copy synthesis

Acoustic features and articulatory gestures have always been studied separately. Articulatory synthesis could offer a nice solution to study both domains simultaneously provided that relevant information can be fed into the acoustic simulation. The first step consisted of connecting the 2D geometry given by mediosagittal images of the vocal tract with the acoustic simulation. Last year we thus developed an algorithm to compute the centerline of the vocal tract, i.e. a line which is approximately perpendicular to the wave front. The centerline is then used to segment the vocal tract into elementary tubes whose acoustic equivalents are fed into the acoustic simulation. A new version of the centerline algorithm [53] has been developed in order to approximate the propagation of a plane wave more correctly.

The work on the development of time patterns used to pilot the acoustic simulation has been continued by improving the choice of relevant X-ray images and the temporal transitions from one image to the following. This procedure has been applied successfully to copy sentences and VCV for four X-ray films of the DOCVACIM database[52]. More difficult transitions, i.e. those corresponding to consonant clusters, will be investigated this year.

In addition to the control of the acoustic simulation we started an informal cooperation with the IADI laboratory www.iadi-nancy.fr in order to record better static images of the vocal tract, and cineMRI, i.e. films, for a number of sentences.

6.1.2. Using articulography for speech animation

We are continuously working on the acquisition and analysis of the articulatory data using electromagnetic articulography (EMA).This year, we have conducted research to use EMA as motion capture data and we showed that it is possible to use it for audiovisual speech animation. In fact, as EMA captures the position and orientation of a number of markers, attached to the articulators, during speech, it performs the same function for speech that conventional motion capture does for full-body movements acquired with optical modalities, a long-time staple technique of the animation industry. We have processed EMA data from a motion-capture perspective and applied to the visualization of an existing multimodal corpus of articulatory data, creating a kinematic 3D model of the tongue and teeth by adapting a conventional motion capture based animation paradigm. Such an animated model can then be easily integrated into multimedia applications as a digital asset, allowing the analysis of speech production in an intuitive and accessible manner. In this work [61], we have addressed the processing of the EMA data, its co-registration with 3D data from vocal tract magnetic resonance imaging (MRI) and dental scans, and the modeling workflow. We will continue our effort in the future to improve this technique.

6.1.3. Acoustic analyses of non-native speech

Within the framework of the project IFCASL, we designed a corpus for the study of French and German, with both languages pronounced by French and German speakers, so as to put into light L1/L2 interferences. The corpus was constructed to control for several segmental and suprasegmental phenomena. German and French, for instance, show different kinds of voicing patterns. Whereas in French, the voicing opposition of stops is realized as voiced versus unvoiced, in German, the same difference is realized mostly as unaspirated versus aspirated. Furthermore, differences between the two language groups are expected with respect to the production of nasal vowels (absent in German), the realization of /h/ (not present in French, but in German). On the suprasegmental level, word stress and focus intonation are central to our investigation. Speakers produce both native and non-native speech, which allows for a parallel investigation of both languages.

We have conducted a pilot study on the realization of obstruents in word-final position -a typical example of L1-L2 interference on the segmental level-, which are subject to devoicing in German, but not in French. First results showed that German learners (beginners) had difficulties to voice French obstruents in this context, and, when listening to French realizations, tend to add a final schwa to achieve the expected realization.

6.1.4. Speech synthesis

We recall that within the framework of the ViSAC project we have developed bimodal acoustic-visual synthesis technique that concurrently generates the acoustic speech signal and a 3D animation of the speaker's outer face. This is done by concatenating bimodal diphone units that consist of both acoustic and visual information. In the visual domain, we mainly focus on the dynamics of the face rather than on rendering. The proposed technique overcomes the problems of asynchrony and incoherence inherent in classic approaches to audiovisual synthesis. The different synthesis steps are similar to typical concatenative speech synthesis but are generalized to the acoustic-visual domain. This year we have performed an extensive evaluation of the synthesis system using perceptual and subjective evaluations. The overall outcome of the evaluation indicates that the proposed bimodal acoustic-visual synthesis technique provides intelligible speech in both acoustic and visual channels [22]. For testing purposes we have also added a simple tongue model that is controlled by the generated phonemes. The purpose is to improve the quality of the audiovisual speech intelligibility.

Morover, we perform feature selection and weight tuning for a given unit-selection corpus to make the ranking given by the target cost function consistent with the ordering given by an objective dissimilarity measure. To find an objective metric highly correlated to perception we analyzed correlation between objective and subjective evaluation results. It shows interesting patterns which might help in designing better tuning metrics and objective evaluation techniques [55].

6.1.5. Phonemic discrimination evaluation in language acquisition and in dyslexia and dysphasia

We keep working on a project concerning identification of early predictors of reading, reading acquisition and language difficulties, more precisely in the field of specific developmental disabilities : dyslexia and dysphasia. A fair proportion of those children show a weakness in phonological skills, particularly in phonemic discrimination. However, the precise nature and the origin of the phonological deficits remain unspecified. In the field of dyslexia and normal acquisition of reading, our first goal was to contribute to identify early indicators of the future reading level of children. We based our work on the longitudinal study - with 85 French children - of [90], [91] which indicates that phonemic discrimination at the beginning of kindergarten is strongly linked to success and specific failure in reading acquisition. We study now the link between oral discrimination both with oral comprehension and written comprehension. Our analyses are based on the follow up of a hundred children for 4 years from kindergarten to end of grade 2 (from age 4 to age 8) [98].

6.1.6. Enhancement of esophageal voice

6.1.6.1. Pitch detection

Over the last two years, we have proposed two new real time pitch detection algorithms (PDAs) based on the circular autocorrelation of the glottal excitation, weighted by temporal functions, derived from the CATE [85] original algorithm (Circular Autocorrelation of the Temporal Excitation), proposed initially by J. Di Martino and Y. Laprie. In fact, this latter algorithm is not constructively real time because it uses a post-processing technique for the Voiced/Unvoiced (V/UV) decision. The first algorithm we developed is the eCATE algorithm (enhanced CATE) that uses a simple V/UV decision less robust than the one proposed later in the eCATE+ algorithm. We propose a recent modified version called the eCATE++ algorithm which focuses especially on the detection of the F0, the tracking of the pitch and the voicing decision in real time. The objective of the eCATE++ algorithm consists in providing low classification errors in order to obtain a perfect alignment with the pitch contours extracted from the Bagshaw or Keele databases by using robust voicing decision techniques. This algorithm has been published in Signal, Image and Video Processing, [14].

6.1.6.2. Real-time pitch detection for application to pathological voices

The work first rested on the CATE algorithm developed by Joseph Di Martino and Yves Laprie, in Nancy, 1999.The CATE (Circular Autocorrelation of the Temporal Excitation) algorithm is based on the computation of the autocorrelation of the temporal excitation signal which is extracted from the speech log-spectrum. We tested the performance of the parameters using Bagshaw database, which is constituted of fifty sentences, pronounced by a male and a female speaker. The reference signal is recorded simultaneously with a microphone and a laryngograph in an acoustically isolated room. These data are used for the calculation of the contour of the pitch reference. When the new optimal parameters from the CATE algorithm were calculated, we carried out statistical tests with the C functions provided by Paul BAGSHAW. The results obtained were very satisfactory and a first publication relative to this work was accepted and presented at the ISIVC 2010 conference [79]. At the same time, we improved the voiced / unvoiced decision by using a clever majority vote algorithm electing the actual F0 index candidate. Recently Fadoua Bahja developed a new algorithm based on wavelet transforms applied to the cepstrum excitation. The preliminary results obtained were satisfactory and a complete description of this latter study is under a submission process in an international journal.

6.1.6.3. Voice conversion techniques applied to pathological voice repair

Voice conversion is a technique that modifies a source speaker's speech to be perceived as if a target speaker had spoken it. One of the most commonly used techniques is the conversion by GMM (Gaussian Mixture Model). This model, proposed by Stylianou, allows for efficient statistical modeling of the acoustic space of a speaker. Let “x” be a sequence of vectors characterizing a spectral sentence pronounced by the source speaker and “y” be a sequence of vectors describing the same sentence pronounced by the target speaker. The goal is to estimate a function F that can transform each source vector as nearest as possible of the corresponding target vector. In the literature, two methods using GMM models have been developed: In the first method (stylianou,98), the GMM parameters are determined by minimizing a mean squared distance between the transformed vectors and target vectors. In the second method (kain,98), source and target vectors are combined in a single vector “z”. Then, the joint distribution parameters of source and target speakers is estimated using the EM optimization technique. Contrary to these two well known techniques, the transform function F, in our laboratory, is statistically computed directly from the data: no needs of EM or LSM techniques are necessary. On the other hand, F is refined by an iterative process. The consequence of this strategy is that the estimation of F is robust and is obtained in a reasonable lapse of time. Recently, we realized that one of the most important problems in speech conversion is the prediction of the excitation. In order to solve this problem we developed a new strategy based on the prediction of the cepstrum excitation pulses. Another very important problem in voice conversion concerns the prediction of the phase spectra. This study is under progress in the framework of an Inria ADT which began in September 2013.

6.1.6.4. Signal reconstruction from short-time Fourier transform magnitude spectra

Joseph Di Martino and Laurent Pierron developed in 2010 an algorithm for real-time signal reconstruction from short-time Fourier magnitude spectra [86]. Such an algorithm has been designed in order to enable voice conversion techniques we are developing in Nancy for pathological voice repair. Recently Mouhcine Chami, an assistant-professor of the INPT institute at Rabat (Morocco) proposed a hardware implementation of this algorithm using FPGAs. This implementation has been published in the SIIE 2012 conference [81]. Maryem Immassi, a PhD student of Mouhcine Chami, is comparing this algorithm with the state of the art RTISI-LA algorithm in the framework of a hardware implementation.

6.1.7. Audio source separation

Audio source separation is the task of extracting one or more target source signals from a given mixture signal. It is an inverse problem, which requires the user to guide the separation process using prior models for the source signals and the mixing filters or for the source spectra and their spatial covariance matrices. We studied the impact of sparsity penalties over the mixing filters [38] and we defined probabilistic priors [20] and deterministic subspace constraints [45] over the spatial covariance matrices. We also wrote a review paper about guided audio source separation for *IEEE Signal Processing Magazine* [28].

This paper highlighted that many guided separation techniques now exist that are closer than ever to successful industrial applications, as exemplified by the ongoing industrial collaborations of the team. In order to exploit our know-how for these real-world applications, we investigated issues such as the impact of audio coding [59], artifact reduction [21], real-time implementation [62], and latency [70]. Two patents have been filed [77], [76]. We also started a new research track on the fusion of multiple source separation techniques [46].

Finally, we pursued our long-lasting efforts on the evaluation of audio source separation by collecting the first-ever publicly available dataset of multichannel real-world noise recordings [71] and by conducting an experimental comparison of the two main families of techniques used for source separation [63].

6.2. Automatic speech recognition

Participants: Dominique Fohr, Jean-Paul Haton, Irina Illina, Denis Jouvet, Odile Mella, Emmanuel Vincent, Arseniy Gorin, Luiza Orosanu, Dung Tran.

stochastic models, acoustic models, language models, automatic speech recognition, speech transcription, training, robustness

6.2.1. Detailed acoustic modeling

Acoustic models aim at representing the acoustic features that are observed for the sounds of the language, as well as for non-speech events (silence, noise, ...). Currently context-dependent hidden Markov models (CD-HMM) constitute the state of the art for speech recognition. However, for text-speech alignment, simpler context-independent models are used as they provide better performance.

The use of larger speech training corpora allows us increasing the size of the acoustic models (more parameters through more Gaussians components per density, and more shared densities) and this leads to improved performance. However, in such approaches, Gaussian components are estimated independently for each density. Thus, after having investigated last year the usage of multiple modeling approaches for better constraining the acoustic decoding space, recent studies have focused on enriching the acoustic models themselves in view of handling trajectory and speaker consistency in decoding.

This year a new modeling approach was developed that takes benefit of the multiple modeling ideas and involves a sharing of parameters. The idea is to use the multiple modeling approach to partition the acoustic space according to classes (manual classes or automatic classification). Then, for each density, some Gaussian components are estimated on the data of each class. These class-based Gaussian components are then pooled to provide the set of Gaussian components of the density. Finally class dependent mixture weights are estimated for each density. The method allows us to better parameterize GMM-HMM without increasing significantly the number of model parameters. The experiments on French radio broadcast news data demonstrate the improvement of the accuracy with such parameterization compared to the models with similar, or even larger number of parameters [43].

Current experiments deal with stranded HMM. The objective of such an approach is to introduce in the GMM-HMM modeling some extra parameters to take into account the transition between the Gaussian components when moving from one frame to the next.

6.2.2. Noise-robust speech recognition

In many real-world conditions, the speech signal is overlapped with noise, including environmental sounds, music, or undesired extra speech. Source separation may then be used as a pre-processing stage to enhance the desired speech signal [64]. In practice, the enhanced signal always includes some distortions compared to the original clean signal. It is important to quantify which parts of the enhanced signal are reliable in order not to propagate these distortions to the subsequent feature extraction and decoding stages. A number of heuristic statistical uncertainty estimators and propagators have been proposed to this aim. We started some work aiming to improve the accuracy of these estimators and propagators. We also showed how to exploit uncertainty in order to train unbiased acoustic models directly from noisy data [24].

In order to motivate further work by the community, we created a new international evaluation campaign on that topic in 2011: the CHiME Speech Separation and Recognition Challenge. This challenge aims to recognize small or medium-vocabulary speech mixed with noise recorded in a real family home over the course of several weeks. We analyzed the outcomes of the first edition [16] which led to a special issue of *Computer Speech and Language* [15] and we organized a second edition in 2013 [66] which illustrated the progress made in two years over small-vocabulary speech and the remaining challenges towards robust recognition of medium-vocabulary speech [65].

6.2.3. Linguistic modeling

Usually the lexicon used by a speech recognition system refers to word entries, where each entry in the pronunciation lexicon specifies a possible pronunciation of a word, and the associated language model specifies the probability of a word knowing preceding words. However, whatever the size of the lexicon is, the size is always finite, and the speech recognition system cannot recognize properly words that are not present in the lexicon. In such cases, the unknown word is typically replaced by a sequence of short words which is acoustically similar to the unknown speech portion.

6.2.3.1. Random indexing

This year we studied the introduction of semantic information through the Random Indexing paradigm (RI) in statistical language models used in speech recognition. Random Indexing is a scalable alternative to LSA (Latent Semantic Analysis) for analyzing relationships between a set of documents and the terms they contain. We determined the best methods and parameters by minimizing the perplexity of a realistic corpus of 290000 words. We investigated 4 methods for training RI matrices, 4 weighting functions, several matrix sizes and how balancing the 4-gram and RI language model. We only obtained a relative gain of 3% [42].

6.2.3.2. Continuous language models

Language modeling plays an important role in automatic speech recognition because it constrains the decoder to search the most likely sequences of words according to a given language and a given task. A limitation of N-grams models is that they represent the words in a discrete space. It would be interesting to represent words in a continuous space where semantically close words would be projected in the same region of space. This projection can be achieved by recurrent neural networks. Moreover they are able to learn long-term dependencies with the recurrent layer that can store a record of the past. During his master internship, Othman Zennaki integrated this new language model in our speech recognition system ANTS.

6.2.3.3. Linguistic units for embedded systems

In the framework of the RAPSODIE project, speech recognition is to be used to help communication with hard of hearing people. Because of requirements on memory and CPU (almost real time processing), various modeling approaches have been investigated with respect to linguistic units. The first approach has focused on analyzing the achieved phonetic decoding performance of various linguistic units (phonemes, syllables, words). Best phonetic decoding performance is achieved using word units and associated tri-gram language model, but at the expense of large CPU and memory requirements. Using directly phoneme units leads to the smallest models and requires little CPU, however, this also leads to the worst performance. The proposed approach relying on syllable units provides results which are rather close to the word based approach, but requires much less CPU [58], [57].

Further experiments are now focusing on combining word and syllable units, in view of having frequent words covered by the word units, and using syllables for decoding unknown words.

6.2.3.4. OOV proper name retrieval

Proper name recognition is a challenging task in information retrieval in large audio/video databases. Proper names are semantically rich and are usually key to understanding the information contained in a document.

In the framework of the ContNomina project, we focus on increasing the vocabulary coverage of a speech transcription system by automatically retrieving proper names from contemporary diachronic text documents. We proposed methods that dynamically augment the automatic speech recognition system vocabulary, using lexical and temporal features in diachronic documents. We also studied different metrics for proper name selection in order to limit the vocabulary augmentation and therefore the impact on the ASR performances. Recognition results show a significant reduction of the word error rate using augmented vocabulary [56].

6.2.4. Speech transcription

The first complete version of the speech transcription system ANTS (see section 5.5) has been initially developed in the framework of the Technolangue project ESTER, and since then, the system has been regularly enriched through the integration of research results. The latest version can handle either HTK-based acoustic models through the Julius decoder, or Sphinx-based acoustic models with the CMU Sphinx decoders. In the last version, a Perl script encapsulates all the calls to the various tools used for diarization, model adaptation and speech recognition, and takes benefit of the multiple CPU available on the computer for parallelizing the different tasks as much as possible.

6.2.4.1. Combining recognizers

Last year in the context of the ETAPE speech transcription evaluation campaign, the Sphinx-based and Julius-based decoders have been further improved, and it was observed that combining the recognition outputs of several Sphinx-based and Julius-based decoder lead to a significant word error rate reduction compared to the best individual system.

More controlled experiments have then been performed to understand what was the main reason of the large performance improvement observed when combining Julius-based and Sphinx-based transcription system results. The Sphinx decoder processes the speech data in a forward pass, whereas the Julius decoder ends its decoding process by a backward pass. The Sphinx training and decoding scripts have been modified to process the speech material in a reverse time order; and various systems were developed by using different sets of acoustic features and different sets of acoustic units. It was then observed that combining several Sphinx-forward and several Sphinx-reverse decoders lead to much better results than combining the same amount of only Sphinx-forward decoders or only Sphinx-reverse decoders; and the achieved word error rate was consistent with the one obtained by combining the Sphinx-based (forward) and Julius-based (backward) decoders [49]. Hence, the improvement is mainly due to the fact that forward-based and backward-based processing are combined. Because heuristics are applied during decoding to limit the acoustic space that is explored, some hypotheses might be wrongly pruned when processing the data one way, and may be kept in the active beam search when processing the other way. This is corroborated by the analysis of the word graph which show a large dissimilarity in the distribution of the number of words starting and ending in each frame [48].

Experiments have also shown that when the forward and backward decoders yield the same word hypothesis, this word is likely to be a correct answer. Recent experiments are investigating how far such behavior could help for unsupervised learning of acoustic models.

6.2.4.2. Spontaneous speech

During his master intership, Bruno Andriamiarina focuses on the new challenges brought by this spontaneity of the speech, making it difficult to be transcribed by the existing automatic speech recognition systems. He studied how to improve global performance of automatic speech recognition systems when dealing with spontaneous speech by adapting language model and pronunciation dictionary to this particular type of speech. He also studied the detection of disfluent speech portions (produced by spontaneous speech) in speech signal using a Gaussian Mixture Model (GMM)-based classifier trained on prosodic features covering the main prosodic characteristics (duration, fundamental frequency and energy).

6.2.4.3. Towards a structured output

The automatic detection of the prosodic structure of speech utterances has been investigated. The algorithm relies on a hierarchical representation of the prosodic organization of the speech utterances, and detects prosodic boundaries whether they are followed or not by pause. The detection of the prosodic boundaries and of the prosodic structures is based on an approach that integrates little linguistic knowledge and mainly uses the amplitude of the F0 slopes and the inversion of the F0 slopes as well as phone durations. The approach was applied on a corpus of radio French broadcast news and also on radio and TV shows which are more spontaneous speech data. The automatic prosodic segmentation results were then compared to a manual prosodic segmentation made by an expert phonetician [37].

Further work has focused on analyzing the links between manually set punctuation marks and this automatically detected prosodic structure, in view of using the prosodic structure for helping an automatic punctuation process.

6.2.5. Speech/text alignment

6.2.5.1. Alignment with non-native speech

Non-native speech alignment with text is one critical step in computer assisted foreign language learning. The alignment is necessary to analyze the learner's utterance, in view of providing some prosody feedback (as for example bad duration of some syllables - too short or too long -). However, non-native speech alignment with text is much more complicated than native speech alignment. This is due to the pronunciation deviations observed on non-native speech, as for example the replacement of some target language phonemes by phonemes of the mother tongue, as well as errors in the pronunciations.

In the case of French speakers learning English, we conducted a detailed analysis that has showed the benefit of taking into account non-native variants, and lead to determining the classes of phonemes whose temporal boundaries are the most accurate and which should be favored in the design of exercises for language learning[18].

In the framework of the IFCASL project, we proposed to use a two-step approach for automatic phone segmentation. The first step consists in determining the phone sequence that best explains the learner's utterance. This is achieved by force aligning the learner's speech utterance with a model representing the various possible pronunciation variants of the current sentence (both native and non-native variants need to be considered). In this step detailed acoustic Hidden Markov Models (HMMs) are used, with a rather large number of Gaussian components per mixture density. This kind of detailed acoustic models is the one that provides the best performance in automatic speech recognition. The second step consists in determining the phone boundaries. This is also achieved through a forced alignment process, but this time, the sequence of phones is known (as determined in the first step), and phone acoustic models with only a few Gaussians components per mixture density are used because it has been shown that they provide better temporal precision than detailed acoustic models. For the training of the models used for both forced alignment steps, the speech of native and non-native speakers could be used, either directly or by MLLR (Maximum Likelihood Linear Regression) adaptation.

6.2.5.2. Alignment with spontaneous speech

In the framework of the ANR ORFEO, we addressed the problem of the alignment of spontaneous speech. The ORFEO audio files were recorded under various conditions with a large SNR range and contain extra speech phenomena and overlapping speech. As regards overlapping speech, the orthographic transcription of the audio files only provides a rather imprecise time information of the overlapping speech segment. As a first approach, among the different orthographic transcripts corresponding to the overlapping area, we determined as the main transcript the one that best matches the audio signal, the others are kept in other tiers with the same time boundaries.

6.3. Machine translation and language modeling

Participants: Kamel Smaïli, David Langlois, Denis Jouvet, Emmanuel Vincent, Motaz Saad, Cyrine Nasri.

machine translation, statistical models

6.3.1. Language modeling

6.3.1.1. Vocabulary selection

In the framework of the ETAPE evaluation campaign a new machine learning based process was developed to select the most relevant lexicon to be used for the transcription of the speech data (radio and TV shows). The approach relies on a neural network trained to distinguish between words that are relevant for the task and those that are not. After training, the neural network (NN) is applied to each possible word (text tokens extracted from a very large text corpus). Then the words that have the largest NN output score are selected for creating the speech recognition lexicon. Such an approach can handle counts of occurrences of the words in various data subsets, as well as other complementary information, and thus offer more perspectives than the traditional unigram-based selection procedures [50].

6.3.1.2. Music language modeling

Similarly to speech, music involves several levels of information, from the acoustic signal up to cognitive quantities such as composer style or key, through mid-level quantities such as a musical score or a sequence of chords. The dependencies between mid-level and lower- or higher-level information can be represented through acoustic models and language models, respectively. We pursued our pioneering work on music language modeling, with a particular focus on log-linear interpolation of multiple conditional distributions. We applied it to the joint modeling of "horizontal" (sequential) and "vertical" (simultaneous) dependencies between notes for polyphonic pitch estimation [26] and to the joint modeling of melody, key and chords for automatic melody harmonization [25]. We also proposed a new Bayesian n-gram topic modeling and estimation technique, which we applied to genre-dependent modeling of chord sequences and to music genre classification [74].

6.3.2. Quality estimation of machine translation

In the scope of Confidence Measures, we participated to the World Machine Translation evaluation campaign for the second year (WMT2013 http://www.statmt.org/wmt13/quality-estimation-task.html). More precisely, we proposed a Quality Estimation system to the Quality Estimation shared task. The goal was to predict the quality of translations generated by an automatic system. Each translated sentence is given a score between 0 and 1. The score is obtained by using several numerical or boolean features calculated according to the source and target sentences. We performed a linear regression of the feature space against scores in the range [0 ;1], to this end, we use a Support Vector Machine with 66 features. In this new participation, we proposed to increase the size of the training corpus. For that, we decided to use the post-edited and reference corpora in the training step after assigning a score to each sentence of these corpora. Then, we tune these scores on a development corpus. This leads to an improvement of 10.5% on the development corpus, in terms of Mean Average Error (average difference between reference and predicted scores), but achieves only a slight improvement on the test corpus. This work has been published in [51].

6.3.3. Comparable corpora and multilingual sentiment analysis

In the PhD Thesis of Motaz Saad, we work on collecting comparable corpora. For that purpose we presented a method which extracts and aligns comparable corpora at the article level from Wikipedia encyclopedia based on interlanguage links. To evaluate the closeness of corpora we proposed several comparability measures. Our evaluations show that the proposed comparability measures are able to capture the comparability degree of any comparable corpora [60]. We go further on the comparability of multilingual corpora by studying their comparability in terms of sentiment. The final objective is to propose a multilingual press review concerning a given topic. This review should use several multilingual resources (electronic newspapers), and should class resources according to the including sentiments (fear, joy...about the subject), polarity (against or not to the subject)...This conducts to study opinions across different languages by comparing the underlying messages written by different people having different opinions. We propose "Sentiment based Comparability Measures" to compare opinions in multilingual comparable articles without translating source/target into the same language [27].

6.3.4. Machine translation of arabic dialect

The translation of Arabic dialect constitutes a real challenge since it is an under-resourced language. In fact, Modern Standard Arabic is as any other evoluated language, it means it could be processed by the available tools but unfortunately in Arabic countries people speak an Arabic language which is inspired from the standard one but is different. Our objective is then to propose a speech to speech system converting modern standard Arabic to Algerian dialect. After collecting corpus, we decided to propose a method allowing to diacritize dialects in order to be able in the following to develop an acoustic model. For that, we considered the issue of diacritization as a machine translation issue, and we have developed a statistical machine translation which learns to transform an undiacritized corpus into a diacritized one [44].

PARSIFAL Project-Team

6. New Results

6.1. Substitution as Proof Compression

Participants: Lutz Straßburger, Novak Novakovic.

In previous work [58] we have shown how the calculus of structures can accommodate Tseitin extension without relying on the cut (or modus ponens). Thus, cut and extension can be studied independently as proof compression mechanisms. Another such proof compression mechanism is substitution. It has been shown by Cook, Reckhow, Krajíček and Pudlák that in the presence of cut, extension and substitution are equally powerful with respect to proof complexity. This year we succeeded in showing that this is also the case in the absence of cut. I.e., we have shown that the cut-free system with extension and the cut-free system with substitution p-simulate each other. This result is presented in [34].

6.2. Herbrand Confluence

Participants: Lutz Straßburger, Stefan Hetzl.

In the result on Herbrand confluence from last year [46], the endsequent of a proof had to be an existential sentence in prenex form. This year we were able to relax this restriction and to extend our result to arbitrary endsequent. This work has been published in [15].

6.3. Nested Sequents for Intuitionistic Modal Logics

Participant: Lutz Straßburger.

We present cut-free deductive systems without labels for the intuitionistic variants of the modal logics obtained by extending IK with a subset of the axioms d, t, b, 4, and 5. For this, we use the formalism of nested sequents, which allows us to give a uniform cut elimination argument for all 15 logic in the intuitionistic S5 cube. This work (published in [25]), is an improvement of the result on intuitionistic modal logic from 2011: the deductive systems the cut elimination proof are much simpler now.

6.4. First efforts at designing proof certificates

Participants: Hichem Chihani, Quentin Heath, Dale Miller, Fabien Renaud.

Work on the ERC Advance Grant ProofCert has progressed along two lines.

Given earlier work within the team [6], [7], there now exists a flexible and well understood concept of focused proof for classical and intuitionistic first-order logics. Chihani, Miller, and Renaud have been working to use that notion of proof as a means of providing flexible definition of *proof evidence* for those two logics. Initial results along those directions have been reported in the [19] and [20]. In those papers, several examples definitions of the semantics of *proof certificates* (formal documents providing the details of some proof evidence) are provided in such a way that a single, simple proof checker can formally elaborate that evidence into a focused sequent calculus. Such an elaboration thus guarantees the soundness of that proof. These papers also describe a "reference proof checker" that has been built with the expectation that its formal correctness can be established. That checker is also able to do bounded *proof reconstruction* as well as allow both deterministic and non-deterministic computation to be mixed with deduction.

Our understanding of focused proofs in the presence of both induction and co-induction (inference rules found in model checkers and most theorem provers) is less well developed. As a result, Miller and Tiu have been studied a simple approach of proof certificate in the setting of model checking in the hope of identifying the relevant proof theory designs that need to be developed. In [33], they showed how tabled deduction in model checking can be used to provide a formal proof certificate for a range of co-inductively defined predicates.

6.5. Combinations of classical and intuitionistic logic

Participant: Dale Miller.

Chuck Liang and Miller have been studying the question of how one can mix intuitionistic and classical logic into a single logic. The initial motivation for considering this problem arose from the concerns raised by the ProofCert project of how best to deal with both classical and intuitionistic logic and their associated proof evidence. Will there need to be two different kinds of checkers and two different kinds of libraries for these two different kinds of logics? Will we be able to mix theorems and proofs in one logic with those in the second logic in rich and useful ways?

One way we have considered answering this question is to actually consider a third logic that combines these other two logics. Our work on such combinations is reported in [16], where a thorough analysis of the semantics and proof theory of such a combination is provides, and in [24], where significant examples of the computational aspects of proofs are explored in detail.

6.6. Formal meta theory of sequent calculus

Participant: Dale Miller.

Keeping with the ProofCert theme of finding global, eternal, and formal mechanisms representing proof evidence, Miller and Pimentel describe in [17] a way in which linear logic can be used to formally specify inference rules for a wide range of proof system in several logics. They were able to show that adequacy of their encodings and to provide sufficient conditions for both cut-elimination and initial-elimination to hold for the resulting proof systems. The fact that these elimination results hold or not is an important characteristic for judging a proof system. Using this work, these important questions can be resulted automatically for a wide range of such proof systems.

6.7. The correctness of program using finite precision

Participants: Ivan Gazeau, Dale Miller.

Programs dealing with real number quantities must live with the fact that such numbers are represented using only finite precision. As such, programs that might be considered correct over the abstract field of infinite precision arithmetic can display chaotic and incorrect behaviors when run on actual computer hardware.

One such problem with finite precision is that programs can "leak" information about values that are intended to be hidden or at least obfuscated as happens in the area differential privacy. In [22], Gazeau, Miller, and Palamidessi illustrated just how such attacks on information hiding can be made and how it is possible to add noise to reported data values in such a way that only appropriate amounts of information leakage occurs.

In his PhD thesis, *Safe Programming in finite precision: Controlling the errors and information leaks* (École Polytechnique, 2013 [11]), Gazeau develops that theme further as well as shows how techniques from rewriting theory can be applied to show that, in some situations, the chaotic behavior of finite precision programs can be expected to converge in acceptable time to acceptable answers.

6.8. Sequent Calculus with Calls to a Decision Procedure

Participants: Mahfuza Farooque, Stéphane Graham-Lengrand.

In the PSI project, a version of the focused sequent calculus (for first-order classical logic) has been designed, which can call external decision procedures. Several results were achieved in 2013 since the last Activity Report:

Firstly, a bug was discovered in the proof of cut-elimination, which was used to prove the logical completeness of the calculus. Fixing the bug required minor changes in the definition of the system, but incurred a major re-development of the meta-theory. Out of this technical work, one idea emerged: in presence of a non-trivial theory, changing the polarity of literals may change the provability of formulas. This was quite unexpected, but it led to interesting issues, such as finding sufficient conditions on polarities to guarantee cut-elimination and logical completeness. An substantial achievement in this research topic was to successfully address such issues, which gave rise to a new version of the report [30].

Secondly, more techniques from automated reasoning were captured as proof-search in this sequent calculus (the incremental construction of proof-trees): besides the SMT-solving algorithm DPLL(T) treated successfully in 2012 (which was written down and published this year in [21]), the techniques of *clause tableaux* and *connection tableaux* were captured this year. This includes in particular a notion of *clause tableaux modulo theories* that C. Tinelli introduced in 2007 [60]. This new range of captured techniques is interesting as clause tableaux are designed to handle quantifiers, which DPLL(T) does not. This gives a new hope to combine the efficiency of SAT-solvers for propositional reasoning with the handling of quantifiers.

6.9. Path Functors in the Category of Small Categories

Participant: François Lamarche.

In [31] François Lamarche gives a detailed description of two path functors in the category of small categories, which he calls $\mathbf{Pe}$ and $\mathbf{P}$, and proves some of their important properties. The second of these is the functor which is used to model the Martin-Löf identity type in [47]; it associates to every small category X an internal category structure whose object of objects is X; one important theorem which is proved in [31] is that the category of internal (co- or contravariant) presheaves on $\mathbf{P}X$ coincides with the category of Grothendieck bifibrations over the base X. Thus, through a trivial use of monadic abstract nonsense, we can say that $\mathbf{P}X$ is the free bifribration over X. The category $\mathbf{P}X$ is obtained by taking the bigger $\mathbf{Pe}X$, which is a little more than just a category, being poset-enriched, and getting rid of the order enrichment by quotienting. $\mathbf{Pe}X$ is a more general kind of bifibration than an ordinary Grothendieck bifibration, and the enrichment is necessary to describe its properties, thus taking us outside of the theory 1-categores.

6.10. Subformula Linking as an Interaction Method

Participant: Kaustuv Chaudhuri.

We showed how to generalize the *calculus of structures*, a *deep inference* formalism, for classical linear logic to a *calculus of linking* [18]. This generalization simplifies the calculus by eliminating most of its inference rules. In its place we add a notion of annotation with *links* and a *link resolution* procedure. We show that this is sound and complete with respect to the usual calculus of structures. The linking calculus is the foundational basis of the *Profound* tool described in 5.1 .

6.11. Recovering Proof Structures in the Sequent Calculus

Participants: Kaustuv Chaudhuri, Stefan Hetzl, Dale Miller.

The *sequent calculus* is often criticized as a proof syntax because it contains a lot of noise. It records the precise minute sequence of operations that was used to construct a proof, even when the order of some proof steps in the sequence is irrelevant and when some of the steps are unnecessary or involve detours. These features lead to several technical problems: for example, cut-elimination in the classical sequent calculus LK, as originally developed by Gentzen, is not confluent, and hence proof composition in LK is not associative. Many people choose to discard the sequent calculus when attempting to design a better proof syntax with the desired properties.

In recent years, there has been a project at Parsifal to recover some of these alternative proof syntaxes by imposing a certain abstraction over sequent proofs. Our technique, pioneered at Parsifal, involves the use of *maximal multi-focusing* which gives a syntactic characterization of those sequent proofs that: (1) have a "don't care" ordering of proof steps where the order does not matter, and (2) groups larger logical steps, called *actions*, into a maximally parallel form where only important orderings of actions are recorded. The earliest example of this technique was in [40], where we showed a class of sequent proofs that were isomorphic to proof nets for multiplicative linear logic. In 2012, we were able to obtain a similar result for first-order classical logic, wherein we defined a class of sequent proofs that are isomorphic to expansion proofs, a generalization of Herbrand disjunctions that is in some sense a minimalistic notion of proof for classical logic. This result was published in a preliminary form at the CSL 2012 conference [39].

In 2013 we published an extended paper on this result in the Journal of Logic and Computation [14]. The major contribution here was a detailed proof of the result that gives a precise account of the proof identifications made by expansion proofs.

PERCEPTION Team

6. New Results

6.1. High-resolution depth maps based on TOF-stereo fusion

The combination of range sensors with color cameras can be very useful for a wide range of applications, e.g., robot navigation, semantic perception, manipulation, and telepresence. Several methods of combining range- and color-data have been investigated and successfully used in various robotic applications. Most of these systems suffer from the problems of noise in the range-data and resolution mismatch between the range sensor and the color cameras, since the resolution of current range sensors is much less than the resolution of color cameras. High-resolution depth maps can be obtained using stereo matching, but this often fails to construct accurate depth maps of weakly/repetitively textured scenes, or if the scene exhibits complex self-occlusions. Range sensors provide coarse depth information regardless of presence/absence of texture. The use of a calibrated system, composed of a time-of-flight (TOF) camera and of a stereoscopic camera pair, allows data fusion thus overcoming the weaknesses of both individual sensors. We propose a novel TOF-stereo fusion method based on an efficient seed-growing algorithm which uses the TOF data projected onto the stereo image pair as an initial set of correspondences. These initial "seeds" are then propagated based on a Bayesian model which combines an image similarity score with rough depth priors computed from the low-resolution range data. The overall result is a dense and accurate depth map at the resolution of the color cameras at hand. We show that the proposed algorithm outperforms 2D image-based stereo algorithms and that the results are of higher resolution than off-the-shelf color-range sensors, e.g., Kinect. Moreover, the algorithm potentially exhibits real-time performance on a single CPU. Two journal papers were submitted in 2013 and currently they are under review.

6.2. Continuous action recognition

Continuous action recognition is more challenging than isolated recognition because classification and segmentation must be simultaneously carried out. We build on the well known dynamic time warping (DTW) framework and devise a novel video alignment technique, dynamic *frame* warping (DFW), which performs isolated recognition based on a per-frame representation of videos and on aligning a test sequence with a model sequence. Next we devise two extensions which are able to perform action recognition and video segmentation in a concomitant manner, namely one-pass DFW and two-pass DFW. Both these algorithms have their roots in the continuous speech recognition domain but, to the best of our knowledge, their extension to visual recognition of actions and activities has been overlooked. We test and illustrate the proposed methods with several public-domain datasets and we compare both the isolated and continuous recognition algorithms with several recently published methods. One journal paper was submitted in 2013 and currently is under review.

6.3. High-dimensional regression

We addressed the problem of approximating high-dimensional data with a low-dimensional representation. We make the following contributions. We propose an inverse regression method which exchanges the roles of input and response, such that the low-dimensional variable becomes the regressor, and which is tractable. We introduce a mixture of locally-linear probabilistic mapping model that starts with estimating the parameters of inverse regression, and follows with inferring closed-form solutions for the forward parameters of the high-dimensional regression problem of interest. Moreover, we introduce a partially-latent paradigm, such that the vector-valued response variable is composed of both observed and latent entries, thus being able to deal with data contaminated by experimental artifacts that cannot be explained with noise models. The proposed probabilistic formulation could be viewed as a latent-variable augmentation of regression. We devise expectation-maximization (EM) procedures based on a data augmentation strategy which facilitates the maximum-likelihood search over the model parameters. We propose two augmentation schemes and we describe in detail the associated EM inference procedures that may well be viewed as generalizations of a number of EM regression, dimension reduction, and factor analysis algorithms. The proposed framework is validated with both synthetic and real data. We provide experimental evidence that our method outperforms several existing regression techniques. See [26], [12].

6.4. Simultaneous sound-source separation and localization

Human-robot communication is often faced with the difficult problem of interpreting ambiguous auditory data. For example, the acoustic signals perceived by a humanoid with its on-board microphones contain a mix of sounds such as speech, music, electronic devices, all in the presence of attenuation and reverberations. We proposed a novel method, based on a generative probabilistic model and on active binaural hearing, allowing a robot to robustly perform sound-source separation and localization. We show how interaural spectral cues can be used within a constrained mixture model specifically designed to capture the richness of the data gathered with two microphones mounted onto a human-like artificial head. We describe in detail a novel expectation-maximization (EM) algorithm that alternates between separation and localization, we analyze its initialization, speed of convergence and complexity, and we assess its performance with both simulated and real data. Subsequently, we studied the *binaural manifold*, i.e., the low-dimensional space of sound-source locations embedded in the high-dimensional space of perceived interaural spectral features, and we provided a method for mapping interaural cues onto source locations. See [21], [12]. A journal paper was submitted in 2013 and accepted with minor revisions.

6.5. The geometry of non-coplanar microphone arrays

We addressed the problem of sound-source localization from time-delay estimates using arbitrarily-shaped non-coplanar microphone arrays. A novel geometric formulation is proposed, together with a thorough algebraic analysis and a global optimization solver. The proposed model is thoroughly described and evaluated. The geometric analysis, stemming from the direct acoustic propagation model, leads to necessary and sufficient conditions for a set of time delays to correspond to a unique position in the source space. Such sets of time delays are referred to as *feasible sets*. We formally prove that every feasible set corresponds to exactly one position in the source space, whose value can be recovered using a closed-form localization mapping. Therefore we seek for the optimal feasible set of time delays given, as input, the received microphone signals. This time delay estimation problem is naturally cast into a programming task, constrained by the feasibility conditions derived from the geometric analysis. A global branch-and-bound optimization technique is proposed to solve the problem at hand, hence estimating the best set of feasible time delays and, subsequently, localizing the sound source. Extensive experiments with both simulated and real data are reported; we compare our methodology to four state-of-the-art techniques. This comparison clearly shows that the proposed method combined with the branch-and-bound algorithm outperforms existing methods. These in-depth geometric understanding, practical algorithms, and encouraging results, open several opportunities for future work. See [18], [25], [11].

6.6. Audiovisual calibration and alignment

We addressed the problem of aligning visual (V) and auditory (A) data using a sensor that is composed of a camera-pair and a microphone-pair. The original contribution of the paper is a method for AV data aligning through estimation of the 3D positions of the microphones in the visual-centred coordinate frame defined by the stereo camera-pair. We exploit the fact that these two distinct data sets are conditioned by a common set of parameters, namely the (unknown) 3D trajectory of an AV object, and derive an EM-like algorithm that alternates between the estimation of the microphone-pair position and the estimation of the AV object trajectory. The proposed algorithm has a number of built-in features: it can deal with A and V observations that are misaligned in time, it estimates the reliability of the data, it is robust to outliers in both modalities, and it has proven theoretical convergence. We report experiments with both simulated and real data. See [24] (this work received the best paper award).

6.7. Audiovisual fusion for human-robot interaction

Natural human-robot interaction in complex and unpredictable environments is one of the main research lines in robotics. In typical real-world scenarios, humans are at some distance from the robot and the acquired signals are strongly impaired by noise, reverberations and other interfering sources. In this context, the detection and localisation of speakers plays a key role since it is the pillar on which several tasks (e.g.: speech recognition and speaker tracking) rely. We address the problem of how to detect and localize people that are both seen and heard by a humanoid robot. We introduce a hybrid deterministic/probabilistic model. Indeed, the deterministic component allows us to map the visual information into the auditory space. By means of the probabilistic component, the visual features guide the grouping of the auditory features in order to form AV objects. The proposed model and the associated algorithm are implemented in real-time (17 FPS) using a stereoscopic camera pair and two microphones embedded into the head of the humanoid robot NAO. We performed experiments on (i) synthetic data, (ii) a publicly available data set and (iii) data acquired using the robot. The results we obtained validate the approach and encourage us to further investigate how vision can help robot hearing. See [19], [20], [27], [11], [13]

PHOENIX Project-Team

6. New Results

6.1. Design-Driven Development Methodology for Resilient Computing

Critical systems have to face changes, either to meet new user requirements or because of changes in the execution context. Often, these changes are made at runtime because the system is too critical to be stopped. Such systems are called *resilient systems*. They have to guarantee dependability despite runtime evolution. For example, in the domain of pervasive computing, building management systems (*e.g.,* anti-intrusion, fire protection system, access control) have to be resilient as they are in charge of people safety and have to run in a continuous way.

To mitigate faults at runtime, dependable systems are augmented with fault tolerance mechanisms such as replication techniques or degraded modes of operation. However, these mechanisms cover a large spectrum of areas ranging from hardware to distributed platforms, to software components. As a consequence, the need of fault-tolerance expertise is spread throughout the software development process, making it difficult to trace the dependability requirements. The fault tolerance mechanisms have to be systematically and rigorously applied in order to guarantee the conformance between the application runtime behavior and the dependability requirements. This integration becomes even more complex when taking into account runtime adaptation. Indeed, a change in the execution context of an application may require to adapt the fault tolerance mechanisms. For example, a decrease of the network bandwidth may require to change the replication mechanism for one requiring less network bandwidth (*e.g.,* Leader-Follower Replication instead of Primary-Backup Replication).

Without a clear separation of the functional and fault-tolerance concerns, ensuring dependability becomes a daunting task for programmers, whose outcome is unpredictable. In this context, design-driven development approaches are of paramount importance because the design drives the development of the application while ensuring the traceability of the dependability requirements. However, because most existing approaches are general purpose, their guidance is limited, causing inconsistencies to be introduced in the design and along the development process. This situation calls for an integrated development process centered around a conceptual framework that allows to guide the development process of a resilient application in a systematic manner.

In this work, we propose a novel approach that relies on a design language which is extended with fault-tolerance declarations. To further raise the level of abstraction, our development approach revolves around the Sense-Compute-Control paradigm. The design is then compiled into a customized programming framework that drives and supports the development of the application. To face up changes in the execution context, our development methodology relies on a component-based approach, allowing fine-grained runtime adaptation. This design-driven development approach ensures the traceability of the dependability requirements and preserves the separation of concerns, despite runtime evolution.

This work was funded by the Inria collaboration program (in french, actions de recherches collaboratives). The Serus ARC includes the Phoenix project-team (Bordeaux), the ADAM project-team (Lille) and the TSF-LAAS group (Toulouse). These accomplishments were part of Quentin Enard's PhD studies [14]. This work has been published at the International Conference on Component-based Software Engineering (CBSE'13) [23].

6.2. A Case for Human-Driven Software Development

Human-Computer Interaction (HCI) defines a range of principles and methodologies to design User Interfaces (UIs), aiming (1) to improve the interaction between users and computers, (2) to address how interfaces are implemented, leveraging techniques such as program generation and component architectures, and (3) to propose methods to evaluate and compare interfaces.

Despite the many successes of HCI, when it comes to software development, this domain expertise often does not go beyond guidelines (*e.g.,* ISO/TR 22411:2008 addressing the needs of the elderly and users with disabilities). Sometimes, guidelines are mapped into UI design artifacts. However, for a lack of tools, these artifacts remain contemplative. As a consequence, there exists a gap between UI design and software development. This gap is not typical of the HCI domain. Yet, its consequences are dramatically increasing in importance as software systems intertwine with our daily activities, both professional and domestic. Nowadays, a host of systems are playing a critical role for users in terms of safety, privacy, *etc*.

To bridge the gap between UI design and software development, our approach consists in making UI design a full-fledged dimension of software design. We introduce a language dedicated to designing UIs in a high-level manner, while capturing the key requirements of user interaction. We go beyond a contemplative approach and process a UI design artifact to produce a dedicated programming framework that supports the implementation of all the dimensions expressed in a design artifact. This programming framework guides the stakeholders during the development process, while ensuring the conformance between the UI design and its implementation over time.

This work has been published at the International Conference on Software Engineering (ICSE'13, NIER track) [21].

6.3. Technological Support for Self-Regulation of Children with Autism

Children with Autism Spectrum Disorders (ASD) have difficulties to self-regulate emotions, impeding their inclusion in a range of mainstreamed environments. Self-regulating emotions has been shown to require recognizing emotions and invoking specific coping strategies.

In the context of the School+ research project, we have developed an application dedicated to self-regulating emotions in children with ASD. Ten children with ASD have experimentally tested this tablet-based application over a period of three months in a mainstreamed school. A collaborative learning approach, involving parents, teachers and a school aid, was used 1) to train students to operate the tablet and our application autonomously, and 2) to facilitate the adoption of our intervention tool.

This study shows that our application was successful in enabling students with ASD to self-regulate their emotions in a school environment. Our application helped children with autism to recognize and name their emotions, and to regulate them using idiosyncratic, parent-child, coping strategies, supported by multimedia contents.

This work is in the context of the School+ national research project funded by the French Ministry of National Education. This work is part of Charles Fage's PhD studies.

PI.R2 Project-Team

5. New Results

5.1. Proof-theoretical and effectful investigations

Participants: Pierre Boutillier, Guillaume Claret, Pierre-Louis Curien, Yann-Régis Gianas, Hugo Herbelin, Guillaume Munch-Maccagnoni, Ludovic Patey, Pierre-Marie Pédrot, Alexis Saurin.

5.1.1. Sequent calculus and computational duality

Categorical semantics.

During his collaboration with Marcelo Fiore and Pierre-Louis Curien, Guillaume Munch-Maccagnoni characterised the polarised evaluation order through a categorical structure where the hypothesis that composition is associative is relaxed. Duploid is the name of the structure, as a reference to Jean-Louis Loday's duplicial algebras. The main result, in the lineage of Führmann's [38] direct-style characterisation of monadic models, is a reflection Adj→Dupl where Dupl is a category of duploids and duploid functors, and Adj is the category of adjunctions and pseudo maps of adjunctions. The result suggests that the various biases in denotational semantics: indirect, call-by-value, call-by-name... are a way of hiding the fact that composition is not always associative. This work was accepted for publication in FoSSaCs 2014 [53].

Pierre-Louis Curien, in connection with his increasing interests in operads and algebraic structures of various kinds, found out that the core syntax of system L (underlying the duality of computation) could be used with profit to describe the wiring structures underlying operads, dioperads, cyclic operads, and more generally Lamarche's structads [48]. He also showed a syntactic equivalence between Munch-Maccagnoni's (pre)duploids and system L syntax. These results were presented in his invited talks at the Loday's Mathematical Legacy workshop in Strasbourg and at the workshop Algebra and Computation in Lyon, in January 2014.

Duality of construction.

Paul Downen and Zena Ariola developed a generalized theory of the sequent calculus for understanding the concepts of evaluation strategy and of data (for example, pairs in ML) and co-data (for example, functions) in programming languages. This theory provides a single framework for user-defined data and co-data types as well as a generalized treatment of evaluation strategies, including call-by-value, call-by-name, and call-by-need, that are given as parameters to the theory. In the end, the framework encompasses the previously known duality of call-by-name and call-by-value in the sequent calculus, both by Curien and Herbelin [3] and Wadler [59] , while also including call-by-need and its dual. Additionally, the framework reveals connections with approaches by Zeilberger [60], Munch-Maccagnoni [6], and Curien and Munch-Maccagnoni [33], for using polarization and focalization to provide deterministic strategies for classical computation with structures and pattern matching. This work will be presented at ESOP 2014 [15].

Luke Maurer and Zena Ariola in collaboration with Daniele Varacca studied the connections between π-calculus encodings of the λ-calculus and similar continuation-passing style (CPS) transformations, extending the connections for call-by-value and call-by-name encodings to include the call-by-need π-calculus encoding as well. This development revealed a better understanding of the computational effect needed in the λ-calculus to model call-by-need evaluation, which better reflects the way that memoization for call-by-need is implemented. The work is going to be submitted to RTA-TLCA.

Constructive interpretation of an involutive negation.

Guillaume Munch-Maccagnoni developped a syntax of delimited control operators that exposes a formulae-as-types correspondence between an involutive negation in classical natural deduction, and the idea that captured contexts, unlike continuations, can be inspected. This decomposes technical artefacts found in call-by-name classical realisability, and simplifies witness extraction from proofs of Σ formulae. This work has been submitted and appears in his PhD thesis [11].

5.1.2. Dependent monads

Guillaume Claret and Yann Régis-Gianas are developping a monadic translation from functional programs with effects to Coq that uses a dependent monad. The aim of this work is to allow to reason about effectful programs directly in Coq.

5.1.3. Linear dependent types

Pierre-Marie Pédrot developped a dependent version of the Dialectica translation, that gives interesting insights into the possibility to design linear dependent types. Indeed, Dialectica can be decomposed as a translation acting on linear types instead of intuitionistic ones.

5.1.4. Delimited continuations, polarity and computational effects

Guillaume Munch-Maccagnoni's polarised decomposition of delimited control calculi appeared in his PhD thesis [11].

5.1.5. Reverse mathematics

Ludovic Patey studied with Laurent Bienvenu and Paul Shafer the deep connections between algorithmic randomness and reverse mathematics by defining formally the ability of computing a solution to a problem by probabilistic means within the framework of reverse mathematics, the No Randomized Algorithm property (NRA). They provided a classification of the whole revese mathematics zoo created by Damir Dzhafarov in terms of having the NRA property or not, answering to some open separation questions.

Ludovic Patey stated two dichotomy theorems about satisfiability problems within reverse mathematics and proved them using clone theory. The corresponding paper is submitted to Computability in Europe 2014. He studied also ramseyan theorems related to the Rainbow Ramsey theorem and provided characterizations in terms of diagonally non-computable functions, algorithmic randomness, and related it to the Erdös Moser theorem and Thin set theorem.

5.1.6. Gödel's functional interpretation

Pierre-Marie Pédrot showed that the Dialectica translation could be explained in terms of the Krivine abstract machine, in a way similar to the usual presentation of classical realizability. This opens the door to a better understanding of related translations, as well as adding semi-classical effect into PTS.

5.1.7. Logical foundations of call-by-need evaluation

Alexis Saurin and Pierre-Marie Pédrot developed a structured reconstruction of call-by-need based on linear head reduction which arose in the context of linear logic. This opens new directions both to extend call-by-need to control and to apply linear logic proof-theory (and particularly proof-nets) to call-by-need evaluation.

5.1.8. Streams and classical logic

Alexis Saurin and Fanny He have been working on transfinite term rewriting in order to model stream calculi and their connections with lambda-calculi for classical logic.

Jaime Gaspar identified the eight simplest variants (some already known) of the Kuroda negative translation that translate classical logic into minimal logic.

5.2. 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.

5.2.1. Substitutions and isomorphisms

Pierre-Louis Curien completed his joint work with Richard Garner and Martin Hofmann on relating syntax unstrictification through coercions with model strictification (cf. πr^2 report 2012), adding a careful treatment of identity types. The corresponding paper was accepted for publication in the TCS special issue for Glynn Winskel's anniversary.

5.2.2. Homotopy type theory

Hugo Herbelin, Matthieu Sozeau and Pierre-Louis Curien participated to the univalent foundations program. A collaborative book [18] on the results of this program has been published.

5.2.3. Models of type theory

Simplicial sets and their extensions as Kan complexes can serve as models of homotopy type theory. Hugo Herbelin developed a concrete type-theoretic formalisation of semi-simplicial sets following ideas from Steve Awodey, Peter LeFanu Lumsdaine and other researchers both at Carnegie-Mellon University and at the Institute of Advanced Study. This has been accepted for publication in a special issue of MSCS on homotopy type theory [22].

The technique he used generalises to provide type-theoretic constructions for arbitrary presheaves on Reedy categories, thus including simplicial sets. In particular, this provides with a formulation of simplicial sets where degeneracies are decidable, which is not the case with the definition as a presheaf.

Hugo Herbelin also investigated hybrid constructive definitions of simplicial sets where face maps are axiomatised but degeneracies are built. Again, this provides with a formulation of simplicial sets where it is decidable whether a given simplex is degenerate or not.

5.2.4. Internalizing the setoid model of type theory

As an example use of the new polymorphic universe extension of Coq, Matthieu Sozeau developed together with Nicolas Tabareau (Inria Ascola team, École des mines Nantes) a complete groupoid model of type theory, following the seminal work of Hofmann and Streicher. A preliminary paper presenting a partial generalization of this model to 2-goupoids was written and will be resubmitted [23].

A completed version of this model has since been formalized and will be submitted shortly. This model showcases the use of the polymorphic universes: in the course of its formalization we uncovered hidden assumptions in the interpretation of substitution and sigma types in the original presentation thanks to the universe system.

5.2.5. Proof irrelevance, eta-rules

Matthieu Sozeau finished his implementation of a proof-irrelevant system but did not publish it. Indeed, the homotopy type theory interpretation suggests new ways to introduce proof-irrelevance using bracket types that seem to significantly depart from the syntactic treatment developped by Werner and himself. An investigation of the relationship between the presentation of the calculus of inductive constructions given by Hugo Herbelin and Arnaud Spiwack in [44] which includes the bracket construction and the aforementioned syntactic version will be part of a master's internship supervised by Matthieu Sozeau in 2014.

5.3. Homotopy of rewriting systems

Participants: Cyrille Chenavier, Pierre-Louis Curien, Yves Guiraud, Maxime Lucas, Philippe Malbos.

5.3.1. The homotopical completion-reduction procedure

In [39], Stéphane Gaussent (Institut Camille Jordan), Yves Guiraud and Philippe Malbos have introduced the homotopical completion-reduction procedure as a higher-dimensional rewriting method to compute coherent presentations of monoids. The results of this procedure on Artin monoids of spherical type have been implemented by Yves Guiraud in a Python library, available on his webpage. The procedure is currently improved towards the explicit computation of full polygraphic resolutions of Artin monoids to provide a purely algebraic and constructive account of well-known geometric objects, such as Caylay graphs and Salvetti complexes.

In [16], Yves Guiraud, Philippe Malbos and Samuel Mimram (CEA Saclay) have further investigated the homotopical completion-reduction procedure, extended with the adjunction/elimination of redundant generators, with successful application to two new classes of monoids: the plactic and the Chinese monoids. This work has been implemented by Samuel Mimram and Yves Guiraud into a prototype, that can be tested at http://www.pps.univ-paris-diderot.fr/~smimram/rewr, and has been presented to RTA 2013 by Philippe Malbos, where it has received the best paper award.

5.3.2. New methods for the computation of coherent presentations

During his M2 internship, Maxime Lucas, supervised by Yves Guiraud, has improved the rewriting method used in [43] for the computation of homotopy bases of monoids and categories. This allows a more effective computation in several cases, based on the notion of Anick chain [25] instead of the broader notion of critical branching. Maxime Lucas has now started a PhD thesis, supervised by Yves Guiraud and Pierre-Louis Curien, and currently investigates the use of Garside-like structures [35] to further improve the computation of coherent presentations for higher-dimensional categories.

5.3.3. Higher-dimensional linear rewriting

Cyrille Chenavier, Pierre-Louis Curien, Yves Guiraud and Philippe Malbos investigate with Eric Hoffbeck (LAGA, Université Paris 13) and Samuel Mimram (CEA Saclay) the links between set-theoretic rewriting theory and the computational methods known in symbolic algebra, such as Gröbner bases [28]. This interaction is supported by the Focal project of the IDEX Sorbonne Paris Cité. Yves Guiraud, Eric Hoffbeck and Philippe Malbos are currently working on an improvement, based on the homotopical completion-reduction procedure, of the methods known in algebra to compute homological invariants of algebras and operads. Cyrille Chenavier has started a PhD thesis, supervised by Yves Guiraud and Philippe Malbos, to use Berger's theory of reduction operators [27] to design new methods for the study of rewriting systems.

5.3.4. Homotopical and homological finiteness conditions

Yves Guiraud and Philippe Malbos have written a comprehensive introduction [21] on the links between higher-dimensional rewriting, the homotopical finiteness condition "finite derivation type" and the homological finiteness condition "FP_3", from the point of view of higher categories and polygraphs. The purpose of this work is to provide an introduction to the field, formulated in a contemporary language, and with new, more formal proofs of classical results.

In [19], Yves Guiraud and Philippe Malbos have introduced a notion of identities among relations for higher categories presented by polygraphs. This notion is well-known in combinatorial group theory, where it is linked to the explicit computation of homological invariants and of formal representations of groups as crossed complexes. The main result of [19] is a procedure based on higher rewriting to compute generators of the identities among relations. They have related the facts that the natural system of identities among relations is finitely generated and that the higher category has finite derivation type (a homotopical finiteness condition introduced in [43] for higher categories after Squier's work for monoids [57]).

5.4. Coq as a functional programming language

Participants: Pierre Boutillier, Guillaume Claret, Lourdes Del Carmen González Huesca, Hugo Herbelin, Pierre Letouzey, Matthias Puech, Yann Régis-Gianas, Matthieu Sozeau.

5.4.1. Type classes and libraries

Type Classes are heavily used in the HoTT/Coq library (http://github.com/HoTT/coq) developed by the Univalent Foundations program at the IAS, to which Matthieu Sozeau participated.

5.4.2. Dependent pattern-matching

How to encode structurally dependent pattern matching into case analysis by hand has been described by Jean François Monin in [52]. Pierre Boutillier, with the help of Thomas Braibant (GALLIUM team), has mechanized this process and exhibited a missing part to make it scale. These are the main results presented in Pierre Boutillier's forthcoming thesis.

5.4.3. Incrementality in proof languages

Lourdes González and Yann Régis-Gianas studied incremental computing and self-adjusting computation [24] as a starting point to develop an applicative notion of change over data structures, to be applied to lambda-terms. They formalized in Coq a notion of derivative of an inductive function to define how to compute a new result from an input that has changed, this is done by using the derivative of the function and the difference on inputs and old outputs. They are working out a technique that allows a specification of functions using derivatives and old inputs and outputs including a cost analysis of the benefits of reusing previous computations.

5.4.4. Lightweight proof-by-reflection

In collaboration with Beta Ziliani (MPI), In the context of the ANR project Paral-ITP, Lourdes del Carmen González 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. This work has been presented at ITP 2013 ([14]).

POEMS Project-Team

6. New Results

6.1. Time domain wave propagation problems

6.1.1. Numerical methods in electromagnetism

Participant: Gary Cohen.

In the framework of contract GREAT, we implemented and compared two discontinuous Galerkin methods to solve Maxwell's equations for time dependent problems, the first using tetrahedral meshes (first used by Hesthaven), the second using hexahedral meshes with mass lumping. This comparison showed the undeniable superiority of the second method, 4-7 times faster (for orders from 2 to 4) for the same accuracy.

The ultimate goal of this program was the hybridization of those two types of meshes because the construction of purely hexahedral mesh for complex geometries is often difficult or almost impossible. A first approach was studied in the thesis Morgane Bergot, where the transition between the two grids was performed by the use of pyramids. The implementation of such elements is difficult and costly, we were interested in a transition mortars elements capable of hybridizing directly tetrahedra with flat faces and hexahedral with non-planar faces. This approach is promising and should lead to a rapid and efficient method. A theoretical study of the error and stability is conducted in collaboration with Eric Chung of CUHK (Chinese University of Hong Kong).

Moreover, always with E. Chung, we became interested in the construction of a discontinuous Galerkin method on hexahedral meshes offset for solving Maxwell's equations. This approach has two advantages : firstly, the shift naturally removes the spurious waves which appear with other approaches (which usually requires the introduction of a dissipative term to remove them). On the other hand, a phenomenon of super-convergence appears which should lead to a substantial time saving. A first study of the dispersion of this method led to a publication.

6.1.2. Solving the Homogeneous Isotropic Linear Elastodynamics Equations Using Potentials and Finite Elements.

Participants: Aliénor Burel, Patrick Joly.

The aim of this subject, investigated in collaboration with Marc Duruflé (Inria Bordeaux) and Sébastien Imperiale (Inria Saclay), is to use the classical theoretical decomposition of the elastodynamic displacement into two potentials referring to the pressure wave and the shear wave, and use it in a numerical framework. During the past two years, a method has been proposed for solving the Dirichlet problem (clamped boundary), successfully analysed and implemented, and for the free boundary conditions, we proposed an original method considering these boundary conditions as a perturbation of the Dirichlet conditions. This approach performs successfully in the time-harmonic regime but appears to give rise to severe instabilities in the time-dependent case after space and time discretization. Our investigations seem to prove that this instability is already present in the semi-discrete problem in space, but we are still looking for an explanation of this phenomenon.

6.1.3. Limiting amplitude principle in a two-layered medium composed of a dielectric and a metamaterial

Participants: Maxence Cassier, Christophe Hazard, Patrick Joly, Valentin Vinoles.

We are investigating this problem from both theoretical and numerical points of view. This is also the object of a collaboration with B. Gralak from the Institut Fresnel in Marseille.

This work is the time-domain counterpart of the research done at Poems about frequency domain analysis of metamaterials in electromagnetism, in the framework of the ANR Project METAMATH. One fundamental question is the link between the evolution / time harmonic problems via the limiting amplitude principle, in particular in the cases where the time harmonic problem fails to be well posed. This occurs, at certain frequencies, when one considers a transmission problem between a standard dielectric material and a dispersive material obeying for instance to the Drude model (other models as Lorentz materials or their generalization also give rise to the same results). Indeed, for well-chosen coefficients (which we refer as the critical case), there exist critical frequencies (only one frequency ω_c for the Drude model) for which the metamaterial behaves as a material whose equivalent electric permittivity and magnetic permeability are negative and precisely opposite to the ones of the dielectric medium : in such a situation, in the case of a plane interface, it is known that the time harmonic transmission problem is strongly ill-posed.

We have considered the evolution problem in a two-layered medium, when we consider a source term $f(x)\, e^{i\,\omega t}$ with frequency $\omega > 0$. In the non critical case, the limiting amplitude principle holds : for large times, the solution $u(x,t)$ of the evolution problem "converges" to a time harmonic solution of the form $u^\infty(x)\, e^{i\,\omega t}$. In the critical case, the limiting amplitude principle no longer holds. If $\omega \neq \omega_c$, the solution of the evolution problem behaves when $t \to +\infty$ to a "double frequency" solution of the form

$$u^\infty(x)\, e^{i\,\omega t} + u^\infty_{c,+}(x)\, e^{i\,\omega_c t} + u^\infty_{c,-}(x)\, e^{-i\,\omega_c t}.$$

If $\omega = \omega_c$, the solution blows up linearly at infinity :

$$u(x,t) \sim t\, u^\infty_c(x)\, e^{i\,\omega_c t} \quad (t \to +\infty)$$

where the function $u^\infty_c(x)$ is "concentrated" near the interface : this can be interpreted as an "interface resonance" phenomenon. We have performed various numerical experiments (using in particular the stabilized PMLs evoked in section 6.3.2) that illustrate this resonance phenomenon (cf figure 1).

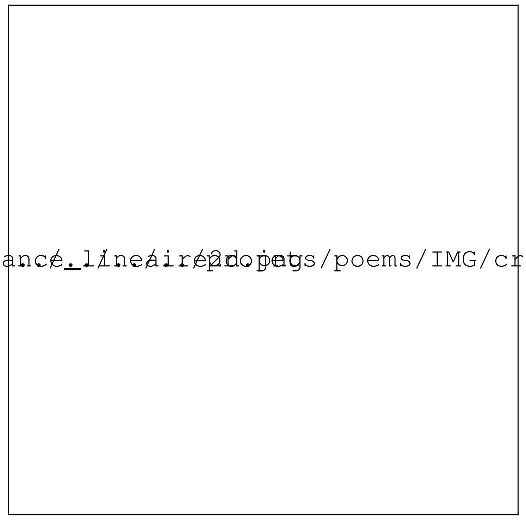


Figure 1. Left : the interface wave. Right : solution at one point as a function of time.

From the mathematical point of view, the method we have used consists in rewriting the original problem as an abstract Schrödinger equation

$$i \frac{du}{dt} + Au = F\, e^{i\,\omega t}$$

where A is a self-adjoint operator in an appropriate Hilbert space H. The key of the analysis is the spectral theory of the operator A. This permits a quasi-explicit representation of the solution via the (generalized) diagonalization of A. This is achieved by combining a partial Fourier transform along the interface with Sturm-Liouville type techniques in the orthogonal direction. In the critical case, the resonance phenomenon appears to be linked to the fact that A admits a (single) eigenvalue of infinite multiplicity.

6.1.4. Finite differences method for nonlinear acoustic waves with fractional derivatives

Participant: Jean-François Mercier.

This subject is developed in collaboration with Bruno Lombard from LMA.

We develop a numerical method to study the wave propagation in a 1-D guide with an array of Helmholtz resonators, considering large amplitude waves and viscous boundary layers. The model consists in two coupled equations: a nonlinear PDE for the velocity in the tube (Burgers like equation) and a linear ODE describing the pressure oscillations in the Helmholtz resonators. The dissipative and dispersive effects in the tube and in the necks of the resonators are modelled by fractional derivatives expressed as convolution products with singular kernels. Based on a diffusive representation, the convolution kernels of the fractional derivatives are replaced by a finite number of memory variables that satisfy local ordinary differential equations. The procedure to compute weights and nodes of the diffusive representation of fractional derivatives is optimized. Moreover an adequate coupling between the PDE and the ODE is introduced to be sure that the discrete energy is decreasing. A splitting strategy is then applied to the evolution equations to obtain a stable scheme under the optimal CFL condition: the propagative part is solved by a standard TVD scheme for hyperbolic equations, whereas the diffusive part is solved exactly. This approach is validated by comparisons with exact solutions. The properties of the full nonlinear solutions are investigated numerically. In particular, the existence of acoustic solitary waves, due to the competition between dispersion and nonlinear effects, is confirmed.

6.2. Time-harmonic diffraction problems

6.2.1. Numerical computation of variational integral equation methods

Participants: Marc Lenoir, Nicolas Salles.

The discretization of 3-D scattering problems by variational boundary element methods leads to the evaluation of such elementary integrals as

$$\int_{S\times T} G\left(x,y\right)\, v\left(x\right) w\left(y\right)\, dx\, dy \text{ and } \int_{S\times T} \frac{\partial}{\partial n_y} G\left(x,y\right)\, v\left(x\right) w\left(y\right)\, dx\, dy \tag{24}$$

where v and w are polynomial basis functions, G is the Green kernel and S and T two planar polygons from the discretization of the boundary. Due to the singularity of the kernel, the numerical evaluation of these integrals may lead to inaccurate results when S and T are close to each other. We split G and its gradient into a regular part which involves classical numerical techniques and a singular part subject to our method. This new method consists in integrating exactly integrals such as

$$I = \int_{S\times T} v(x) \frac{1}{\|x-y\|} w(y)\, dx\, dy \text{ and } J_\zeta = \int_{S\times T} \frac{x-y}{\|x-y\|^{1+\zeta}}\, dx\, dy,\ \zeta \in \{0,2\}. \tag{25}$$

or numerically integrals such as:

$$\mathcal{L} = \int_{S\times T} v(x) \frac{e^{ik\|x-y\|}}{\|x-y\|} w(y)\, dx\, dy \tag{26}$$

where v and w are basis functions of order 0 or 1. The general approach relies on two steps.

Basic formulas : let $f\,(x, d)\ \ : \Omega \subset \mathbb{R}^n \times \mathbb{R} \longrightarrow \mathbb{R}$ a positively homogeneous function of degree q. By Euler's formula and Green's theorem we have the function $I(d)$ satisfies :

$$(q+n)\, I\,(d) = dI'\,(d) + \int_{\partial\Omega} (\overrightarrow{z}\,|\overrightarrow{\nu}\,)\, f\,(z, d)\, d\gamma_z, \quad \text{with } I\,(d) = \int_{\Omega} f\,(z, d)\, dz \tag{27}$$

where $\overrightarrow{\nu}$ is the exterior normal to Ω. Provided $d^{-(q+n)} \int_{\Omega} f\,(z, d)\, dz \to 0$ as $d \to +\infty$ one obtains

$$I\,(d) = d^{q+n} \int_{\partial\Omega} (\overrightarrow{z}\,|\overrightarrow{\nu}\,) \int_{d}^{+\infty} \frac{f\,(z, t)}{t^{q+n+1}} dt d\gamma_z. \tag{28}$$

When $f\,(z, d)$ does not depend on d and $q+n \neq 0$ then

$$I = \frac{1}{q+n} \int_{\partial\Omega} (\overrightarrow{z}\,|\overrightarrow{\nu}\,)\, f\,(z)\, d\gamma_z. \tag{29}$$

As long as the inner integral in (5) can be explicitly evaluated, both formulas reduce an n-dimensional integral to an $(n-1)$ one. When Ω is an n-dimensional polyhedron (such as $S \times T$ with $n = 4$), $(\overrightarrow{z}\,|\overrightarrow{\nu}\,)$ is constant on each $(n-1)$-face of Ω, a simplification of crucial importance in the sequel.

The reduction process : we have obtained formulas for three types of geometrical configurations: S and T are (i) coplanar, (ii) in secant planes and (iii) in parallel planes. All these cases are treated using formulas (6) or (5) or both, depending on the relative positions of S and T. As an example, we present the simple but significant result for the self-influence coefficient ($S = T$). Let A_i be a vertex of the triangle, α_i the opposite side, B_i the orthogonal projection of A_i on α_i and $\gamma_i = \|A_i B_i\|$. After 3 successive reductions using formula (6), one obtains

$$I = \int_{S\times S} \frac{1}{\|x-y\|} dx dy = \frac{2\,|S|}{3} \sum_{i=1,3} \gamma_i R\,(A_i, \alpha_i)\,, \tag{30}$$

where $R(A_i, \alpha_i)$ is given analytically by ((i, jk) being a circular permutation of $(1, 2, 3)$)

$$R\,(A_i, \alpha_i) = \int_{\alpha_i} \frac{1}{\|A_i - y\|} dy = \arg\sinh \frac{\|B_i A_k\|}{\gamma_i} - \arg\sinh \frac{\|B_i A_j\|}{\gamma_i}. \tag{31}$$

Results for the 3-D Helmholtz equation with piecewise constant density have been obtained for all pairs of panels. Integral $\mathcal{L}$ (see formula (3)) can be reduced to a linear combination of 1-D or 2-D integrals when triangles have at least one common vertex; the resulting integrals have to be evaluated numerically but the final integrands are simple and regular on the domain of integration. For example, when $T = S$ and with piecewise constant basis functions, one has:

$$\mathcal{L} = \int_{S\times S} \frac{e^{ik\|x-y\|}}{\|x-y\|}\, dx\, dy = 4|S| \sum_{i=1}^{3} \gamma_i \int_{\alpha_i} f\,(\|A_i - y\|)\, ds_y \tag{32}$$

where $f(r) = i\dfrac{e^{ikr} - 1 - ikr + k^2 r^2/2}{k^3 r^4}$.

The extension to linear basis functions is in progress. Our method works also for 3-D Maxwell's equations with linear edge basis functions (for MFIE and EFIE). Despite some (possibly) lengthy calculations, the principle is rather straightforward and the method is quite flexible, leading to the reduction of 4-D integrals to a linear combination 1-D regular integrals which can be numerically or even explicitly evaluated. It is possible to use our method for Collocation technique, 2-D BEM and volume integral equations. A high degree of accuracy can be obtained, even in the case of nearly singular integrals. We will present the method and some results for 3-D Helmholtz equation.

6.2.2. Integral equations for modelling eddy current non destructive testing experiments

Participants: Marc Bonnet, Audrey Vigneron.

This work in collaboration with E. Demaldent (CEA LIST) is concerned with developing boundary element solvers for modelling eddy current non destructive testing experiments, taking into account the probe, the probed part and the surrounding air. Attention is focused in implementing Galerkin-type formulations, overcoming ill-conditioning arising in configurations involving high contrasts, and fast solvers. Among several possible integral formulations based on either Maxwell's equations or the eddy-current model, a weighted coupled formulation using a loop-tree decomposition of the trial and test spaces was found to perform adequately over the whole range of values of physical parameters typical of eddy-current NDT experiments.

6.2.3. Elastodynamic fast multipole method for semi-infinite domains.

Participants: Marc Bonnet, Stéphanie Chaillat.

The use of the elastodynamic half-space Green's tensor in the FM-BEM is a very promising avenue for enhancing the computational performances of 3D BEM applied to analyses arising from e.g. soil-structure interaction or seismology. This work is concerned with a formulation and computation algorithm for the elastodynamic Green's tensor for the traction-free half-space allowing its use within a Fast Multipole Boundary Element Method (FM-BEM). Due to the implicit satisfaction of the traction-free boundary condition achieved by the Green's tensor, discretization of (parts of) the free surface is no longer required. Unlike the full-space fundamental solution, the elastodynamic half-space Green's tensor cannot be expressed in terms of usual kernels such as $e^{ikr/r}$ or $1/r$. Its multipole expansion thus cannot be deduced from known expansions, and is formulated in this work using a spatial two-dimensional Fourier transform approach. The latter achieves the separation of variables which is required by the FMM. To address the critical need of an efficient quadrature for the 2D Fourier integral, whose singular and oscillatory character precludes using usual (e.g. Gaussian) rules, generalized Gaussian quadrature rules have been used instead. The latter were generated by tailoring for the present needs the methodology of Rokhlin's group. Extensive numerical tests have been conducted to demonstrate the accuracy and numerical efficiency of the proposed FMM. In particular, a complexity significantly lower than that of the non-multipole version was shown to be achieved. A full FM-BEM based on the proposed acceleration method for the half-space Green's tensor is currently under way. This treatment of the Green's tensor will be extended to other cases, e.g. layered semi-infinite media.

6.2.4. Domain decomposition methods for time harmonic wave propagation

Participants: Patrick Joly, Mathieu Lecouvez.

This work is motivated by a collaboration with the CEA-CESTA (B. Stupfel) through the PhD thesis of M. Lecouvez and is the object of a collaboration with F. Collino, co-advisor of the thesis with P. Joly.

We have considered first the case of the scalar Helmholtz equation for which we have developed a non overlapping iterative domain decomposition method based on the use of Robin type transmission conditions, in the spirit of previous works in the 90's by Collino, Desprès, and Joly.

The novelty of our approach consists in using new transmission conditions using some specific impedance operators in order to improve the convergence properties of the method (with respect to more standard Robin conditions). Provided that such operators have appropriate functional analytic properties, the theory shows that one achieves geometric convergence (in opposition the the slow algebraic convergence obtained with standard methods). These properties prevent the use of local impedance operator, a choice that was commonly done for the quest of optimized transmission conditions (following for instance the works of Gander, Japhet, Nataf). We propose a solution that uses nonlocal integral operators using appropriate Riesz potentials, the important feature of which being their singularity at the origin. To overcome the disadvantage of dealing with completely nonlocal operators, we suggest to work with truncated kernels, involving adequate smooth cut-off function. The results we have obtained are

- A complete theoretical justification of the exponential convergence of the algorithm in the 2D case for smooth enough interfaces. The extension to 3D is in progress : the case of a spherical interface is in particular completely understood.
- An heuristic analysis of the influence of the truncation procedure (several choices are possible) on the convergence of the method, together with a (semi-analytical) search for optimal values of the parameters involved in the method to improve the convergence rate.
- The implementation of the method in 2D and an intensive campaign of numerical validation of the method that appear to provide very good performance and seem to indicate that the method is quite robust with respect to increasing frequency (which remains to be proven). Let us however mention that, not so important but unexpected phenomena, due to space discretization, have been observed and remain to be explained. The implementation in 3D, in cooperation with M. Duruflé, is in progress.

The relevant application at CESTA being electromagnetism, the extension of the method to 3D Maxwell's equations, which proposes new non trivial difficulties, has been initiated.

As the development and the theoretical understanding of these new domain decomposition methods clearly exceed the content of one single thesis, we have proposed an ANR project on this topic, in collaboration with X. Claeys (Paris VI).

6.2.5. Time harmonic aeroacoustics

Participant: Jean-François Mercier.

This subject is treated in collaboration with Florence Millot (Cerfacs). We are still working on the numerical simulation of the acoustic radiation and scattering in presence of a mean flow. Up to now we have considered Galbrun's equation, but for 3D configurations it requires to introduce many unknowns. Therefore we focus now on the alternative model of Goldstein's equations. When the fluid flow and the source are potential, the acoustic perturbations are also potential and the velocity potential φ satisfy a simple scalar model. For a general flow, this model is slightly modified and is called Goldstein's equations. A new vectorial unknown ξ is introduced, satisfying a transport equation coupled to the velocity potential. φ satisfies the same modified Helmholtz's equation than in the potential flow case, in which ξ plays the role of a source term. The advantage of Goldstein's formulation compared to Galbrun's model is that the vectorial unknown vanishes in the areas where the flow is potential.

For a general flow ξ can be expressed versus φ as a convolution formula along the flow streamlines. The situation is much simpler for slow flows since the convolution formula can be simplified and the link between ξ and φ becomes explicit. Then Goldstein's equations can be solved by using continuous finite element (discontinuous elements must be used in the general case). We have proved theoretically that when replacing the general convolution formula by the "slow flow" approximation, the error on the velocity potential is small, bounded by the square of the flow velocity. This has been done for a simpler case, a shear flow, for which the streamlines are just parallel lines. Numerical tests have confirmed the square law for the error.

6.2.6. Mathematical and numerical analysis of metamaterials

Participants: Patrick Joly, Anne-Sophie Bonnet-Ben Dhia, Patrick Ciarlet, Sonia Fliss, Camille Carvalho, Valentin Vinoles, Christian Stohrer.

Metamaterials are artificial composite materials having the extraordinary electromagnetic property of negative permittivity and permeability at some frequencies . Both of sign-changing coefficients and high contrast homogenization raise new mathematical and numerical challenges. The ANR METAMATH is devoted to the study of those problems. We perform analysis both in time domain (see sections 6.1.3 and 6.3.2) and harmonic domain.

6.2.6.1. Time-harmonic transmission problems involving metamaterials

A special interest is devoted to the transmission of an electromagnetic wave between two media with opposite sign dielectric and/or magnetic constants. As a matter of fact, applied mathematicians have to address challenging issues, both from the theoretical and the discretization points of view. In particular, it can happen that the problem is not well-posed in the classical frameworks (H^1 for the scalar case, $H(curl)$ for the vector case). During 2013, we addressed the issues below.

The numerical analysis of the well-posed scalar eigenproblem discretized with a classical, H^1 conforming, finite element method, for arbitrarily shaped interfaces can be carried out with the help of T-coercivity. This work complements the paper Chesnel-Ciarlet, published in Numerische Mathematik, which handled simpler interface configurations (see also §6.2.6.2).

As a second topic, we investigated the case of a scattering problem with a 2D corner interface which can be ill-posed (in the classical H^1 framework). When this is the case, the part of the solution which does not belong to H^1 can be described as a wave which takes an infinite time to reach the corner: this "black-hole" phenomenon is observed in other situations (elastic wedges for example). We have proposed a numerical approach to approximate the solution which consists in adding some Perfectly Matched Layers in the neighbourhood of the corner. As an alternate choice, a T-coercivity approach is also being currently developed to solve the discrete problem.

Last, we studied the transmission problem in a purely 3D electromagnetic setting from a theoretical point of view. We proved that the Maxwell problem is well-posed if and only if the two associated scalar problems (with Dirichlet and Neumann boundary conditions) are well-posed. Numerical analysis of the discretized models (edge elements) is under way.

L. Chesnel left our project in March 2013 after he completed his PhD thesis on these topics. He is currently a post-doc fellow at Aalto University (Finland).

6.2.6.2. Modeling of plasmonic devices

Plasmonic surface waves occur at the interface between the vacuum (or a dielectric) and a metal, at optical frequencies, when the dielectric permittivity ε of the metal has a small imaginary part and a large negative real part. Neglecting the dissipation effects, we have to study electromagnetic problems with a sign-changing ε. An in-depth analysis has been done by Lucas Chesnel during his PhD. In the context of the PhD of Camille Carvalho, we extended the results obtained previously by Lucas Chesnel to more realistic configurations. First, we studied the diffraction of a transversely polarized plane wave by a cylindrical metallic inclusion, when the section of the inclusion presents edges (cf. §6.2.6.1). Then, we considered a related spectral problem in view of studying plasmonic guided waves. The spectral theory is far from obvious. In particular, we have to introduce a non-selfadjoint formulation which provides physical real eigenvalues and complex spurious ones. For both the diffraction problem and the spectral problem, a MATLAB code has been developed, where Perfectly Matched Layers are introduced at the corners to take into account the presence of black-hole waves seemingly absorbed by the corners. The convergence of the finite element discretization (including convergence of the eigenvalues) has been proved (see §6.2.6.1).

6.2.6.3. Study of metamaterials via numerical homogenization

Recently, we have started to study the numerical approximation of the full models, using the HMM (Heterogeneous Multiscale Method). Recall that the full model is obtained via periodization of a local model that includes slow and fast variations. With this HMM approach, computations are carried out on a global mesh, whereas the action of the test-functions is computed at a local level to take into account the fast variations. As a first step, we have begun by the application of HMM for the time-harmonic scalar problem. The case of uniformly bounded coefficients has been addressed. The more general case of non-uniformly bounded coefficients, also called the high-contrast case, is now under scrutiny. It is hoped that one can recover some extra-ordinary properties of the metamaterials with this latter case.
C. Stohrer arrived as a post-doc fellow this fall.

6.3. Absorbing boundary conditions and absorbing layers

6.3.1. New transparent boundary conditions for time harmonic acoustic and elastic problems in anisotropic media

Participants: Anne-Sophie Bonnet-Ben Dhia, Antoine Tonnoir, Sonia Fliss.

This topic is developed in collaboration with Vahan Baronian (CEA). Non destructive testing (NDT) is a common method to check the quality of structures and is widely used in industrial applications. Typically, in aircraft design, it is required to control structures like plates. Efficient and accurate numerical methods are required to simulate NDT experiments.

In our case, we want to study the diffraction of a time harmonic wave by a bounded defect in an infinite anisotropic elastic plate. The difficulty is to find a way to restrict the finite element computation to a small box containing the defect. Indeed classical methods such as the perfectly matched layers fail when the medium is anisotropic.

Up to now we considered the simpler case of an infinite dissipative 2D medium.

Our idea, inspired by the work of Sonia Fliss and Patrick Joly for periodic media, is to consider five domains recovering the whole plane:

- a square that surrounds the defect in which we have a finite element representation of the solution,
- and four half-spaces parallel to the four edges of the square, in which we can give an analytical representation of the solution thanks to the Fourier transform.

The different unknowns are coupled by well-chosen transmission relations which ensure the compatibility between the five representations.

The method has been validated successfully in the case of anisotropic acoustic media and the implementation for the case of elasticity is in progress. The mathematical properties of the formulation and the efficiency of the method strongly depend on the presence or not of overlaps between the finite element box and the four half-planes. The formulation with overlaps has good Fredholm properties but the well-posedness for all frequencies is proved only for the formulation without overlaps.

6.3.2. Perfectly Matched Layers in negative index metamaterials

Participants: Patrick Joly, Eliane Bécache, Valentin Vinoles.

The simulation of waves in unbounded domains requires methods to artificially truncate the computational domain. One of the most popular ones to do so is the Perfectly Matched Layers (PMLs) which are effective and stable for non dispersive isotropic media. For non dispersive anisotropic media, we established a necessary stability condition in 2004 : PMLs are unstable in presence of so called-backward waves.

We are interested here in dispersive media and more specifically in Negative Index Metamaterials (NIMs), also called left-handed media. Those media have negative permittivity and permeability at some frequencies due to microscopic resonating structures. Since the 1990s, NIMs are the subject of active researches due to their promising applications : superlens, cloaking, improved antenna, etc.

In a first step, we consider a simple model of NIMs : the Drude model. For this model, a plane wave analysis shows the simultaneous presence of both forward and backward waves and numerical simulations confirm the instability of standard PMLs (cf figure 2) that result from complex changes of variable leading to the following modification of the spatial derivatives

$$\partial_x \longrightarrow \left(1+\frac{\sigma_x(x)}{i\omega}\right)^{-1}\partial_x \quad \text{and} \quad \partial_y \longrightarrow \left(1+\frac{\sigma_y(y)}{i\omega}\right)^{-1}\partial_y$$

where $\sigma_x(x) > 0$ and $\sigma_y(y) > 0$ are the damping terms. Inspired by works of the physics community, we propose more general changes of variable

$$\partial_x \longrightarrow \left(1+\frac{\sigma_x(x)}{i\omega\psi(\omega)}\right)^{-1}\partial_x \quad \text{and} \quad \partial_y \longrightarrow \left(1+\frac{\sigma_y(y)}{i\omega\psi(\omega)}\right)^{-1}\partial_y$$

where $\psi(\omega)$ is a function to be chosen judiciously. We have generalised the previous necessary stability condition for those new PMLs, called Stabilized Perfectly Matched Layers, for dispersive media. This analysis allows us to understand the instabilities observed for standard PMLs in NIMs and to propose a choice of functions $\psi(\omega)$ which take into account the backward waves and stabilize the PMLs as confirmed by numerical simulations (cf figure 2).

../../../../projets/poems/IMG/pml_instables.png ../../../../projets/poems/IMG/pml_stables.png

Figure 2. Left : the standard PMLs are unstable. Right : the Stabilized PMLs are stable.

6.3.3. Perfectly Matched Layers in plasmas

Participants: Patrick Joly, Eliane Bécache, Valentin Vinoles.

This work was done during the internship of Guillaume Chicaud in the framework of the ANR CHROME which concerns the study of electromagnetic wave propagation in plasmas. Our aim is to develop efficient and robust codes to simulate wave propagation in unbounded plasmas models. The simulation of waves in plasmas requires technics to bound the computational domain. As plasmas are dispersive media where backward waves may occur, the difficulties to construct stable PMLs are analogous to the ones encountered for Negative Index Metamaterials (cf 6.3.2). This work is a preliminary study of this topics, in a simplified model, the case of a 2D anisotropic uniaxial plasma. It consists first in analyzing the presence of backward waves with a plane wave analysis. The second step was to implement the equations using standard PMLs and to confirm the expected instabilities. Finally, we proposed stabilized PMLs (SPMLs), inspired by the work done in metamaterials (see section 6.3.2).

The continuation of this project will constitute the subject of the post-doc of Maryna Kachanovska.

6.4. Waveguides, resonances, and scattering theory

6.4.1. An improved modal method in non uniform acoustic waveguides

Participant: Jean-François Mercier.

This topic is developed in collaboration with Agnès Maurel (Langevin Institute ESPCI).
We develop modal methods to study the scattering of an acoustic wave in a non uniform waveguide. Usual modal approaches are efficient only when a rather large number of evanescent modes are taken into account. An improved representation has been proposed in which an additional transverse mode and an additional unknown modal component are introduced. This so called boundary mode helps to better satisfy the Neumann boundary conditions at the varying walls. A system of coupled ordinary differential equations is obtained and is found to remain coupled in the straight part of the waveguide which implies that the classical radiation condition cannot be applied directly at the inlet/outlet of the scattering region.

We revisit the coupled mode equations in order to derive an improved system, in which the additional mode can be identified as evanescent mode, and then adapted to define radiation conditions. This makes possible the implementation of efficient numerical multimodal methods (like the admittance matrix method) and also approximate solutions can be found using the Born approximation. The numerical tests have shown that our method is very efficient to reduce the number of degree of freedom: adding to the boundary mode, it is sufficient to take only the propagative modes to get very good results. This is in particular interesting at low frequency when only the plane mode propagates. In the low frequency regime, the system can be solved analytically, using the Born approximation, leading to improved approximate equations compared to the usual Webster's approximation.

6.4.2. Construction of non scattering perturbations in a waveguide

Participants: Anne-Sophie Bonnet-Ben Dhia, Eric Lunéville.

This work is done in collaboration with Sergei Nazarov from Saint-Petersbourg University and during the internship of Yves Mbeutcha. We consider a two-dimensional homogeneous acoustic waveguide and we aim at designing deformations of the boundary which are invisible at a given frequency (or more generally at a finite number of given frequencies) in the sense that they are non scattering. To find such invisible perturbations, we take advantage of the fact that there are only a finite number of propagative modes at a given frequency in a waveguide. As a consequence, the invisibility is achieved by canceling a finite number of scattering coefficients, and an invisible deformation only produces an exponentially decreasing scattered field, not measurable in the far field.

The first step consists in studying the effect of a small deformation, of amplitude ε. The asymptotic analysis allows to derive the first order terms of the scattering coefficients, as integrals involving the function describing the deformation. This leads to express the deformation as a linear combination of some explicit (compactly supported) functions, so that invisibility is satisfied if and only if the coefficients of the linear combination are solution of a fixed point equation. The key point is that we can prove, using the results of the asymptotic analysis, that the function of this fixed point equation is a contraction for ε small enough. This proves the existence of invisible deformations of amplitude ε. Moreover, it provides a natural algorithm to compute the invisible deformation.

This has been tested numerically and the results are in perfect agreement with the theory. At low frequency, the good news is that ε can be taken quite large (the amplitude of the deformation may be half the size of the guide). But this deteriorates when the frequency increases.

6.4.3. Localized modes in unbounded perturbed periodic media

Participants: Patrick Joly, Sonia Fliss, Elizaveta Vasilevskaya.

This topic is investigated in collaboration with Bérangère Delourme (Univ. Paris XIII) and constitutes the subject of the E. Vasilevskaya's PhD thesis. We are interested in a 2D propagation medium which is a localized perturbation of a reference homogeneous periodic reference medium. This reference medium is a "thick graph", namely a thin structure (the thinness being characterized by the parameter $\delta > 0$) whose limit when δ tends to 0 is a periodic graph. This is for instance the case of the thick periodic ladder and the thick periodic rectangular grid of figure. The perturbation consists in changing only the geometry (and not the material properties) of the reference medium by modifying the thickness of one of the lines of the reference medium as illustrated by figure with the perturbed ladder and perturbed grid (see figure 3). The question we investigate is whether such a geometrical perturbation is able to produce localized eigenmodes (for the ladder) or guided modes (for the grid). We have investigated this question when the propagation model is the scalar Helmholtz equation with Neumann boundary conditions (in opposition to Dirichlet conditions that have been more studied in the literature - see the works by S. Nazarov for instance). This amounts to solving an eigenvalue problem for the Laplace operator in an unbounded domain : the associated self-adjoint operator has a continuous spectrum with a band gap structure and the eigenvalues are searched in the gaps.

../../../../projets/poems/IMG/ladder.png ../../../../projets/poems/IMG/grid.png

Figure 3. Left : periodic ladder (non perturbed/perturbed). Right : periodic "thick graph" (non perturbed/perturbed). The propagation domain is in grey.

With Neumann boundary conditions, we can use for the theoretical study an asymptotic analysis with respect to δ: indeed, it is well known (see in particular the works by Exner, Kuchment, Post) the limit model when δ tends to 0, is the Helmholtz equation on the graph : 1D Helmholtz equations on each branch completed by continuity and Kirchoff transmission conditions at each node. The geometrical perturbation of the original medium results into a perturbation of the Kirchoff conditions on the nodes of the modified line. The spectral analysis of the limit problem can be done completely by hand and the existence of eigenmodes for the thick medium is ensured, for δ small enough, by the existence of corresponding eigenmodes for a limit "1D operator" whose spectrum appears to possess an infinity of band gaps in each of which eigenvalues can exist, due to the perturbation. Following this idea, we have been able to prove the existence of localized modes in the case of the ladder provided that the geometrical perturbation consists in diminishing the width of one rung. One can even prove that one can produce more and more localized modes, corresponding to larger and larger frequencies, when δ is smaller and smaller. On the contrary, we conjecture that there is no localized modes when we enlarge the rung. The extension of these results to the existence of guided modes in the case of the grid in progress.

For the numerical computation of such localized modes, we have adapted the DtN approach discussed in the activity report of 2012. We gave in figure 4 an example of computed localized mode in the case of the ladder : this mode is geometrically confined at the neighbourhood of the modified rung.

../../../../projets/poems/IMG/Modelocalise.png

Figure 4. Localized mode in the perturbed ladder.

6.5. Asymptotic methods and approximate models

6.5.1. Homogenization and interfaces

Participants: Sonia Fliss, Valentin Vinoles.

This topic is developed in collaboration with Xavier Claeys (LJLL, Paris VI).

The mathematical modelling of electromagnetic metamaterials and the homogenization theory are intimately related because metamaterials are precisely constructed by a periodic assembly of small resonating micro-structures involving dielectric materials presenting a high contrast with respect to a reference medium. In the framework of the ANR Metamath (see 6.2.6), we wish to look carefully at the treatment of boundaries and interfaces that are generally poorly taken into account by the first order homogenization.

This question is already relevant for standard homogenization (ie without high contrast). Indeed, the presence of a boundary induces a loss of accuracy due to the inadequateness of the standard homogenization approach to take into account boundary layer effects. Our objective is to construct approximate effective boundary conditions that would restore the desired accuracy.

We first considered a plane interface between a homogeneous and the periodic media in the standard case without high-contrast. We obtained high order transmission conditions between the homogeneous media and the periodic media. The technique we used involves matched asymptotic expansions combined with standard homogenization ansatz. Those conditions are non standard : they involve Laplace-Beltrami operators at the interface and requires to solve cell problems in infinite periodic waveguides. The derivation of the corresponding error estimates is in progress. The analysis is based on a original combination of Floquet-Bloch and a periodic version of Kondratiev technique.

The next step will be to consider the same problem but with a high-contrast periodic media in collaboration with Guy Bouchitté, a french expert in high contrast homogenization.

6.5.2. Effective boundary conditions for thin periodic coatings

Participants: Mathieu Chamaillard, Patrick Joly.

This topic is the object of a collaboration with Houssem Haddar (CMAP École Polytechnique). We are interested in the construction of "equivalent" boundary condition for the diffraction of waves by an obstacle with smooth boundary Γ covered with a thin coating of width δ whose physical characteristics vary "periodically" along Γ with a period proportional to the small parameter δ. For a general boundary Γ, the notion of periodicity is ambiguous: we have chosen to define the coating as the image, or the deformation, by a smooth mapping ψ_Γ of a flat layer of width δ (the reference configuration) that preserves the normals, which appears consistent with a manufacturing process. The electromagnetic parameters in the coating are then defined as the images through ψ_Γ of periodic functions in the reference configuration.

We have first considered the case of the scalar wave equation when the homogeneous Neumann condition is applied on the boundary of the obstacle. Using an asymptotic analysis in δ, which combines homogenization and matched asymptotic expansions, we have been able to establish a second order boundary condition of the form

$$\partial_\nu u + (\delta B_\Gamma^1 + \delta^2 B_\Gamma^2)u = 0,$$

where B_Γ^1 and B_Γ^2 are second order tangential differential operators along Γ. The coefficients of these operators depend on both the geometrical characteristics of Γ (through the curvature tensor), the deformation mapping ψ_Γ and the material properties of the coating, through the resolution of particular unbounded cell problems in the flat reference configuration. When the coating is homogeneous, we have checked that one recovers the well known second order thin layer condition. We have moreover proven that this approximate condition provides in $\mathcal{O}(\delta^3)$.

6.5.3. Thin Layers in Isotropic Elastodynamics

Participants: Marc Bonnet, Aliénor Burel, Patrick Joly.

This research is concerned with the numerical modelling of non-destructive testing experiments using ultrasonic waves. Some materials, e.g. composite materials, involve thin layers of resin. The numerical modelling of such thin layers can be problematic as they result in very small spatial mesh sizes. To alleviate this difficulty, we develop an approach based on an asymptotic analysis with respect to the layer thickness ε, aiming to model the thin layer by approximate effective transmission conditions (ETCs), which remove the need to mesh the layer. So far, ETCs that are second-order accurate in ε have been formulated, justified, implemented and numerically validated, for 2-D and 3-D configurations involving planar interfaces of constant thickness. In particular, the continuous evolution problem is shown to be stable, and a time-stepping scheme that essentially preserves the stability requirement on the time step is proposed. Extension of this work to 2-D and 3-D configurations involving a curved layer is ongoing.

6.5.4. Mathematical modelling of electromagnetic wave propagation in electric networks.

Participants: Geoffrey Beck, Patrick Joly.

This topic is developed in collaboration with S. Imperiale (Inria Saclay) in the framework of the ANR project SODDA, in collaboration with CEA-LETI, about the non destructive testing of electric networks. This is the subject of the PhD thesis of G. Beck.

We investigate the question of the electromagnetic propagation in thin electric cables from a mathematical point of view via an asymptotic analysis with respect to the (small) transverse dimension of the cable: as it has been done in the past in mechanics for the beam theory from 3D elasticity, we use such an approach for deriving simplified effective 1D models from 3D Maxwell's equations.

During last year, we have achieved some progress in various directions:

- Single wire coaxial cables. This is the direct continuation of what has been done last year. Concerning the lowest order, the telegraphist's model, we have extended the error analysis, previously restricted to non lossy cylindrical cables to very general cases. Technically, this relies on time Laplace transform and new, parameter dependent, Poincaré-Friedrichs inequalities. From the numerical point of view, in collaboration with M. Duruflé, we have initiated a quantitative comparison between the full 3D model and our 1D model. Furthermore we have derived and studied a higher order generalized telegraphist's equation that include dispersive effects through nonlocal capacity and inductance operators. The corresponding mathematical analysis is in progress.
- Multiple wires cables. The objective here was to extend our approach to cables containing N conducting wires. Our results into a vectorial generalized telegraphist's model with $2N$ (2 for each wire) 1D unknowns, N electrical potentials and N currents. This model involves in particular a capacity matrix C, an inductance matrix L, a resistance matrix R and a conductance matrix G, whose properties have been deeply investigated, which allowed us to justify rigorously and extend some results from the electrical engineering literature. In the most general case, the effective models also involve time memory terms with matrix valued convolution kernels.
- Junction of cables. This is a new and essential step towards the modelling of networks. We have started the case of junctions of single wire cables via the method of matched asymptotic expansions in the spirit of the PhD thesis of A. Semin.

6.5.5. Elastic wave propagation in strongly heterogeneous media

Participants: Simon Marmorat, Patrick Joly.

This subject enters our long term collaboration with CEA-LIST on the development of numerical methods for time-domain non destructive testing experiments using ultra-sounds, and is realized in collaboration with Xavier Claeys (LJLL, Paris VI). We aim at developing an efficient numerical approach to simulate the propagation of waves in a medium made of many small heterogeneities, embedded in a smooth (or piecewise smooth) background medium, without any particular assumption (such as periodicity) on the spatial distribution of these heterogeneities. The figure 5 is a snapshot of a simulation inside such a medium, computed thanks to classical simulation tools: to reach satisfying accuracy, one has to use mesh refinement in the vicinity of the heterogeneities, which greatly increases the computational cost of the method.

../../../../projets/poems/IMG/snap2.png

Figure 5. Snapshot of a simulation in the medium of interest, using high-order finite element method as well as local mesh refinement and local time stepping around the heterogeneities.

By considering the medium with defects as a perturbation of the smooth one, we have derived an auxiliary model in the acoustic case, involving the defect-free wave operator and some volume Lagrange multipliers which account for the presence of the defects. These Lagrange multipliers are unknown functions defined on the defects and live in some infinite dimensional functional space. Exploiting the smallness of the defects, we have shown thanks to matched asymptotic analysis that the aforementioned functional space may be well described by a finite number N of profile functions: we propose an asymptotic model by looking for the Lagrange multipliers into the space spanned by these N profile functions, and we have shown that the error hence made is controlled by ε^N, ε being the characteristic size of the defects, assumed to be small.

On a computational point of view, the asymptotic model is much easier to solve than the original one since it can be discretized using a computation mesh that ignores the presence of the heterogeneities, the Lagrangian multipliers being computed by solving a linear system of size N. A resolution of this model has been implemented in the 1D and in the 2D case, and a rigorous error estimate has been established.

6.6. Imaging and inverse problems

6.6.1. Sampling methods in waveguides

Participants: Laurent Bourgeois, Sonia Fliss, Eric Lunéville, Anne-Claire Egloffe.

First, we have adapted the modal formulation of sampling methods (Linear Sampling Method and Factorization Method) to the case of a periodic waveguide in the acoustic case. This study is based on the analysis of the far field of scattering solutions in cylindrical waveguides, in particular for the fundamental solution, which enables us to obtain a far field formulation of sampling methods, and then a modal formulation of such methods. The aim of the inverse problem is to retrieve a defect (that is a loss of periodicity) from the scattered fields which correspond to the incident fields formed by the Floquet modes. Some convincing numerical experiments have shown the feasibility of the method. Secondly, going back to the homogeneous waveguide in the acoustic case, we have started a study of the sampling methods in a more realistic situation, that is the data (emission and reception) are measured on the boundary of the waveguide in the time domain. This was the subject of Anne-Claire Egloffe's post-doc. The aim is to use the modal formulation of the sampling methods at all frequencies and recompose the best possible image of the defect. Some first encouraging results have been obtained when the spectrum of the incident signal is centered at a rather low frequency (corresponding to 3 propagating guided modes).

6.6.2. Space-time focusing on unknown scatterers

Participants: Maxence Cassier, Patrick Joly, Christophe Hazard.

This topic concerns the studies about time-reversal in the context of Maxence Cassier's thesis. We are motivated by the following challenging question: in a propagative medium which contains several unknown scatterers, how can one generate a wave that focuses selectively on one scatterer not only in space, but also in time? In other words, we look for a wave that 'hits hard at the right spot'. Such focusing properties have been studied in the frequency domain in the context of the DORT method ("Decomposition of the Time Reversal Operator"). In short, an array of transducers first emits an incident wave which propagates in the medium. This wave interacts with the scatterers and the transducers measure the scattered field. The DORT method consists in doing a Singular Value Decomposition (SVD) of the scattering operator, that is, the operator which maps the input signals sent to the transducers to the measure of the scattered wave. It is now well understood that for small and distant enough scatterers, each singular vector associated with a non zero singular value generates a wave which focuses selectively on one scatterer. Can we take advantage of these spatial focusing properties in the frequency domain to find the input signals which generate a time-dependent wave which would be also focused in time? Since any frequency superposition of a family of singular vectors associated with a given scatterer leads to a spatial focusing, the main question is to synchronize them by a proper choice of their phases. The method we propose is based on a particular SVD of the scattering operator related to its symmetry. The signals we obtain do not require the knowledge of the locations of the scatterers. We compare it with some "optimal" signals which require this knowledge. Our study is illustrated by a two dimensional acoustic model where both scatterers and transducers are assumed pointlike (see figure 6).

6.6.3. The exterior approach to retrieve obstacles

Participant: Laurent Bourgeois.

This theme is a collaboration with Jérémi Dardé from IMT (Toulouse). The aim is to find a fixed Dirichlet obstacle in a bounded domain by using some redundant boundary conditions (Cauchy data) on the accessible part of the boundary, while the boundary conditions are unknown on the inaccessible part of the boundary. We wish to adapt the exterior approach developped for the Laplace equation and the Stokes system to the case of time evolution problems, in particular the heat equation. The exterior approach consists in defining a decreasing sequence of domains that converge in some sense to the obstacle. More precisely, such iterative approach is based on a combination of a quasi-reversibility method to update the solution of the ill-posed Cauchy problem outside the obstacle obtained at previous iteration and of a level set method to update the obstacle with the help of the solution obtained at previous iteration. We have already introduced two different mixed formulations of quasi-reversibility for the ill-posed heat equation with lateral Cauchy data in order to use standard Lagrange finite elements.

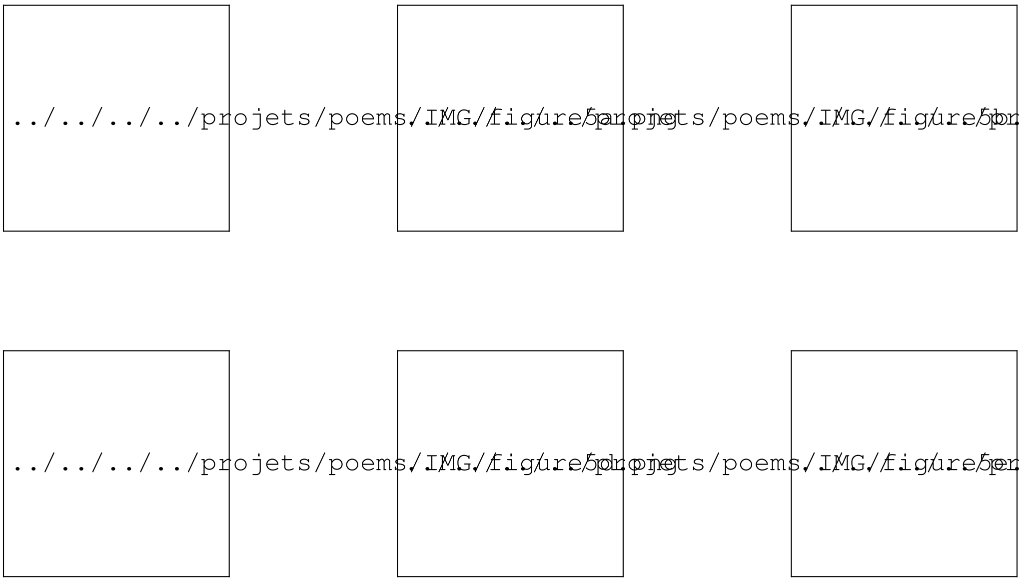


Figure 6. The case of a scattering reference medium perturbed by two obstacles (the white circles) : modulus of the field generated by 128 transducers (left edge of each figure) at different times.

6.6.4. Uniqueness and stability of inverse problems

Participant: Laurent Bourgeois.

In collaboration with Laurent Baratchart and Juliette Leblond from APICS (Nice), we have proved uniqueness for the inverse Robin problem with a boundary coefficient in L^∞ in the 2D case, for the Laplace equation in the divergence form. The result is based on complex analysis. We have also established an abstract Lipschitz stability result for inverse problems such that the set of parameters is a compact and convex subset of a finite dimensional space. In particular, such result can be applied to the previous inverse Robin problem.

6.6.5. Interior transmission problem

Participant: Anne-Sophie Bonnet-Ben Dhia.

This work is in collaboration with Lucas Chesnel (Aalto University, Finland). During this year, we investigated a two-dimensional interior transmission eigenvalue problem for an inclusion made of a composite material. This problem plays a central role in the theory of the corresponding inverse problem. We considered configurations where the difference between the parameters of the composite material and the ones of the background change sign on the boundary of the inclusion. In a first step, under some assumptions on the parameters, we extended the variational approach of the T-coercivity to prove that the transmission eigenvalues form at most a discrete set. In the process, we also provided localization results. Then, we study what happens when these assumptions are not satisfied. The main idea is that, due to very strong singularities that can occur at the boundary, the problem may lose Fredholmness in the natural H^1 framework. Using Kondratiev theory, we proposed a new functional framework where the Fredholm property is restored.

6.6.6. Flaw identification using elastodynamic topological derivative or transmission eigenvalues

Participants: Marc Bonnet, Rémi Cornaggia.

This work is in collaboration with C. Bellis (LMA, CNRS, Marseille), F. Cakoni (Univ. of Delaware, USA) and B. Guzina (Univ. of Minnesota, USA). The concept of topological derivative (TD) quantifies the perturbation induced to a given cost functional by the nucleation of an infinitesimal flaw in a reference defect-free body, and may serve as a flaw indicator function. In this work, the TD is derived for three-dimensional crack identification exploiting over-determined transient elastodynamic boundary data. This entails in particular the derivation of the relevant polarization tensor, here given for infinitesimal trial cracks in homogeneous or bi-material elastic bodies. Simple and efficient adjoint-state based formulations are used for computational efficiency, allowing to compute the TD field for arbitrarily shaped elastic solids. The latter is then used as an indicator function for the spatial location of the sought crack(s). The heuristic underpinning TD-based identification, which consists in deeming regions where the TD is most negative as the likeliest locations of actual flaws and on formulating higher-order topological expansions in the elastodynamic case, has (with C. Bellis and F. Cakoni) been given a partial justification within the limited framework of acoustic inverse scattering using far-field data. Current investigations (M. Bonnet, R. Cornaggia) include setting up and justifying the formulation of higher-order topological expansions for the elastostatic and elastodynamic cases.

Another ongoing research on a related topic addresses the use of transmission eigenvalues (TEs), i.e. values of the wave number for which the homogeneous interior transmission problem (ITP) related to the scattering scattering of time-harmonic elastic waves by aninhomogeneity D admits non-trivial solutions. This works (R. Cornaggia, in collaboration with C. Bellis, F. Cakoni, B. Guzina) aims on the one hand to understand better how to compute the TEs -if any- in the case where D's characteristics vary periodically. On the other hand it looks for how a previously obtained knowledge of the TE set could be the basis of an identification process. In a preliminary study considering 1-D elastic beams with periodically varying section over a length L, gradient elasticity was found to be a well-suited homogenization model to both compute the TEs and identify L, the periodic cell length and the damage parameter from available values of the TEs.

6.6.7. Topological derivative in anisotropic elasticity

Participant: Marc Bonnet.

This work is in collaboration with G. Delgado (PhD student, CMAP Ecole Polytechnique and EADS IW). Following up on previous work on the topological derivative (TD) of displacement-based cost functionals in anisotropic elasticity, a TD formula has been derived and justified for general cost functionals that involve strains (or displacement gradients) rather than displacements. The small-inclusion asymptotics of such cost functionals are quite different than in the previous case, due to the fact that the strain perturbation inside an elastic inclusion has a finite, nonzero asymptotic value in the limit of a vanishingly small inclusion. Cost functionals of practical interest having this format include von Mises equivalent stress (often used in plasticity or failure criteria) and energy-norm error functionals for coefficient-identification inverse problems. This TD formulation has been tested on 2D and 3D numerical examples, some of them involving anisotropic elasticity and nonquadratic cost functionals.

6.6.8. Energy functionals for elastic medium reconstruction using transient data

Participant: Marc Bonnet.

This work is in collaboration with W. Aquino (Duke Univ., USA). Energy-based misfit cost functionals, known in mechanics as error in constitutive relation (ECR) functionals, are known since a long time to be well suited to (electrostatic, elastic,...) medium reconstruction. In this ongoing work, a transient elastodynamic version of this methodology is developed, with emphasis on its applicability to large time-domain finite element modeling of the forward problem. The formulation involves coupled transient forward and adjoint solutions, which greatly hinders large-scale computations. A computational approach combining an iterative treatment of the coupled problem and the adjoint to the discrete Newmark time-stepping scheme is found to perform well on cases where both the FE model and the identification problem are of large size (2D and 3D elastodynamic numerical experiments made so far involve up to half a million unknown for the discretized inverse problem), making the time-domain ECR functional a worthwhile tool for medium identification.

POLSYS Project-Team

6. New Results

6.1. Fundamental Algorithms and Structured Systems

6.1.1. Structured polynomial systems: the quasi-homogeneous case

Let $\mathbb{K}$ be a field and $(f_1, ..., f_n) \subset \mathbb{K}[X_1, ..., X_n]$ be a sequence of quasi-homogeneous polynomials of respective weighted degrees $(d_1, ..., d_n)$ w.r.t a system of weights $(w_1, \cdots, w_n)$. Such systems are likely to arise from a lot of applications, including physics or cryptography. In [29], we design strategies for computing Gröbner bases for quasi-homogeneous systems by adapting existing algorithms for homogeneous systems to the quasi-homogeneous case. Overall, under genericity assumptions, we show that for a generic zero-dimensional quasi-homogeneous system, the complexity of the full strategy is polynomial in the weighted Bézout bound $\prod_{i=1}^{n} d_i / \prod_{i=1}^{n} w_i$. We provide some experimental results based on generic systems as well as systems arising from a cryptography problem. They show that taking advantage of the quasi-homogeneous structure of the systems allow us to solve systems that were out of reach otherwise.

6.1.2. Structured polynomial systems: the determinantal case

In [13], We study the complexity of solving the *generalized MinRank problem*, i.e. computing the set of points where the evaluation of a polynomial matrix has rank at most r. A natural algebraic representation of this problem gives rise to a *determinantal ideal*: the ideal generated by all minors of size $r + 1$ of the matrix. We give new complexity bounds for solving this problem using Gröbner bases algorithms under genericity assumptions on the input matrix. In particular, these complexity bounds allow us to identify families of generalized MinRank problems for which the arithmetic complexity of the solving process is polynomial in the number of solutions. We also provide an algorithm to compute a rational parametrization of the variety of a 0-dimensional and radical system of bi-degree $(D, 1)$. We show that its complexity can be bounded by using the complexity bounds for the generalized MinRank problem.

6.1.3. On the Complexity of the Generalized MinRank Problem

In [13] we study the complexity of solving the *generalized MinRank problem*, i.e. computing the set of points where the evaluation of a polynomial matrix has rank at most r. A natural algebraic representation of this problem gives rise to a *determinantal ideal*: the ideal generated by all minors of size $r + 1$ of the matrix. We give new complexity bounds for solving this problem using Gröbner bases algorithms under genericity assumptions on the input matrix. In particular, these complexity bounds allow us to identify families of generalized MinRank problems for which the arithmetic complexity of the solving process is polynomial in the number of solutions. We also provide an algorithm to compute a rational parametrization of the variety of a 0-dimensional and radical system of bi-degree $(D, 1)$. We show that its complexity can be bounded by using the complexity bounds for the generalized MinRank problem.

6.1.4. On the Complexity of Computing Gröbner Bases for Quasi-homogeneous Systems

Let $\mathbb{K}$ be a field and $(f_1, ..., f_n) \subset \mathbb{K}[X_1, ..., X_n]$ be a sequence of quasi-homogeneous polynomials of respective weighted degrees $(d_1, ..., d_n)$ w.r.t a system of weights $(w_1, \cdots, w_n)$. Such systems are likely to arise from a lot of applications, including physics or cryptography.

In [29], we design strategies for computing Gröbner bases for quasi-homogeneous systems by adapting existing algorithms for homogeneous systems to the quasi-homogeneous case. Overall, under genericity assumptions, we show that for a generic zero-dimensional quasi-homogeneous system, the complexity of the full strategy is polynomial in the weighted Bézout bound $\prod_{i=1}^{n} d_i / \prod_{i=1}^{n} w_i$.

We provide some experimental results based on generic systems as well as systems arising from a cryptography problem. They show that taking advantage of the quasi-homogeneous structure of the systems allow us to solve systems that were out of reach otherwise.

6.1.5. Gröbner bases of ideals invariant under a commutative group : the non-modular case

In [30], we propose efficient algorithms to compute the Gröbner basis of an ideal $I \subset k[x_1, \cdots, x_n]$ globally invariant under the action of a commutative matrix group G, in the non-modular case (where $char(k)$ doesn't divide $|G|$). The idea is to simultaneously diagonalize the matrices in G, and apply a linear change of variables on I corresponding to the base-change matrix of this diagonalization. We can now suppose that the matrices acting on I are diagonal. This action induces a grading on the ring $R = k[x_1, \cdots, x_n]$, compatible with the degree, indexed by a group related to G, that we call G-degree. The next step is the observation that this grading is maintained during a Gröbner basis computation or even a change of ordering, which allows us to split the Macaulay matrices into $|G|$ submatrices of roughly the same size. In the same way, we are able to split the canonical basis of R/I (the staircase) if I is a zero-dimensional ideal. Therefore, we derive *abelian* versions of the classical algorithms F_4, F_5 or FGLM. Moreover, this new variant of F_4/F_5 allows complete parallelization of the linear algebra steps, which has been successfully implemented. On instances coming from applications (NTRU crypto-system or the Cyclic-n problem), a speed-up of more than 400 can be obtained. For example, a Gröbner basis of the Cyclic-11 problem can be solved in less than 8 hours with this variant of F_4. Moreover, using this method, we can identify new classes of polynomial systems that can be solved in polynomial time.

6.1.6. Signature Rewriting in Gröbner Basis Computation

In [27] we introduce the RB algorithm for Gröbner basis computation, a simpler yet equivalent algorithm to F5GEN. RB contains the original unmodified F5 algorithm as a special case, so it is possible to study and understand F5 by considering the simpler RB. We present simple yet complete proofs of this fact and of F5's termination and correctness. RB is parametrized by a rewrite order and it contains many published algorithms as special cases, including SB. We prove that SB is the best possible instantiation of RB in the following sense. Let X be any instantiation of RB (such as F5). Then the S-pairs reduced by SB are always a subset of the S-pairs reduced by X and the basis computed by SB is always a subset of the basis computed by X.

6.1.7. An analysis of inhomogeneous signature-based Gröbner basis computations

In [8] we give an insight into the behaviour of signature-based Gröbner basis algorithms, like F5, G2V or SB, for inhomogeneous input. On the one hand, it seems that the restriction to sig-safe reductions puts a penalty on the performance. The lost connection between polynomial degree and signature degree can disallow lots of reductions and can lead to an overhead in the computations. On the other hand, the way critical pairs are sorted and corresponding s-polynomials are handled in signature- based algorithms is a very efficient one, strongly connected to sorting w.r.t. the well-known sugar degree of polynomials.

6.1.8. Improving incremental signature-based Gröbner basis algorithms

In [9] we describe a combination of ideas to improve incremental signature-based Gröbner basis algorithms having a big impact on their performance. Besides explaining how to combine already known optimizations to achieve more efficient algorithms, we show how to improve them even more. Although our idea has a positive affect on all kinds of incremental signature-based algorithms, the way this impact is achieved can be quite different. Based on the two best-known algorithms in this area, F5 and G2V, we explain our idea, both from a theoretical and a practical point of view.

6.1.9. A new algorithmic scheme for computing characteristic sets

Ritt-Wu's algorithm of characteristic sets is the most representative for triangularizing sets of multivariate polynomials. Pseudo-division is the main operation used in this algorithm. In [18] we present a new algorithmic scheme for computing generalized characteristic sets by introducing other admissible reductions than pseudo-division. A concrete subalgorithm is designed to triangularize polynomial sets using selected admissible reductions and several effective elimination strategies and to replace the algorithm of basic sets (used in Ritt-Wu's algorithm). The proposed algorithm has been implemented and experimental results show that it performs better than Ritt-Wu's algorithm in terms of computing time and simplicity of output for a number of non-trivial test examples

6.2. Solving Systems over the Reals and Applications

6.2.1. On the Boolean complexity of real root refinement

In [32] we assume that a real square-free polynomial A has a degree d, a maximum coefficient bitsize τ and a real root lying in an isolating interval and having no nonreal roots nearby (we quantify this assumption). Then, we combine the *Double Exponential Sieve* algorithm (also called the *Bisection of the Exponents*), the bisection, and Newton iteration to decrease the width of this inclusion interval by a factor of $t = 2^{-L}$. The algorithm has Boolean complexity $\widetilde{O}_B(d^2\tau + dL)$. Our algorithms support the same complexity bound for the refinement of r roots, for any $r \leq d$.

6.2.2. On the minimum of a polynomial function on a basic closed semialgebraic set and applications

In [17] we give an explicit upper bound for the algebraic degree and an explicit lower bound for the absolute value of the minimum of a polynomial function on a compact connected component of a basic closed semialgebraic set when this minimum is not zero. We also present extensions of these results to non-compact situations. As an application, we obtain a lower bound for the separation of two disjoint connected components of basic closed semialgebraic sets, when at least one of them is compact.

6.2.3. Rational solutions to Linear Matrix Inequalities and Sums of Squares

Consider a $(D \times D)$ symmetric matrix A whose entries are linear forms in $\mathbb{Q}[X_1, ..., X_k]$ with coefficients of bit size $\leq \tau$. In [31], we provide an algorithm which decides the existence of rational solutions to the linear matrix inequality $\mathsf{A} \succeq 0$ and outputs such a rational solution if it exists. This problem is of first importance: it can be used to compute algebraic certificates of positivity for multivariate polynomials. Our algorithm runs within $(k\tau)^{O(1)} 2^{O(\min(k,D)D^2)} D^{O(D^2)}$ bit operations; the bit size of the output solution is dominated by $\tau^{O(1)} 2^{O(\min(k,D)D^2)}$. These results are obtained by designing algorithmic variants of constructions introduced by Klep and Schweighofer. This leads to the best complexity bounds for deciding the existence of sums of squares with rational coefficients of a given polynomial. We have implemented the algorithm; it has been able to tackle Scheiderer's example of a multivariate polynomial that is a sum of squares over the reals but not over the rationals; providing the first computer validation of this counter-example to Sturmfels' conjecture.

6.2.4. Exact Voronoi diagram of smooth convex pseudo-circles: General predicates, and implementation for ellipses

In [10] we examine the problem of computing exactly the Voronoi diagram (via the dual Delaunay graph) of a set of, possibly intersecting, smooth convex pseudo-cirlces in the Euclidean plane, given in parametric form. Pseudo-circles are (convex) sites, every pair of which has at most two intersecting points. The Voronoi diagram is constructed incrementally. Our first contribution is to propose robust and efficient algorithms, under the exact computation paradigm, for all required predicates, thus generalizing earlier algorithms for non-intersecting ellipses. Second, we focus on INCIRCLE, which is the hardest predicate, and express it by a simple sparse 5×5 polynomial system, which allows for an efficient implementation by means of successive Sylvester resultants and a new factorization lemma. The third contribution is our CGAL-based C++ software for the case of possibly intersecting ellipses, which is the first exact implementation for the problem. Our code spends about a minute to construct the Voronoi diagram of 200 ellipses, when few degeneracies occur. It is faster than the CGAL segment Voronoi diagram, when ellipses are approximated by k-gons for $k > 15$, and a state-of-the-art implementation of the Voronoi diagram of points, when each ellipse is approximated by more than 1250 points.

6.2.5. Patience of Matrix Games

In [15], for matrix games we study how small nonzero probability must be used in optimal strategies. We show that for $n \times n$ win-lose-draw games (i.e. $(-1, 0, 1)$ matrix games) nonzero probabilities smaller than $n^{-O(n)}$ are never needed. We also construct an explicit $n \times n$ win-lose game such that the unique optimal strategy uses a nonzero probability as small as $n^{-\Omega(n)}$. This is done by constructing an explicit $(-1, 1)$ nonsingular $n \times n$ matrix, for which the inverse has only nonnegative entries and where some of the entries are of value $n^{\Omega(n)}$.

6.2.6. A polynomial approach for extracting the extrema of a spherical function and its application in diffusion MRI

Antipodally symmetric spherical functions play a pivotal role in diffusion MRI in representing sub-voxel-resolution microstructural information of the underlying tissue. This information is described by the geometry of the spherical function. In [14] we propose a method to automatically compute all the extrema of a spherical function. We then classify the extrema as maxima, minima and saddle-points to identify the maxima. We take advantage of the fact that a spherical function can be described equivalently in the spherical harmonic (SH) basis, in the symmetric tensor (ST) basis constrained to the sphere, and in the homogeneous polynomial (HP) basis constrained to the sphere. We extract the extrema of the spherical function by computing the stationary points of its constrained HP representation. Instead of using traditional optimization approaches, which are inherently local and require exhaustive search or re-initializations to locate multiple extrema, we use a novel polynomial system solver which analytically brackets all the extrema and refines them numerically, thus missing none and achieving high precision. To illustrate our approach we consider the Orientation Distribution Function (ODF). In diffusion MRI the ODF is a spherical function which represents a state-of-the-art reconstruction algorithm whose maxima are aligned with the dominant fiber bundles. It is, therefore, vital to correctly compute these maxima to detect the fiber bundle directions. To demonstrate the potential of the proposed polynomial approach we compute the extrema of the ODF to extract all its maxima. This polynomial approach is, however, not dependent on the ODF and the framework presented in this line of work can be applied to any spherical function described in either the SH basis, ST basis or the HP basis.

6.2.7. Improving Angular Speed Uniformity by Reparameterization

In [20] we introduce the notion of angular speed uniformity as a quality measure for parameter-izations of plane curves and propose an algorithm to compute uniform reparameterizations for quadratic and cubic curves. We prove that only straight lines have uniform rational parameterizations. For any plane curve other than lines, we show how to find a rational reparameterization that has the maximum uniformity among all the rational parameterizations of the same degree. We also establish specific results for quadratic and certain cubic Bézier curves.

6.2.8. Formalization and Specification of Geometric Knowledge Objects

[7] presents our work on the identification, formalization, structuring, and specification of geometric knowledge objects for the purpose of semantic representation and knowledge management. We classify geometric knowledge according to how it has been accumulated and represented in the geometric literature, formalize geometric knowledge statements by adapting the language of first-order logic, specify knowledge objects with embedded knowledge in a retrievable and extensible data structure, and organize them by modeling the hierarchic structure of relations among them. Some examples of formal specification for geometric knowledge objects are given to illustrate our approach. The underlying idea of the approach has been used successfully for automated geometric reasoning, knowledge base creation, and electronic document generation.

6.2.9. A Framework for Improving Uniformity of Parameterizations of Curves

In [16] we define quasi-speed as a generalization of linear speed and angular speed for parameterizations of curves and use the uniformity of quasi-speed to measure the quality of the parameterizations. With such conceptual setting, a general framework is developed for studying uniformity behaviors under reparameterization via proper parameter transformation and for computing reparameterizations with improved uniformity of quasispeed by means of optimal single-piece, C^0 piecewise, and C^1 piecewise Möbius transformations. Algorithms are described for uniformity-improved reparameterization using different Möbius transformations with different optimization techniques. Examples are presented to illustrate the concepts, the framework, and the algorithms. Experimental results are provided to validate the framework and to show the efficiency of the algorithms.

6.3. Solving Systems in Finite Fields, Applications in Cryptology and Algebraic Number Theory

6.3.1. On the Complexity of Solving Quadratic Boolean Systems

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. 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.

6.3.2. Decomposing polynomial sets into simple sets over finite fields: The positive-dimensional case

Our work in [19] presents an algorithm for decomposing any positive-dimensional polynomial set into simple sets over an arbitrary finite field. The algorithm is based on some relationship established between simple sets and radical ideals, reducing the decomposition problem to the problem of computing the radicals of certain ideals. In addition to direct application of the algorithms of Matsumoto and Kemper, the algorithm of Fortuna and others is optimized and improved for the computation of radicals of special ideals. Preliminary experiments with an implementation of the algorithm in Maple and Singular are carried out to show the effectiveness and efficiency of the algorithm.

6.3.3. Using Symmetries in the Index Calculus for Elliptic Curves Discrete Logarithm

In 2004, an algorithm is introduced to solve the DLP for elliptic curves defined over a non prime finite field $\mathbb{F}_{q^n}$. One of the main steps of this algorithm requires decomposing points of the curve $E(\mathbb{F}_{q^n})$ with respect to a factor base, this problem is denoted PDP. In [11], we apply this algorithm to the case of Edwards curves, the well-known family of elliptic curves that allow faster arithmetic as shown by Bernstein and Lange. More precisely, we show how to take advantage of some symmetries of twisted Edwards and twisted Jacobi intersections curves to gain an exponential factor $2^{\omega(n-1)}$ to solve the corresponding PDP where ω is the exponent in the complexity of multiplying two dense matrices. Practical experiments supporting the theoretical result are also given. For instance, the complexity of solving the ECDLP for twisted Edwards curves defined over $\mathbb{F}_{q^5}$, with $q \approx 2^{64}$, is supposed to be $\sim 2^{160}$ operations in $E(\mathbb{F}_{q^5})$ using generic algorithms compared to 2^{130} operations (multiplication of two 32-bits words) with our method. For these parameters the PDP is intractable with the original algorithm. The main tool to achieve these results relies on the use of the symmetries and the quasi-homogeneous structure induced by these symmetries during the polynomial system solving step. Also, we use a recent work on a new algorithm for the change of ordering of Gröbner basis which provides a better heuristic complexity of the total solving process.

6.3.4. A Distinguisher for High Rate McEliece Cryptosystems

The Goppa Code Distinguishing (GD) problem consists in distinguishing the matrix of a Goppa code from a random matrix. The hardness of this problem is an assumption to prove the security of code-based cryptographic primitives such as McEliece's cryptosystem. Up to now, it is widely believed that the GD problem is a hard decision problem. We present in [12] the first method allowing to distinguish alternant and Goppa codes over any field. Our technique can solve the GD problem in polynomial-time provided that the codes have sufficiently large rates. The key ingredient is an algebraic characterization of the key-recovery problem. The idea is to consider the rank of a linear system which is obtained by linearizing a particular polynomial system describing a key-recovery attack. Experimentally it appears that this dimension depends on the type of code. Explicit formulas derived from extensive experimentations for the rank are provided for "generic" random, alternant, and Goppa codes over any alphabet. Finally, we give theoretical explanations of these formulas in the case of random codes, alternant codes over any field of characteristic two and binary Goppa codes.

6.3.5. Cryptanalysis of HFE, multi-HFE and variants for odd and even characteristic

We investigate in this paper 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).

6.3.6. Cryptanalysis of a Public-Key Encryption Scheme Based on New Multivariate Quadratic Assumptions

In [24], we investigate the security of a public-key encryption scheme introduced by Huang, Liu and Yang (HLY) at PKC'12. This new scheme can be provably reduced to the hardness of solving a set of quadratic equations whose coefficients of highest degree are chosen according to a discrete Gaussian distributions. The other terms being chosen uniformly at random. Such a problem is a variant of the classical problem of solving a system of non-linear equations (PoSSo), which is known to be hard for random systems. The main hypothesis of Huang, Liu and Yang is that their variant is not easier than solving PoSSo for random instances. In this paper, we disprove this hypothesis. To this end, we exploit the fact that the new problem proposed by Huang, Liu and Yang reduces to an easy instance of the Learning With Errors (LWE) problem. The main contribution of this paper is to show that security and efficiency are essentially incompatible for the HLY proposal. That is, one cannot find parameters which yield a secure and a practical scheme. For instance, we estimate that a public-key of at least 1.03 GB is required to achieve 80-bit security against known attacks. As a proof of concept, we present practical attacks against all the parameters proposed Huang, Liu and Yang. We have been able to recover the private-key in roughly one day for the first challenge (i.e. Case 1) proposed by HLY and in roughly three days for the second challenge (i.e. Case 2).

6.3.7. On the Complexity of the BKW Algorithm on LWE

In [3], we present a study of the complexity of the Blum-Kalai-Wasserman (BKW) algorithm when applied to the Learning with Errors (LWE) problem, by providing refined estimates for the data and computational effort requirements for solving concrete instances of the LWE problem. We apply this refined analysis to suggested parameters for various LWE-based cryptographic schemes from the literature and compare with alternative approaches based on lattice reduction. As a result, we provide new upper bounds for the concrete hardness of these LWE-based schemes. Rather surprisingly, it appears that BKW algorithm outperforms known estimates for lattice reduction algorithms starting in dimension $n \approx 250$ when LWE is reduced to SIS. However, this assumes access to an unbounded number of LWE samples.

6.3.8. Combined Attack on CRT-RSA. Why Public Verification Must Not Be Public?

In [25] we introduce a new Combined Attack on a CRT-RSA implementation resistant against Side-Channel Analysis and Fault Injection attacks. Such implementations prevent the attacker from obtaining the signature when a fault has been induced during the computation. Indeed, such a value would allow the attacker to recover the RSA private key by computing the gcd of the public modulus and the faulty signature. The principle of our attack is to inject a fault during the signature computation and to perform a Side-Channel Analysis targeting a sensitive value processed during the Fault Injection countermeasure execution. The resulting information is then used to factorize the public modulus, leading to the disclosure of the whole RSA private key. After presenting a detailed account of our attack, we explain how its complexity can be significantly reduced by using Coppersmith's techniques based on lattice reduction. We also provide simulations that confirm the efficiency of our attack as well as two different countermeasures having a very small impact on the performance of the algorithm. As it performs a Side-Channel Analysis during a Fault Injection countermeasure to retrieve the secret value, this article recalls the need for Fault Injection and Side-Channel Analysis countermeasures as monolithic implementations.

6.3.9. Polynomial root finding over local rings and application to error correcting codes

GURUSWAMI and SUDAN designed a polynomial-time list-decoding algorithm. Their method divides into two steps. First it computes a polynomial Q in $\mathbb{F}_q[x][y]$ such that the possible transmitted messages are roots of Q in $\mathbb{F}_q[x]$. In the second step one needs to determine all such roots of Q. Several techniques have been investigated to solve both steps of the problem.

The Guruswami and Sudan algorithm has been adapted to other families of codes such as algebraic-geometric codes and alternant codes over fields. Extensions over certain types of finite rings have further been studied for Reed-Solomon codes, for alternant codes, and for algebraic-geometric codes. In all these cases, the two main steps of the Guruswami and Sudan algorithm are roughly preserved, but to the best of our knowledge, the second step has never been studied into deep details from the complexity point of view. In [5], we investigate root-finding for polynomials over *Galois rings*, which are often used within these error correcting codes, and that are defined as non-ramified extension of $\mathbb{Z}/p^n\mathbb{Z}$. We study the cost of our algorithms, discuss their practical performances, and apply our results to the Guruswami and Sudan list decoding algorithm over Galois rings.

POMDAPI Project-Team (section vide)

Popix Team

6. New Results

6.1. Estimation in mixed-effects diffusion models

Participant: Marc Lavielle.

We have coupled the SAEM algorithm and the extended Kalman filter for maximum likelihood estimation in mixed-effects diffusion models: we have considered some general mixed-effects diffusion models, in which observations are made at discrete time points and include measurement errors. In these models, the observed likelihood is generally not explicit, making maximum likelihood estimation of the parameters particularly complex. We have proposed a specific inference methodology for these models. In particular, it combines the SAEM algorithm with the extended Kalman filter to estimate the population parameters. We have also provided some tools for estimating the individual parameters, for recovering the individual underlying diffusion trajectories and for evaluating the model. We evaluated the methods on simulations and applied them to a pharmacokinetics example.

6.2. Estimation in mixtures of models

Participant: Marc Lavielle.

We have proposed an improved SAEM algorithm for maximum likelihood estimation in mixtures of non linear mixed effects models. This involves a new methodology for maximum likelihood estimation in mixtures of non linear mixed effects models (NLMEM). Such mixtures of models include mixtures of distributions, mixtures of structural models and mixtures of residual error models. Since the individual parameters inside the NLMEM are not observed, we have proposed to combine the EM algorithm usually used for mixtures models when the mixture structure concerns an observed variable, with the Stochastic Approximation EM (SAEM) algorithm, which is known to be suitable for maximum likelihood estimation in NLMEM and also has nice theoretical properties. The main advantage of this hybrid procedure is to avoid a simulation step of unknown group labels required by a “full” version of SAEM. The resulting MSAEM (Mixture SAEM) algorithm is now implemented in the MONOLIX software. We have also proposed several criteria for classification of subjects and estimation of individual parameters. Our numerical experiments on simulated data have shown that MSAEM performs well in a general framework of mixtures of NLMEM. Indeed, MSAEM provides an estimator close to the maximum likelihood estimator in very few iterations and is robust with regards to initialization. Our application of the method to pharmacokinetic (PK) data demonstrated the potential of the method for practical applications.

6.3. Moving meshes with freefem++

Participants: Astrid Decoene, Bertrand Maury.

The Arbitrary Lagrangian-Eulerian framework allows to compute free surface flows with the Finite Element functions defined on a fittedmesh which follows the globalmotion of the fluid domain. We have described how freefem++ can be used to implement this method, and we have provided two and three dimensional illustrations in the context of water waves.

6.4. Modeling of the oxygen transfer in the respiratory process

Participant: Bertrand Maury.

We have proposed 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

6.5. Congestion-driven dendritic growth

Participant: Bertrand Maury.

In order to observe growth phenomena in biology where dendritic shapes appear, we have proposed a simple model where a given population evolves feeded by a diffusing nutriment, but is subject to a density constraint. The particles (e.g., cells) of the population spontaneously stay passive at rest, and only move in order to satisfy some constraint, by choosing the minimal correction velocity so as to prevent overcongestion. We treat this constraint by means of projections in the space of densities endowed with the Wasserstein distance, defined through optimal transport. This allows to provide an existence result and suggests some numerical computations, in the same spirit of what the authors did for crowd motion (but with extra difficulties, essentially due to the fact that the total mass may increase). The numerical simulations show, according to the values of the parameter and in particular of the diffusion coefficient of the nutriment, the formation of dendritic patterns in the space occupied by cells.

POTIOC Team

6. New Results

6.1. Navigation techniques in 3D digital cities on mobile touch devices

Participants: Jacek Jankowski, Thomas Hulin, Martin Hachet.

This study aimed at characterizing today's most common interaction techniques for street-level navigation in 3D digital cities, for mobile touch devices, in terms of their efficiency and usability. To do so, we conducted a user study, where we compared target selection (Go-To), rate control (Joystick), position control, and stroke-based control navigation metaphors (see Figure 3). The results suggest that users performed best with the Go-To interaction technique. The subjective comments showed a preference of novices towards Go-To and expert users towards the Joystick technique. This work has been accepted for publication at the 3DUI 2014 conference [15].

../../../../projets/potioc/IMG/mobileTouchNavigation.png

Figure 3. Four techniques for navigating in a 3D city on a mobile touch device.

As part of this project on Navigation in 3D digital cities, the Potioc group also built a tutorial about interaction techniques for 3D environments. It was presented at Eurographics 2013 [13] and Web3D 2013 [14]. The goal of this work is to provide an up-to-date state-of-the-art of this topic to the community.

6.2. Interaction with spatial augmented reality for physical drawing

Participants: Jérémy Laviole, Martin Hachet.

We developed tools that enable precise interactive projection on pieces of paper. The sheets of paper are tracked by a camera while the user's inputs (e.g., touch and hovering events) are detected by a Kinect. The paper acts as a screen, its image coming from an overhead projector. The focus of this work is to use such tools to assist the creation of physical drawings and painting. In this context we propose Digital Construction Lines (DCL), in opposition with physical construction lines. Traditionally, the structure of a physical drawing can be created with construction lines which are light pencil strokes. These strokes are then erased during the drawing process. With DCL, it is not required to erase the construction lines anymore. Furthermore, it is possible to create construction lines on fragile material like a canvas for waterpainting or on fresh paint. It also enables construction lines on a dark canvas. In addition to these projection advantages, it is possible to create these DCL interactively and directly onto the support. Consequently, the DCL complement the physical ones during the creation process.

We investigated in a user study if the DCL could effectively replace the physical construction lines, and compared the performance (speed, cleanliness) between the two kinds of construction lines. In this user study we also evaluated the quality and usability of projection of thin lines in a fully controlled environment with a low-cost setup. The study showed that DCL could effectively replace physical construction lines, even though it might not be desirable. The study also showed that the drawing experience was as pleasant with projection, and with the usual tools. The feedback about the quality of tracking and projection was also positive. The only negative evaluation concerned the size of the projection area, which was limited by the resolution of the projector. This work was published as part of Jérémy Laviole's PhD thesis [4].

PapARt was also used as part of a museum exhibition on the Lascaux caves, together with other 3D UI from Potioc. This exhibitation has provided us with the opportunity to experiment with touch-based interfaces for manipulating 3D virtual objects. We targeted three tasks: observing rare objects with Cubtile, reassembling object fragments with Toucheo, and reproducing artwork with PapARt [7] (see Figure 5). These exhibitions allowed us to experiment our systems in real conditions. It led to a Living Lab, where the visitors can test our devices.

6.3. Rouages: Revealing the Mechanisms of Digital Musical Instruments to the Audience

Participants: Florent Berthaut, Martin Hachet, Pierre-Marie Plans.

We have developped Rouages [10], a mixed-reality display system associated with a 3D visualization application. Rouages reveals the mechanisms of digital musical instruments in two ways. First, by amplifying musicians' gestures with virtual extensions of the sensors. Second, by representing the sound components with 3D shapes and specific behaviors and by showing the impact of musicians' gestures on these components. In addition, we have explored new setups to enhance collaboration between musicians using our VR-based instruments. This is illustrated in Figure 6 .

6.4. Gateway driving simulator

Participants: Florian Larrue, Pauline Davignon, Martin Hachet.

As part of the SIMCA FUI project, the POTIOC team focuses on the design and evaluation of a gateway driving simulator (see Figure 7), to teach drivers how to drive an airport gateway in virtual reality, i.e., in a safe and cost- effective environment. We conducted a comprehensive user study to assess the impact of various parameters on user performances. This study allowed us to provide a set of recommendations for the design of an actual simulator.

../../../../projets/potioc/IMG/touch.png

Figure 4. Using Digital Construction Lines for spatial augmented reality-based physical drawing.

../../../../projets/potioc/IMG/BoisDeRenne.png

Figure 5. Manipulation of a 3D model and lighthing conditions for drawing on a prehistoric object in a museum.

../../../../projets/potioc/IMG/collaborationMusic.jpg

Figure 6. Collaborative setup for enhancing interaction between immersed musicians.

../../../../projets/potioc/IMG/simca.jpg

Figure 7. Gateway simulator.

6.5. Training Approaches for Brain-Computer Interfaces

Participants: Fabien Lotte, Florian Larrue, Christian Mühl.

While recent research on Brain-Computer Interfaces (BCI) has highlighted their potential for many applications, they remain barely used outside laboratories due to a lack of robustness. Spontaneous BCI (i.e., mental imagery-based BCI) often rely on mutual learning efforts by the user and the machine, with BCI users learning to produce stable EEG patterns (spontaneous BCI control being widely acknowledged as a skill) while the computer learns to automatically recognize these EEG patterns, using signal processing. Most research so far was focused on signal processing, mostly neglecting the human in the loop. However, how well the user masters the BCI skill is also a key element explaining BCI robustness. Unfortunately, despite the importance of BCI training protocols, they have been scarcely studied so far, and used mostly unchanged for years. In our work, we advocate that current human training approaches for spontaneous BCI are most likely inappropriate. We notably studied instructional design literature in order to identify the key requirements and guidelines for a successful training procedure that promotes a good and efficient skill learning. This literature study highlighted that current spontaneous BCI user training procedures satisfy very few of these requirements and hence are likely to be suboptimal. We therefore identified the flaws in BCI training protocols according to instructional design principles. We also proposed new research directions that are theoretically expected to address some of these flaws and to help users learn the BCI skill more efficiently. This work has been published in the Frontiers in Neuroscience journal [9].

On a related topic, together with colleagues from Inria Rennes (A. Lécuyer and L. Bonnet, Hybrid team) we explored the design and evaluation of multiuser BCI applications, notably to see their impact on user training and performance. We created a multiuser videogame called BrainArena in which two users can play a simple football game by means of two BCIs. They can score goals on the left or right side of the screen by simply imagining left or right hand movements (see Figure 8). To add another interesting element, the gamers can play in a collaborative manner (their two mental activities are combined to score in the same goal),

Figure 8. Multiplayer BCI-based gaming and feedback [5].

or in a competitive manner (the gamers must push the ball in opposite directions). Interestingly enough, our results showed that compared to a single player version of the same BCI-game, collaborative multiplayer BCI-gaming increased the motivation and performance of the most skilled of each player pairs, while leaving the performance of the other players unchanged, hence proving a useful tool to improve BCI training. This work has been published in the IEEE Transactions on Computational Intelligence and AI in Games journal [5].

6.6. Inducing, measuring and estimating mental and psychosocial stress from physiological signals

Participants: Christian Mühl, Camille Jeunet, Fabien Lotte.

Stress is a major societal issue with negative impacts on health and economy. Physiological computing offers a continuous, direct, and unobtrusive method for stress level assessment and computer-assisted stress management. However, stress is a complex construct and its physiology can vary depending on its source, for example cognitive load or social evaluation. To study the feasibility of physiology-based load-invariant psychosocial stress-detection, we designed a stress-induction protocol able to independently vary the relevant types of psychophysiological activity: mental and psychosocial stress. In [27], [17], we validate the efficacy of our protocol to induce psychosocial and mental stress. Our participants (N=24) had to perform a cognitive task associated with two workload conditions (low/high mental stress), in two contexts (low/high psychosocial stress), during which we recorded subject's self-reports, behavior, physiology and neurophysiology. Questionnaires showed that the subjectively perceived level of anxiety varied with the psychosocial stress induction, while perceived arousal and mental effort levels vary with mental stress induction. Behavior and physiology corroborated the validity of our protocol further. Heart rate and skin conductance globally increased after psychosocial stress induction relative to the non-stressful condition. Moreover, we demonstrated that higher workload tasks (mental stress) led to decrease in performance and a marked increase of heart rate.

../../../../projets/potioc/IMG/stressSetup.png

Figure 9. The setup of the experiment to induce and measure mental and psychosocial stress.

Based on this protocol, we also explored the effect of stress on workload estimation. Workload estimation from electroencephalographic signals (EEG) offers a highly sensitive tool to adapt the human-computer interaction to the user state. To create systems that reliably work in the complexity of the real world, a robustness against contextual changes (e.g., mood), has to be achieved. To study the resilience of state-of-the-art EEG-based workload classification against stress, we test the capability of the workload classifier to generalize across affective contexts (stress/non-stress). We show that the classifier is able to transfer between affective contexts, though performance suffers. However, cross-context training is a simple and powerful remedy allowing the extraction of features more resilient to task-unrelated variations in signal characteristics, leading to a performance comparable to within-context training and testing.

6.7. Exploring electroencephalography as an evaluation method for human-computer interaction

Participants: Jérémy Frey, Léonard Pommereau, Fabien Lotte, Christian Mühl, Martin Hachet.

Evaluating human-computer interaction is essential as a broadening population uses machines, sometimes in sensitive contexts. However, traditional evaluation methods may fail to combine real-time measures, an objective approach and data contextualization. We presented a review seeking how neuroimaging techniques can respond to such needs. We focused on electroencephalography (EEG), as it could be handled effectively during a dedicated evaluation phase. We identified workload, attention, vigilance, fatigue, error recognition, emotions, engagement, flow and immersion as being recognizable by EEG. We find that workload, attention and emotions assessments would benefit the most from EEG. Moreover, we advocate to study further error recognition through neuroimaging to enhance usability and increase user experience. This review paper was published in the proceeding of the Physiological Computing Systems (PhyCS) conference [12].

../../../../projets/potioc/IMG/photo2_xp_stereo.jpg

Figure 10. Stereoscopic dispaly studied through EEG by varying apparent depth.

Along this line of research, we also explored whether it was possible to assess the zone of comfort in stereoscopic displays using electroencephalography. Indeed, the conflict between vergence (eye movement) and accommodation (crystalline lens deformation) occurs with every stereoscopic display. It could cause important stress outside the "zone of comfort", when stereoscopic effect is too strong. This conflict has already been studied using questionnaires, during viewing sessions of several minutes. We built an experimental protocol (see Figure 10) which compares two different comfort conditions using electroencephalography over short viewing sequences. Analyses showed significant differences both in event-related potentials and in frequency bands power. By extending our protocol it should be possible to study at the same time comfort and depth perception, having a better understanding of stereoscopy.

Prima Project-Team

5. New Results

5.1. Attention-Based Navigation

Participants: Adrian Bourgaud, Carlos Di Pietro, Thierry Fraichard, Rémi Paulin, Patrick Reignier, Andre Van Den Berg.

Assistant robots and robot companions are designed to share the human living space, to navigate among and interact with human beings. From the mobility point of view, roboticists have recently striven to develop navigation scheme geared towards achieving so-called “socially acceptable motions”. To that end, various concepts borrowed from environmental psychology and anthropology have been used, the “personal space” concept from Proxemics being perhaps the most widely used. The purpose of our work here is to further the research in this area by taking into account other factors such as human activities, interaction configurations and intentions. An attentional model derived from cognitive psychology is used to dynamically determine the “focus of attention” of the persons involved in a given task. Depending on the task at hand, the robot uses the attention information in order to decide its future course of action so as, for instance, to attract one person’s attention or, on the contrary, to minimize the disturbance caused. In 2013, a paper describing the first results obtained was presented during the Israeli Conf. on Robotics [14].

5.2. Qualitative approaches for building energy management

Participant: Patrick Reignier.

Reducing housing energy costs is a major challenge of the 21st century. In the near future, the main issue for building construction is the thermal insulation, but in the longer term, the issues are those of "renewable energy" (solar, wind, etc.) and "smart buildings". Home automation system basically consists of household appliances linked via a communication network allowing interactions for control purposes. Thanks to this network, a load management mechanism can be carried out: it is called distributed control. An optimal home energy management system is still a goal to aim for, because lots of aspects are still not completely fulfilled. Most of the energy systems respect only the energy needs, but they don’t tackle the user needs or satisfaction. Energy systems also have a lack when it comes to the dynamicity of the environments (the system ability to adapt).The problem is similar for the existing HMI (Human User Interface) of those Home Automation Systems where only experts can understand the data coming from the sensors and most important, the energy plan coming from management system (How? and Why?). The goal of this study is to propose a house energy model that can be both used to predict at some level energy evolution and that can be understood by the end user. The house energy model is based on Fuzzy Cognitive Maps representing cause-effects relations. It is first designed by an expert and then automatically tuned to a particular house using machine learning approaches. Preliminary experiments have been done this year using the Predis Plateform datasets.

5.3. Ikio, a sociable kiosk

Participants: Rémi Barraquand [correspondant], Jiří Pytela, Johan Girod.

In the Personal Assisted Living project we investigate the design of iKio: a sociable kiosk. A simple sketch of the iKio is illustrated in figure 6 . The general idea is to enhance the interaction ability of tablet and smartphone. What motivates the choice of this type of devices is the observation that people have come to treat these gadgets as their own body appendage. As pointed out by the recent study conducted by the Pew Research Center [2], people are starting to use their phones and tablets for more sensitive activities that were almost considered taboo in the past, also these devices are becoming substitute for other traditional devices like photo and video cameras.

[2] http://pewinternet.org/

The design of iKio is therefore influenced by this emerging form of symbiosis and aims to enhance both user-experience and human-technologies interaction. As follow, iKio does not have a fixed body per se, instead it is embodied in a tablet which can be carried along with people in their daily activities but which can also be docked into any mechanical structure that will provide it with enhanced abilities. Using such mechanical structure iKio can express emotion and interact more easily in the physical space of people. The core of iKio is specifically designed to handle and to support ostensive-inferential communication which is characteristic of human communication in contrast with the code model of communication argued to be the main reason of unadapted and autistic interaction between technologies and human. An early prototype of iKio is illustrated in figure 6 . It was constructed using the Bioloid [3] construction kit.

5.4. Limits and performances of embedded RGBD sensors on mobile robots for social interaction

Participants: Amaury Nègre, Dominique Vaufreydaz [correspondant].

While working on sociably acceptable companion robots, we highlighted some problems of embedding RGBD sensors on mobile robots. Performances of our algorithms can be severely decreased by intrinsic parameters of the robot: linear and angle speeds, height and angle of view of the mounted RGBD sensor, etc. We are currently conducting experiments on influence of these parameters on our perception of humans within a home-like environment. As an extra expected results, we will provide to the research community a corpus that can be used as benchmark for several tasks in mobility: 2D and 3D face detection, body and skeleton detection, fall detection and engagement detection.

[3] http://www.robotis.com/xe/bioloid_en

../../../../projets/prima/IMG/ikio_illustration.png

Figure 6. Preliminary sketch of the iKio together with an early prototype, both the 3d model and its realization using the Bioloid Kit

../../../../projets/prima/IMG/RobulabExperiment.png

Figure 7. Robotic platform within the home-like environment for mobile RGBD experiments.

PRIVATICS Team

5. New Results

5.1. Online Social Networks Tracking

Participants: Mohamed Ali Kaafar, Abdelberi Chaabane.

Behavioural advertisement, profiling, adver-gaming and social advertisement illustrate how user personal information and social relations have been integrated to the market model. In other words, user information is *commodified*: the user identity becomes a commodity to be sold and bought. This radical change raises several privacy questions and leads to clamouring for better understanding, regulation and protection of user privacy. Within this context, there is both a long-term and a short-term dimension to our work. For the sort term, I showed that OSN can present a real threat to user privacy as the *full* control of user data – both access and dissemination – is hard to achieve. For the long term, our work calls for a better educational approach to privacy as well as a stricter regulation.

5.2. Behavioural advertisement

Participants: Mohamed Ali Kaafar, Abdelberi Chaabane.

Online Social Networks Tracking. I examined web user tracking capabilities of the three major global OSNs. I studied the mechanisms which enable these services to persistently and accurately follow users web activity, and evaluate to which extent this phenomena is spread across the web. Through a study of the top 10K websites, our findings indicate that OSN tracking is diffused among almost all website categories, independently from the content and the audience. I also evaluated the tracking capabilities in practice and demonstrated – by analysing a real traffic traces – that OSNs can reconstruct a significant portion of users web profile and browsing history. I finally provided insights into the relation between the browsing history characteristics and the OSN tracking potential, highlighting the high risk properties. This work shows that web tracking in combination with personal information from social networks represents a serious privacy violation that shifts the tracking from a virtual tracking (i.e. the user is virtual) to a real "physical" tracking (i.e. based on user personal information).

5.3. Selling Off Privacy at Auction

Participants: Claude Castelluccia, Lukasz Olejnik, Cédric Lauradoux, Minh-Dung Tran.

The first one is a privacy analysis of Real-Time Bidding (RTB) and Cookie Matching (CM). RTB is a technology that allows ad buyers (advertisers) and ad sellers (publishers) to buy and sell ad spaces at real-time auctions through ad exchanges. In RTB, when user visits a publisher page, the ad impression (i.e. one ad display in an ad space) and the user information are immediately broadcast by the ad exchange to a number of bidders (i.e. advertisers or their representatives) for them to bid for the chance to serve ads to this user. CM protocol allows the ad exchange and the bidder to synchronize their cookies of the same user, thus facilitating their exchange of user data.

In [41], we characterize and quantify the potential user web history leakage from ad exchanges to bidders in RTB as a result of exchanging user data. We also discuss and quantify the extent to which companies can potentially collude to increase their tracked user profiles using CM. In addition, we leverage a design characteristic of RTB to observe the winning price of each RTB auction. By analyzing these prices, we show how advertisers evaluate the value of user privacy. This work (titled Selling Off Privacy at Auction) will be presented in NDSS 2014, San Diego, USA in February, 2014.

5.4. Wi-Fi and privacy

Participants: Cédric Lauradoux, Mathieu Cunche, Levent Demir.

Active service discovery in Wi-Fi involves wireless stations broadcasting their Wi-Fi fingerprint, i.e. the SSIDs of their preferred wireless networks. The content of those Wi-Fi fingerprints can reveal different types of information about the owner. In [5], we focus on the relation between the fingerprints and the links between the owners. Our hypothesis is that social links between devices owners can be identified by exploiting the information contained in the fingerprint. More specifically we propose to consider the similarity between fingerprints as a metric, with the underlying idea: similar fingerprints are likely to be linked. We have studied the performances of several similarity metrics on a controlled dataset and then apply the designed classifier to a dataset collected in the wild. Our study is based on a dataset collected in Sydney, Australia, composed of fingerprints belonging to more than 8000 devices.

Extending this problem, we present a set of attacks that allow an attacker to link a Wi-Fi device to its owner identity. We present two methods that, given an individual of interest, allow identifying the MAC address of its Wi-Fi enabled portable device. Those methods do not require a physical access to the device and can be performed remotely, reducing the risks of being noticed. We present in [4], [35] scenarios in which the knowledge of an individual MAC address could be used for mischief.

5.5. Sensor security and privacy

Participants: Claude Castelluccia, Marine Minier, Cédric Lauradoux, Mathieu Cunche.

Wireless sensor networks (WSNs) are composed of a large number of low-cost, low-power, and multi-functional sensor nodes that communicate at short distance through wireless links. They are usually deployed in an open and uncontrolled environment where attackers may be present. Due to the use of low-cost materials, hardware components are not tamper-resistant and an adversary could access to a sensor's internal state.

In [7], we consider packet pollution attack. Packet pollution attack is considered as the most threatening attack model against network coding based sensor networks. A widely held belief says that, in a single source multi-destination dissemination scenario, the total number of polluted packets in the network will grow with the length of the transmission path, and the decoding failure (DF) rate at the further destination nodes are relatively lower. In this work, we first obtain an opposite result by analyzing the pollution attack in multicast scenarios, and find out a convergence trend of pollution attack by network coding system, and quantify the network resiliency against the pollution attacks which happen at any place along the source-destination paths. Then, the analysis result is proved by our simulations on two most widely deployed buffer strategies, Random-In Random-Out (RIRO) and First-in First-Out (FIFO). Finally, it is proved that RIRO has a much advanced security feature than FIFO in constraining the pollution attack gradually, and almost vanished in the end.

An adversary can easily capture even a single node and inserts duplicated nodes at any location in the network. If no specific detection mechanisms are established, the attacker could lead many insidious attacks such as subverting data aggregation protocols by injecting false data, revoking legitimate nodes and disconnecting the network if the replicated nodes are judiciously placed in the network. In [8], we first introduce the algorithm already published in PIMRC 2009 that describes a new hierarchical distributed algorithm for detecting node replication attacks using a Bloom filter mechanism and a cluster head selection. This mechanism could be efficiently used in a WSN as soon as the network is built with a clustering algorithm creating a three tiers hierarchy. We extend the results of our previous results by a theoretical discussion on the bounds of our algorithm. We also perform extensive simulations of our algorithm for random topologies and we compare those results with other proposals of the literature. Finally we show the effectiveness of our algorithm and its energy efficiency.

Finding entropy sources is a major issue to design non-deterministic random generators for headless devices. Our goal in [22] is to evaluate a collection of sensors (e.g. thermometer, accelerometer, magnetometer) as potential sources of entropy. A challenge in the analysis of these sources is the estimation of min-entropy. We have followed the NIST recommendations to obtain pessimistic estimations from the dataset collected during our campaign of experiments. The most interesting sensors of our study are: the accelerometer, the magnetometer, the vibration sensor and the internal clock. Contrary to previous results, we observe far less entropy than it was expected before. Other sensors which measures phenomena with high inertia such as the temperature or air pressure provide very little entropy.

In [12], we propose a key certification protocol for wireless sensor networks that allows nodes to autonomously exchange their public keys and verify their authenticity using one-way accumulators. We examine and compare different accumulator implementations for our pro- tocol on the Sun SPOT platform. We observe that our protocol performs best with accumulators based on Elliptic Curve Cryptography (ECC): ECC-based accumulators have roughly the same speed as Secure Bloom filters, but they have a smaller memory footprint.

5.6. Buidling blocks

Participant: Marine Minier.

In [17], we develop a complete library of lightweight block ciphers dedicated to security applications in wireless sensor networks (WSNs). Choosing best algorithms in terms of energy-efficiency and of small memory requirements is a real challenge because the sensor networks must be autonomous. We study on a dedicated platform of sensors most of the recent lightweight block ciphers as well as some conventional block ciphers. First, we describe the design of the chosen block ciphers with a security summary and we then present some implementation tests performed on our platform. The library is available online: http://bloc.project.citi-lab.fr/library.html.

In [23], we present two related key impossible differential attacks against 14 rounds of Piccolo-80 and 21 rounds of Piccolo-128 without the whitening layers. Piccolo is a new lightweight block cipher proposed by SONY at CHES 2011. The attack against Piccolo-80 has a time and data complexity of $2^{68.19}$ whereas the time/data complexity of the attack against Piccolo-128 is $2^{117.77}$.

While Generalized Feistel Networks have been widely studied in the literature as a building block of a block cipher, we propose in [13] a unified vision to easily represent them through a matrix representation. We then propose a new class of such schemes called Extended Generalized Feistel Networks well suited for cryptographic applications. We instantiate those proposals into two particular constructions and we finally analyze their security.

We also obtain, in [24] a result concerning an integral distinguisher on the SHA-3 finalist Grøstl-512 v3.

5.7. Formal and legal issues of privacy

Participants: Thibaud Antignac, Denis Butin, Daniel Le Métayer.

- **Privacy by design** The privacy by design approach is often praised by lawyers as well as computer scientists as an essential step towards a better privacy protection. The general philosophy of privacy by design is that privacy should not be treated as an afterthought but rather as a first-class requirement during the design of a system. The approach has been applied in different areas such as smart metering, electronic traffic pricing, ubiquitous computing or location based services. More generally, it is possible to identify a number of core principles that are widely accepted and can form a basis for privacy by design. For example, the Organization for Economic Co-operation and Development (OECD) has put forward principles such as the consent, limitation of use, data quality, security and accountability. One must admit however that the take-up of privacy by design in the industry is still rather limited. This situation is partly due to legal and economic reasons: as long as the law does not impose binding commitments, ICT providers and data collectors do not have sufficient incentives to invest into privacy by design. The situation on the legal side might change in Europe though because the regulation proposed by the European Commission in January 2012 (to replace the European Directive 95/46/EC), which is currently under discussion, includes binding commitments on privacy by design.

  But the reasons for the lack of adoption of privacy by design are not only legal and economic: even though computer scientists have devised a wide range of privacy enhancing tools, no general methodology is available to integrate them in a consistent way to meet a set of privacy requirements. The next challenge in this area is thus to go beyond individual cases and to establish sound foundations and methodologies for privacy by design. As a first step in this direction, we have focused on the data minimization principle which stipulates that the collection should be limited to the pieces of data strictly necessary for the purpose, and we have proposed a framework to reason about the choices of architecture and their impact in terms of privacy. The first strategic choices are the allocation of the computation tasks to the nodes of the architecture and the types of communications between the nodes. For example, data can be encrypted or hashed, either to protect their confidentiality or to provide guarantees with respect to their correctness or origin. The main benefit of a centralized architecture for the "central" actor is that he can trust the result because he keeps full control over its computation. However, the loss of control by a single actor in decentralized architectures can be offset by extra requirements ensuring that errors (or frauds) can be detected *a posteriori*. In order to help the designer grasp the combination of possible options, our framework provides means to express the parameters to be taken into account (the service to be performed, the actors involved, their respective requirements, etc.) and an inference system to derive properties such as the possibility for an actor to detect potential errors (or frauds) in the computation of a variable. This inference system can be used in the design phase to check if an architecture meets the requirements of the parties or to point out conflicting requirements.

- **Accountability**

  The principle of accountability, which was introduced three decades ago in the OECD guidelines, has been enjoying growing popularity over the last few years as a solution to mitigate the loss of control by increasing transparency of data processing. At the European level, the Article 29 Working Group published an opinion dedicated to the matter two years ago and the principle is expected to be enshrined in the upcoming European data protection regulation. But the term "accountability" is used with different meanings by different actors and the principle itself has been questioned by some authors as providing deceptive protections and also possibly introducing new risks in terms of privacy. We have studied the different interpretations of the notion of accountability following a multidisciplinary approach and we have argued that *strong accountability* should be a cornerstone of future data protection regulations. By *strong accountability* we mean a principle of accountability which

  - applies not only to policies and procedures, but also to practices, thus providing means to oversee the effective processing of the personal data, not only the promises of the data controller and its organisational measures to meet them;
  - is supported by precise binding commitments enshrined in law;
  - involves audits by independent entities.

  Strong accountability should benefit all stakeholders: data subjects, data controllers, and even data protection authorities whose workload should be considerably streamlined.

  But accountability is a requirement to be taken into account from the initial design phase of a system because of its strong impact on the implementation of the log architecture. Using real-world scenarios, we have shown that decisions about log architectures are actually nontrivial. We have addressed the question of what information should be included in logs to make their a posteriori compliance analysis meaningful. We have shown how log content choices and accountability definitions mutually affect each other and incites service providers to rethink up to what extent they can be held responsible. These different aspects are synthesized into guidelines to avoid common pitfalls in accountable log design. This analysis is based on case studies performed on our implementation of the PPL policy language.

- **Verification of privacy properties** The increasing official use of security protocols for electronic voting deepens the need for their trustworthiness, hence for their formal verification. The impossibility of linking a voter to her vote, often called voter privacy or ballot secrecy, is the core property of many such protocols. Most existing work relies on equivalence statements in cryptographic extensions of process calculi. We have proposed the first theorem-proving based verification of voter privacy which overcomes some of the limitations inherent to process calculi-based analysis. Unlinkability between two pieces of information is specified as an extension to the Inductive Method for security protocol verification in Isabelle/HOL. New message operators for association extraction and synthesis are defined. Proving voter privacy demanded substantial effort and provided novel insights into both electronic voting protocols themselves and the analysed security goals. The central proof elements have been shown to be reusable for different protocols with minimal interaction.

- **Privacy and discrimination**

  The interactions between personal data protection, privacy and protection against discriminations are increasingly numerous and complex. For example, there is no doubt that misuses of personal data can adversely affect privacy and self-development (for example, resulting in the unwanted disclosure of personal data to third parties, in identity theft, or harassment through email or phone calls), or lead to a loss of choices or opportunities (for example, enabling a recruiter to obtain information over the Internet about political opinions or religious beliefs of a candidate and to use this information against him). It could even be suggested that privacy breaches and discriminations based on data processing are probably the two most frequent and the most serious types of consequences of personal data breaches. We have studied these interactions from a multidisciplinary (legal and technical) perspective and argued that an extended application of the application of non-discrimination regulations could help strengthening data protection. We have have analysed and compared personal data protection, privacy and protection against discriminations considering both the types of data concerned and the *modus operandi* (*a priori* versus *a posteriori* controls, actors in charge of the control, etc.). From this comparison, we have drawn some conclusions with respect to their relative effectiveness and argued that *a posteriori* controls on the use of personal data should be strengthened and the victims of data misuse should get compensations which are significant enough to represent a deterrence for data controllers. We have also advocated the establishment of stronger connections between anti-discrimination and data protection laws, in particular to ensure that any data processing leading to unfair differences of treatments between individuals is prohibited and can be effectively punished.

PROSECCO Project-Team

6. New Results

6.1. Verification of Security Protocols with Lists in the Symbolic Model

Participants: Bruno Blanchet, Miriam Paiola.

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 in this model.

We designed a novel automatic technique for proving secrecy and authentication properties for security protocols that manipulate lists of unbounded length, for an unbounded number of sessions. This result is achieved by extending the Horn clause approach of the automatic protocol verifier ProVerif. We extended the Horn clauses to be able to represent lists of unbounded length. We adapted the resolution algorithm to handle the new class of Horn clauses, and proved the soundness of this new algorithm. We have implemented our algorithm and successfully tested it on several protocol examples, including XML protocols coming from web services. This work has been published in [22] and our prototype is available at http://prosecco.inria.fr/personal/bblanche/publications/BlanchetPaiolaCCS13.html.

Last year, we published a conference paper that shows that, for a limited class of protocols, if a protocol is proven secure by ProVerif with lists of length one, then it is secure for lists of unbounded length. A journal version [50] of this paper has now been accepted.

6.2. Generation of Implementations Proved Secure in the Computational model

Participants: Bruno Blanchet, David Cadé.

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 continued working on our compiler from CRYPTOVERIF speficications to OCaml. Using CryptoVerif and this compiler, we can prove security properties of specifications of protocols in the computational model and generate runnable implementations from such proved specifications. We have published a journal paper on our implementation of SSH generated using this compiler [13] and a proof that this compiler preserves security [23], and we have submitted a journal version of this proof. David Cadé also defended his PhD thesis on this topic [44].

6.3. Computationally Complete Symbolic Attacker and Key Exchange

Participants: Gergely Bana [correspondant], Koji Hasebe, Mitsuhiro Okada.

Around year 2000, various research groups started looking into the relevance of symbolic verification techniques to computational security. If a symbolic verification technique results computational guarantees, we say that computational soundness holds. One of the major concerns has been that the usual Dolev-Yao symbolic attacker that automated symbolic tools used exclusively (to search for attacks) at that time did not seem to allow satisfactory soundness results, only with serious limitations. One possible promising approach to overcome this problem is to derive security guarantees directly as CryptoVerif or F7 does. As an alternative approach, in 2012, Bana and Comon-Lundh introduced a notion they called computationally complete symbolic attacker. With this technique, elimination of the possibility of a computational attack would also mean that computational attack does not exist without the limitations that the Dolev-Yao technique required. Their symbolic attacker can do everything that is not forbidden by conditions derived from standard computational assumptions on the primitives. In this current work, based on predicates for "key compromise", we provided such conditions to handle secure encryption even keys are allowed to be sent. We examined both IND-CCA2 and KDM-CCA2 encryptions, both symmetric and asymmetric situations as well as INT-CTXT encryptions. We verified (by hand) a number of protocols as the symmetric Needham-Schroeder protocol, Otway-Rees protocol, Needham-Schroeder-Lowe protocol. Furthermore, we also made some improvements in the computational semantics, and have established a relationship between the computational semantics of Bana and Comon-Lundh and Fitting's embedding of classical logic into S4. This work was published at CCS'13 [19].

6.4. Formal Models and Concrete Attacks on Web Applications

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

Modern web applications are built as a combination of mostly static servers that host user data and highly dynamic client-side applications that process and present the data to the user. These client-side applications may be hosted as JavaScript within a browser or within custom applications written, say, for smartphones. Hence, in addition to traditional server-side mechanisms, the security of these applications increasingly depends on the correct use of browser-based security mechanisms, client-side access control, and cryptography. These mechanisms are often new, ad hoc, and deserving of close analysis.

Our approach is to formally model various client- and server-side security mechanisms for web applications and rigorously analyze their real-world deployments. When our formal analyses find attacks, we test them against example web applications, report vulnerabilities to various vendors, design countermeasures, and use automated security protocol analysis tools formally verify that our countermeasure resists a large class of attacks. This year, we published three papers in this area. At ESSoS, we formally modeled the authorization policies of common Android apps, found new attacks, and proposed a verified authorization framework [27]. At POST, we formally modeled various cloud-based encrypted storage applications and found both cryptographic and web attacks on them, resulting in patches to these websites and novel countermeasures [20]. At Usenix Security, we proposed a new, safer language for security-critical web components [25]. Defensive JavaScript is a subset of JavaScript that guarantees isolation from other (potentially untrusted) scripts on the same page. This enables, for the first time, the design of cryptographic and single sign-on components that can be formally guaranteed to preserve its secrets even if the hosting website is subject to a cross-site scripting attack.

6.5. Attacks and Proofs for TLS Implementations

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

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 present the main published results below, other papers have been submitted for review.

We have developed a verified reference implementation of TLS 1.2, called miTLS. Our code fully supports its wire formats, ciphersuites, sessions and connections, re-handshakes and resumptions, alerts and errors, and data fragmen- tation, 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 was published at IEEE S&P 2013 [21].

In parallel with this long-term constructive project, we have been analyzing the use of TLS in existing web applications, and our analyses uncovered a number of attacks, leading to patched in popular browsers like Chrome, Internet Explorer, and Firefox, as well as websites like Google and Akamai.

One of these classes of attacks was published at WOOT'13 [29]. In this paper, we identify logical web application flaws which can be exploited by TLS truncation attacks to desynchronize the user- and server-perspective of an application's state. It follows immediately that servers may make false assumptions about users, hence, the flaw constitutes a security vulnerability. Moreover, in the context of authentication systems, we exploit the vulnerability to launch the following practical attacks: we exploit the Helios electronic voting system to cast votes on behalf of honest voters, take full control of Microsoft Live accounts, and gain temporary access to Google accounts.

RAP Project-Team

4. New Results

4.1. Algorithms: Bandwidth Allocation in Optical Networks

Participants: Christine Fricker, Jelena Pesic, 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. 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).

Our work on access networks proposed an original dynamic bandwidth allocation (DBA) algorithm and demonstrated its excellent performance. This DBA algorithm was then adapted to a meshed metropolitan network based on TWIN and implementing flow-aware resource sharing. Extensions using a concept called “multipath” were shown to offer an energy efficient solution for wide area networks.

In 2013, we contributed to the Celtic Plus project called SASER/SAVENET. 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 responsibility for the work package (where alternative optical network architectures are also evaluated) is with Alcatel-Lucent Bell Labs.

In 2013, Inria edited the M12 milestone document of Task 6.4 "TWIN implementations and preliminary MAC protocol specifications". A paper on applying the network architecture and MAC/DBA protocols proposed by the team to the domain of data center interconnects has been submitted.

RAP has continued to work on a two-year research contract with Orange Labs on further developing the multipath architecture (20012-2013). The main contribution in 2013 has been to propose the use of tunable receivers in addition to tunable transmitters. This technological evolution is possible with recent developments in coherent transmission and offers greater flexibility and enhanced efficiency. Work is continuing on evaluating this architecture by simulation (using Onmet++) and by analytical modelling.

4.2. Algorithms: Content-Centric Networking

Participants: Christine Fricker, Philippe Robert, James Roberts, Nada Sbihi.

RAP participated in an ANR project named CONNECT which contributed to the definition and evaluation of a new paradigm for the future Internet: an information-centric network (ICN) where, rather than interconnecting remote hosts like IP, the network directly manages the information objects that users publish, retrieve and exchange. The project ended in December 2012 but we have continued to work on information-centric networking in 2013.

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. 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. The excellent accuracy of this method over a wide range of practically relevant traffic models has been explained mathematically. 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.

This year we completed the analysis of a star network topology with multiple nodes. Several scalings were used to describe the fluid limit behaviour.

4.4. 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.

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.

This year we have investigated a mathematical model of the production of proteins in prokaryotic cells. Up to now most of the mathematical used to study these problems concern the production of *one* fixed class of proteins. When several classes of proteins are considered, each class requires in fact a fraction of the common and limited resources of the cell. One has therefore to understand how the allocation of the resources within the cell is done. Due to the fact that the cytoplasm of the cell is a quite disorganized medium where the components of the cell move, the whole production process has an important stochastic component. A model describing the allocation of the ribosomes of the cell to produce proteins is investigated via a Markovian representation. Asymptotic results for the equilibrium and for the transient behavior have been obtained under a scaling procedure and a reasonable biological assumption of saturation, i.e. when resources of the cell are tight. The equilibrium and the transient behavior have been investigated, it has been shown in particullar that, in the limit, the number of free ribosomes converges in distribution to a Poisson distribution whose parameter satisfies a fixed point equation.

4.5. Stochastic networks: large bike sharing systems

Participants: Christine Fricker, Hanène Mohamed, Nicolas Servel.

This is a collaboration with Nicolas Gast (EPFL). Bike sharing systems were launched by numerous cities to be a urban mode of transportation, for example Velib in Paris. One of the major issues is the availability of the resources: bikes or free slots to return the bikes. These systems became a hot topic in Operation Research and now the importance of stochasticity of such system behavior is commonly admitted. The problem is to understand their behavior and how to manage them in order to provide both resources to users.

Our model is the first one taking into account the finite number of spots at the stations. In a homogeneous model, mean field limit theorems give the dynamic of a large system. Analytical results are obtained and convergence proved in a standard model via Lyapunov functions. It allows to find the best ratio of bikes per station and to measure the improvement of incentive mechanisms, as choosing among two stations for example. We investigate also redistribution of bikes by trucks. Further results deal with heterogeneous system. By mean field techniques, analytical results were recently obtained on systems consisting in several clusters. In a work with Nicolas Servel, we discuss the improvement of choosing between two stations in the same cluster. Our goal is to propose, via a theoretical study and tests, simple algorithms to improve the system behavior.

With Hanene Mohamed, we study the problem of impact of geometry on incentive mechanisms. Our first model under investigation is very close from the Gates-Westcott crystal growth model with its underlying random deposition process.

4.6. Random Graphs

Participants: Nicolas Broutin, Henning Sulzbach.

4.6.1. 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) 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.

4.6.2. 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.

In a stochastic context, the minimum spanning tree is tigthly 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 permits to describe precisely the metric space scaling limit of a random minimum spanning tree, which identifies a novel continuum random tree which is truely different from that of Aldous.

4.6.3. Analysis of recursive partitions

This is joint work with R. Neininger (Frankfurt)

The techniques that we developped in order to estimate the cost of partial match queries in random quad trees have been used to solve an open question about the recursive lamination of the disk. We have proved that the planar dual of the lamination, which is a tree, converges almost surely when suitably rescaled to a compact random tree encoded by a continuous function. We also pinned down the fractal dimension of the limit object.

REALOPT Project-Team

6. New Results

6.1. Extending the column generation paradigm

Building on our technical review [89] of methods for solving the Lagrangian Dual (with an analysis of the scope for hybridization) we have worked on methodologies that can be understood as an extension of the column generation approach in [22]. Working in an extended variable space allows one to develop tighter reformulations for mixed integer programs. To handle the size of the extended formulation, one can work with inner approximations defined and improved by generating dynamically variables and constraints. This so-called "column-and-row generation" procedure is revisited here in a unifying presentation that generalizes the column generation algorithm and extends to the case of working with an approximate extended formulation. A key benefit of this approach is that lifting pricing problem solutions in the space of the extended formulation permits their recombination into new subproblem solutions and results in faster convergence. The interest of the approach is evaluated numerically on machine scheduling, bin packing, generalized assignment, and multi-echelon lot-sizing problems. We compare a direct handling of the extended formulation, a standard column generation approach, and the “column-and-row generation” procedure. Within the Samba project we further showed that this stabilization offered by the recombination of solutions is complementary and adds up to stabilization techniques based on smoothing that were developed within Samba. These techniques have been applied in [26], [29].

6.2. Interior point cutting plane strategy revisited for column generation

In [89], we identify what are the stabilization features that are built into variants of subgradient algorithms and polyhedral approaches. In [27], we further compare their theoretical performance and discuss their combination. Stabilization procedures for column generation can be viewed as cutting plane strategies in the dual. Exploiting the link between in-out separation strategies and dual price smoothing techniques for column generation, we derive a generic bound convergence property for algorithms using a smoothing feature. Such property adds to existing in-out asymptotic convergence results. In our study on In-Out Separation and Column Generation Stabilization by Dual Price Smoothing, we note that our convergence property adds to existing in-out asymptotic convergence results. Beyond theoretically convergence, we describe in [88] a proposal for effective finite convergence in practice and we develop a smoothing auto-regulating strategy that makes the need for parameter tuning obsolete. Practical speed-up convergence that are observed go from 20% to 500 $\%$. These contributions turn stabilization by smoothing into a general purpose practical scheme that can be used into a generic column generation procedure. We conclude the paper by showing that the approach can be combined with an ascent method, leading to improved performances. Such combination might inspire novel cut separation strategies.

6.3. A MILP approach to minimize the number of late jobs with and without machine availability constraints

The study in [13] investigates scheduling problems that occur when the weighted number of late jobs that are subject to deterministic machine availability constraints have to be minimized. These problems can be modeled as a more general job selection problem. Cases with resumable, non-resumable, and semi-resumable jobs as well as cases without availability constraints are investigated. The proposed efficient mixed integer linear programming approach includes possible improvements to the model, notably specialized lifted knapsack cover cuts. The method proves to be competitive compared with existing dedicated methods: numerical experiments on randomly generated instances show that all 350-job instances of the test bed are closed for the well-known problem $1|r_i|\sum w_i U_i$. For all investigated problem types, 98.4% of 500-job instances can be solved to optimality within one hour.

6.4. Multidimensional dual-feasible functions

Dual-feasible functions have been used in the past to compute lower bounds and valid inequalities for different combinatorial optimization and integer programming problems. Until now, all the dual-feasible functions proposed in the literature were 1-dimensional functions. In [11] we extended the principles of dual-feasible functions to the m-dimensional case by introducing the concept of vector packing dual-feasible function. We explored the theoretical properties of these functions in depth, and we proposed general schemes for generating some instances of these functions. Additionally, we proposed and analyzed different new families of vector packing dual-feasible functions. All the proposed approaches were tested extensively using benchmark instances of the 2-dimensional vector packing problem. Our computational results showed that these functions can approximate very efficiently the best lower bounds for this problem.

6.5. New branch-and-price methods for variants of bin packing problems

We proposed branch-and-price methods for two variants of the well-known bin-packing problem.

The bin packing problem with conflicts consists in packing items in a minimum number of bins of limited capacity while avoiding joint assignments of items that are in con ict. The study in [21] demonstrates that a generic implementation of a Branch-and-Price algorithm using specific pricing oracle yields comparatively good performance for this problem. We use our black-box Branch-and-Price solver BaPCod, relying on its generic branching scheme and primal heuristics. We developed a dynamic programming algorithm for pricing when the conflict graph is an interval graph, and a depth-first-search branch-and-bound approach for pricing when the conflict graph has no special structure. The exact method was tested on instances from the literature where the conflict graph is an interval graph, as well as harder instances that we generated with an arbitrarily conflict graph and larger number of items per bin. Our computational experiment report sets new benchmark results for this problem, closing all open instances of the literature in one hour of CPU time.

In the bin-packing with fragile objects, we are given a set of objects, each characterized by a weight and a fragility, and a large number of uncapacitated bins. Our aim is to find the minimum number of bins needed to pack all objects, in such a way that in each bin the sum of the object weights is less than or equal to the smallest fragility of an object in the bin. The problem is known in the literature as the Bin Packing Problem with Fragile Objects, and appears in the telecommunication field, when one has to assign cellular calls to available channels by ensuring that the total noise in a channel does not exceed the noise acceptance limit of a call. In [10], we propose a branch-and-bound and several branch-and-price algorithms for the exact solution of the problem, and improve their performance by the use of lower bounds and tailored optimization techniques. In addition we also develop algorithms for the optimal solution of the related knapsack problem with fragile objects. We conduct an extensive computational evaluation on the benchmark set of instances, and show that the proposed algorithms perform very well.

6.6. Freight railcar routing

In some countries, the activities of managing railroads and managing a fleet of freight railcars are separated by a law. A state-owned company is in charge of the first activity. The control of freight railcars is separated between several independent companies. The main objective of such company is an effective management of its railcars. As these companies are commercial, the goal is to maximize the profit from the usage of their railcars. The profit of a company is mainly determined by the difference between the total gain it receives from satisfying requests for delivery of goods in railcars and the costs it pays to the state-owned company for exploiting the railroad network.

Consequently, the main optimization problem that every freight railcar management company faces can be formulated as follows. We need 1) to choose a set of transportation demands between stations in a railroad network, and 2) to fulfill these demands by appropriately routing the set of available railcars, while maximizing the total profit. We formulate this problem as a multi-commodity flow problem in a large space-time graph. Three approaches are proposed to solve the Linear Programming relaxation of this formulation: direct solution by an LP solver, a column generation approach based on the path reformulation, and a “column generation for extended formulations" approach [22]. In the latter, the multi-commodity flow formulation is solved iteratively by dynamic generation of arc flow variables. Three approaches have been tested on a set of real-life instances provided by one of the largest freight rail transportation companies in Russia. Instances with up to 10 millions of arc flow variables were solved within minutes of computational time [29], [39].

6.7. Reliable Service Allocation in Clouds with Memory and Capacity Constraints

In [25] we consider allocation problems that arise in the context of service allocation in Clouds. More specifically, on the one part we assume that each Physical Machine (denoted as PM) is offering resources (memory, CPU, disk, network). On the other part, we assume that each application in the IaaS Cloud comes as a set of services running as Virtual Machines (VMs) on top of the set of PMs. In turn, each service requires a given quantity of each resource on each machine where it runs (memory footprint, CPU, disk, network). Moreover, there exists a Service Level Agreement (SLA) between the Cloud provider and the client that can be expressed as follows: the client requires a minimal number of service instances which must be alive at the end of the day, with a given reliability (that can be converted into penalties paid by the provider). In this context, the goal for the Cloud provider is to find an allocation of VMs onto PMs so as to satisfy, at minimal cost, both capacity and reliability constraints for each service. In this paper, we propose a simple model for reliability constraints and we prove that it is possible to derive efficient heuristics.

6.8. On the Theta number of powers of cycle graphs

In [17] we give a closed formula for Lovász's theta number of the powers of cycle graphs C_k^d and of their complements, the circular complete graphs $K_{k/d}$. As a consequence, we establish that the circular chromatic number of a circular perfect graph is computable in polynomial time. We also derive an asymptotic estimate for the theta number of C_k^d.

6.9. Strong chromatic index of planar graphs with large girth

Let Δ be an integer. In [18], we prove that every planar graph with maximum degree Δ and girth at least $10\Delta + 46$ is strong $(2\Delta - 1)$-edge-colorable, that is best possible (in terms of number of colors) as soon as G contains two adjacent vertices of degree Δ. This improves the best previous result when $\Delta \geq 6$.

6.10. Computing clique and chromatic number of circular-perfect graphs in polynomial time

A main result of combinatorial optimization is that clique and chromatic number of a perfect graph are computable in polynomial time (Grötschel et al. in Combinatorica 1(2):169–197,1981). Perfect graphs have the key property that clique and chromatic number coincide for all induced subgraphs; in [19] we address the question whether the algorithmic results for perfect graphs can be extended to graph classes where the chromatic number of all members is bounded by the clique number plus one. We consider a well-studied superclass of perfect graphs satisfying this property, the circular-perfect graphs, and show that for such graphs both clique and chromatic number are computable in polynomial time as well. In addition, we discuss the polynomial time computability of further graph parameters for certain subclasses of circular-perfect graphs. All the results strongly rely upon Lovász's Theta function.

6.11. Computing the clique number of a-perfect graphs in polynomial time

A main result of combinatorial optimization is that clique and chromatic number of a perfect graph are computable in polynomial time (Grötschel, Lovasz and Schrijver 1981). This result relies on polyhedral characterizations of perfect graphs involving the stable set polytope of the graph, a linear relaxation defined by clique constraints, and a semi-definite relaxation, the Theta-body of the graph. A natural question is whether the algorithmic results for perfect graphs can be extended to graph classes with similar polyhedral properties.

In [20] we consider a superclass of perfect graphs, the a-perfect graphs, whose stable set polytope is given by constraints associated with generalized cliques. We show that for such graphs the clique number can be computed in polynomial time as well. The result strongly relies upon Fulkersons's antiblocking theory for polyhedra and Lovasz's Theta function.

REGAL Project-Team

5. New Results

5.1. Introduction

In 2013, we focused our research on the following areas:

- *Distributed algorithms for dynamic and large networks.*
- *Management of distributed data.*
- *Performance and robustness of Systems Software in multicore architectures.*

5.2. Distributed algorithms for dynamic networks

Participants: Luciana Bezerra Arantes [correspondent], Rudyar Cortes, Guthemberg Da Silva Silvestre, Raluca Diaconu, Ruijing Hu, Anissa Lamani, Jonathan Lejeune, Olivier Marin, Sébastien Monnet, Franck Petit [correspondent], Karine Pires, Maria Potop-Butucaru, Pierre Sens, Véronique Simon, Julien Sopena.

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.

5.2.1. 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 [46], [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. In 2013, we studied Ω, the eventual leader election failure detector. Ω ensures that, eventually, each process in the system will be provided by an unique leader, elected among the set of correct processes in spite of crashes and uncertainties. It is known to be weakest failure detector to solve agreement protocols such as Paxos. Then, a number of eventual leader election protocols were suggested. Nonetheless, as far as we are aware of, no one of these protocols tolerates a free pattern of node mobility. In [27] we propose a new protocol for this scenario of dynamic and mobile unknown networks.

[2] Recent work by Leners et al published in SOSP 2011 uses our DSN 2003 paper as basis for performance comparison

5.2.2. 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. 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 and gathering.

5.2.3. 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. In 2013, we have proposed a new distributed algorithm suitable for scale-free random topologies which model some complex real world networks [37], [52].

5.2.4. Peer certification.

In a distributed system, the certification of transactions makes it possible to circumscribe malicious behaviors. Certification requires the use of a trusted third party which must be centralized to guarantee safety. At a large scale, however, centralized certification represents a bottleneck and a single point of attack or failure.

We proposed two decentralized approaches towards certifying transactions with a high probability of success. The first approach replicates transactions over multiple peers and retains identical results from a qualified majority to certify that a service has been carried out for a given client at a given time [30]. The second approach uses distributed reputations to identify trusted nodes and use them as game referees to detect and prevent cheating [57].

5.3. Management of distributed data

Participants: Pierpaolo Cincilla, Guthemberg Da Silva Silvestre, Raluca Diaconu, Jonathan Lejeune, Mesaac Makpangou, Olivier Marin, Sébastien Monnet, Dang Nhan Nguyen, Burcu Külahçioglu Özkan, Karine Pires, Masoud Saeida Ardekani, Thomas Preud'Homme, Pierre Sens, Marc Shapiro, Véronique Simon, Julien Sopena, Gaël Thomas, Mathieu Valero, Mudit Verma, Marek Zawirski.

Storing and sharing information is one of the major reasons for the use of large-scale distributed computer systems. Replicating data at multiple locations ensures that the information persists despite the occurrence of faults, and improves application performance by bringing data close to its point of use, enabling parallel reads, and balancing load. This raises numerous issues:

- where to store or replicate the data, in order to ensure that it is available quickly and remains persistent despite failures and disconnections;
- how many copies are needed to face dynamically-changing demand (load) and offer (elasticity);
- how to parallelize writes and hence how to ensure consistency between replicas;
- tradeoffs between synchronised, consistent but slow updates, and fast but weakly-consistent ones;
- when and how to move data to computation, or computation to data, in order to improve response time while minimizing storage or energy usage;
- etc.

[3] C. Boulinier, F. Petit, and V. Villain. Synchronous vs. asynchronous unison. Algorithmica, 51(1):61-80, 2008

5.3.1. Long term durability

To tolerate failures, distributed storage systems replicate data. However, despite the replication, pieces of data may be lost (i.e. all the copies are lost). We have previously proposed a mechanism, RelaxDHT, to make distributed hash tables (DHT) resilient to high churn rates.

Well sized systems rarely loose data, still, data may be lost: the more the time passes, the greater is the risk of loss. It is thus necessary to study data durability on a long term. To do so, we have implemented an efficient simulator, we can simulate a 100 node system over years within several hours. We have observe that a given system with a given replication mechanism can store a certain amount of data above which the loss rate would be greater than an “acceptable”/fixed threshold. This amount of data can be used as a metric to compare replication strategies. We have studied the impact of the data distribution layout upon the loss rate. The way the replication mechanism distribute the data copies among the nodes has a great impact. If node contents are very correlated, the number of available sources to heal a failure is low. On the opposite, if the data copies are shuffled among the nodes, many source nodes may be available to heal the system, and thus, the system losses less pieces of data. We are also studying the impact of other parameters, like the replication degree or the way a new storer node is chosen.

5.3.2. Adaptative replication

Different pieces of data have different popularity: some data are stored but never accessed while other pieces are very “hot” and are requested concurrently by many clients. This implies that different pieces of data with different popularity should have a different number of copies to efficiently serve the requests without wasting resources. Furthermore, for a given piece of data, the popularity may vary drastically among time. It is thus important that the replication mechanism dynamically adapt the number of replicas to the demand. In the context of the ODISEA2 FUI project, we have made two main contributions. First, we have studied the popularity distribution and evolution of live video streams (Karine Pires thesis). Second, we have designed replication mechanisms able to gracefully adapt the replication degree to the demand, one based on bandwidth reservation, and one using statistical learning (Guthemberg Silvestre thesis).

5.3.3. 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: under standard assumptions (data store accesses are not known in advance, and transactions may access arbitrary objects in the data store), it is impossible to have both SI and GPR. Our impossibility result is based on a novel decomposition of SI which proves that, like serializability, SI is expressible on plain histories. These results are published at the Euro-Par conference [42].

We designed an efficient protocol that maintains side-steps this impossibility but maintains the most important features of SI:

1. (Genuine Partial Replication) only replicas updated by a transaction T make steps to execute T;
2. (Wait-Free Queries) a read-only transaction never waits for concurrent transactions and always commits;
3. (Minimal Commit Synchronization) two transactions synchronize with each other only if their writes conflict.

The protocol also ensures Forward Freshness, i.e., that a transaction may read object versions committed after it started.

Non-Monotonic Snapshot Isolation (NMSI) is the first strong consistency criterion to allow implementations with all four properties. We also present a practical implementation of NMSI called Jessy, which we compare experimentally against a number of well-known criteria. Our measurements show that the latency and throughput of NMSI are comparable to the weakest criterion, read-committed, and between two to fourteen times faster than well-known strong consistencies. This was published in the Symp. on Reliable Distr. Sys. (SRDS) [43].

5.3.4. Distributed Transaction Scheduling

Parallel transactions in distributed DBs incur high overhead for concurrency control and aborts. Our Gargamel system proposes an alternative approach by pre-serializing possibly conflicting transactions, and parallelizing non-conflicting update transactions to different replicas. This system provides strong transactional guarantees. In effect, Gargamel partitions the database dynamically according to the update workload. Each database replica runs sequentially, at full bandwidth; mutual synchronisation between replicas remains minimal. 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.

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.

5.3.5. 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).

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? What are the upper and lower bounds of CRDTs? We co-authored an innovative approach to these questions, to be published at Principles of Programming Languages (POPL) 2014 [29].

We are currently developing an extreme-scale CRDT platform called SwiftCloud; see Section 4.2 .

5.3.6. Mixing commutative and non-commutative updates: reservations

Asynchronous updates are desirable because they ensure the system is available, fast and scalable. CRDTs are asynchronous, but cannot guarantee strong invariants, such as ensuring that a shared counter never goes negative. To solve this problem, we define a novel hybrid model that supports both synchronous and asynchronous updates, "red-blue-purple" consistency. The RPB model classifies updates into commutative, partially-commutative and non-commutative, and distinguishes the (global) states where partially-commutative operations can safely run asynchronously. We use reservation techniques to ensure operation in such states. A reservation promises, to a cache that holds it, that the system is in a state that allows the cache server to perform purple updates asynchronously. Reservations ensure that data is in a known state by caching both data and access permissions over data to make updates. This approach strengthens the safety guarantees in addition to eventual consistency [40].

5.4. Performance and Robustness of Systems Software in Multicore Architectures

Participants: Koutheir Attouchi, Harris Bakiras, Antoine Blin, Florian David, Bertil Folliot, Lokesh Gidra, Julia Lawall, Jean-Pierre Lozi, Gilles Muller [correspondent], Dang Nhan Nguyen, Thomas Preud'Homme, Suman Saha, Peter Senna Tschudin, Marc Shapiro, Julien Sopena, Gaël Thomas, Mudit Verma.

5.4.1. Managed Runtime Environments

Today, multicore architectures are becoming ubiquitous, found even in embedded systems, and thus it is essential that managed runtime environments can scale on multicore processors. We have found that two major scalability bottlenecks are the implementation of highly contented locks and of garbage collectors. On a multicore, a single lock can overload the bus because the cache line that contains the lock bounces between the cores, eliminating all the 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. This work initially targeted C code but is now being adapted to the needs of Java applications in the PhD of Florian David. For garbage collectors, as the memory is physically distributed among a set of memory controllers, a collection saturates the bus when the collector threads access remote memory. This saturation prevents the garbage collector from scaling with the number of cores, making the garbage collector a major bottleneck of managed runtime environments on multicore hardware. As part of the PhD of Lokesh Gidra, we have identified memory placement schemes that decrease the number of remote memory accesses during a collection in OpenJDK 7, thus preventing the bottleneck caused by bus saturation [36].

5.4.2. System software robustness

A widely recognized problem in the area of finding bugs in API usage in systems code is to know what APIs are expected and to identify contexts where these expectations are not satisfied. Indeed, systems code, such as an operating systems kernel, is typically voluminous, amounting to millions of lines of code, and uses many different highly specialized APIs, making it impossible for most developers to keep the usage protocols of all of them in mind. To address this issue, we have developed an approach to inferring API function usage protocols from software, relying on knowledge of common code structures (Software – Practice and Experience [26]). Building on this experience, we have developed an approach to finding resource-release omission faults in systems code that leverages information local to a single function [44]. This approach permits finding hundreds of faults in Linux kernel code as well as a variety of other systems software, with a low rate of false positives. Finally, we have initiated an effort on understanding the range and scope of the oops reports collected in the recently revived Linux kernel oops repository [59].

Beyond finding faults in existing code, we have also considered how systems code is constructed. Specifically, in the context of Linux device drivers, we have identified the notion of a *gene*, as a sequence of code fragments that express a particular device or operating system functionality. We have performed an initial partial sequencing of the genes making up the probe functions of Linux platform drivers [45]. Relatedly, in the context of a Merlion collaboration grant with David Lo of Singapore Management University, we have considered the problem of recommending APIs to developers. We propose one approach based on the set of libraries used by other software having similar properties [47], and a second approach based on the set of libraries used to implement related feature requests [48].

5.4.3. Domain-specific languages for systems software

A challenge in the management of a datacenter is the placement of application replicas, both to avoid a single point of failure and to limit communication costs. We have proposed a novel approach, BtrPlace [23], based on the use of a domain-specific language to express constraints derived from properties of the application and of the datacenter, and the use of a constraint solver to efficiently resolve these constraints. Simulations show that BtrPlace is able to repair a configuration involving 5000 servers after a server failure in 3 minutes.

While the use of domain-specific languages such as that of BtrPlace can ease programming, it is well known that developing, and especially maintaining, a domain-specific language over time is time-consuming and challenging. This is particularly the case when the domain-specific language provides domain-specific verifications, as the code implementing these verifications has to be maintained along with the rest of the language implementation. Furthermore, new domain-specific languages typically must evolve frequently, as the language developer comes to better understand the range and scope of the domain. To address these issues, we have proposed a methodology for domain-specific language implementation development for C-like domain-specific languages [19], based on the use of rewriting rules implemented using Coccinelle. We apply this approach to our previously developed domain specific language z2z for developing network gateways, and find that the resulting language implementation is more concise and easier to extend with new language features.

REGULARITY Project-Team

6. New Results

6.1. Stochastic integration with respect to the Rosenblatt process.

Participant: Benjamin Arras.

From a theoretical perspective to more concrete applications, fractional Brownian motion (fbm) is a fruitful and rich mathematical object. From its stochastic analysis, initiated during the nineties, several theories of stochastic integration have emerged so far. Indeed, fbm is, in general, not a semimartingale neither a Markov process. These theories rely on different properties of the stochastic integrator process and are then of different natures. Despite the quite large number of these strategies, we can group them into two fundamentally distinct categories: the pathwise and the probabilistic approaches. The probabilistic one requires highly evolved stochastic analysis tools. Indeed, the Malliavin calculus as well as Hida's distribution theory have been used in order to define stochastic integration with respect to fractional Brownian motion ([56], [52]) and more general Gaussian processes ([47]). Moreover, fbm belongs to an important class of stochastic processes, namely, the Hermite processes. This class appears in non-central limit theorems for processes defined as integrals or partial sums of non-linear functionals of stationary Gaussian sequences with long-range dependence (see [57]). They admit the following representation for all $d\geq 1$:

$$\forall t>0 \quad Y_t^{H,d}=c(H_0)\int_{\mathbb{R}}...\int_{\mathbb{R}}\left(\int_0^t\prod_{j=1}^d(s-x_j)_+^{H_0-1}ds\right)dB_{x_1}...dB_{x_d}$$

where $c(H_0)$ is a normalizing constant such that $\mathbb{E}[|Y_1^{H,d}|^2]=1$ and $H_0=\frac{1}{2}+\frac{H-1}{d}$ with $H\in(\frac{1}{2},1)$. For $d=1$, one recovers fractional Brownian motion. These processes share many properties with fbm. Namely, they are H-self-similar processes with stationary increments. They possess the same covariance structure, exhibit long range-dependence and their sample paths are almost-surely δ-Hölder continuous, for every $\delta<H$. For $d=2$, the process is called the Rosenblatt process. This process has received lots of interest in the past and more recent years. Stochastic calculus with respect to the Rosenblatt process has been developped in [73] from both, the pathwise type calculus and Malliavin calculus points of view. Even if these two approaches are successful in order to define a stochastic integral with respect to the Rosenblatt process, the Malliavin calculus one fails to give an Itô's formula for the Rosenblatt process in the divergence sense. In [42], by means of white noise distribution theory, we obtain the following result:

Theorem: Let $(a,b)\in\mathbb{R}_+^*$ such that $a\leq b<\infty$. Let F be an entire analytic function of the complex variable verifying:

$$\exists N\in\mathbb{N},\exists C>0,\forall z\in\mathbb{C}\quad |F(z)|\leq C(1+|z|)^N\exp(\frac{1}{\sqrt{2}b^H}|\Im(z)|)$$

Then, we have in $(S)^*$:

$$F(X_b^H)-F(X_a^H)=\int_a^b F^{(1)}(X_t^H)\diamond\dot{X}_t^H dt+\sum_{k=2}^{\infty}\left(H\kappa_k(X_1^H)\int_a^b\frac{t^{Hk-1}}{(k-1)!}F^{(k)}(X_t^H)dt+2^k\int_a^b F^{(k)}(X_t^H)\diamond\dot{X}_t^{H,k}dt\right),$$

where $\{X_t^H\} = \{Y_t^{H,2}\}$, $\{\dot{X}_t^H\}$ is the Rosenblatt noise, $\{\kappa_k(X_1^H); k \geq 2\}$ the non-zero cumulants of the Rosenblatt distribution, $\diamond$ the Wick product and $\{\{X_t^{H,k}\} : k \geq 2\}$ a sequence of processes defined by:

$$\forall t \geq 0 \quad X_t^{H,k} = \int_{\mathbb{R}} \int_{\mathbb{R}} \underbrace{(...((f_t^H \otimes_1 f_t^H) \otimes_1 f_t^H) ... \otimes_1 f_t^H)}_{k-1 \times \otimes_1}(x_1, x_2) dB_{x_1} dB_{x_2}$$

with $f_t^H(x_1, x_2) = c(H) \int_0^t \prod_{j=1}^2 (s - x_j)_+^{\frac{H}{2}-1} ds$ and $\otimes_1$ is the contraction of order 1.
Moreover, in the same setting, we obtain the following "isometry" result for the Rosenblatt noise integral of sufficiently "good" integrand processes:
Theorem: Let $\{\phi_t; t \in I\}$ be a stochastic process such that for all $t \in I$ (I an interval), $\phi_t \in (L^2)$ and such that the Rosenblatt noise integral of $\{\phi_t\}$ exists in $(S)^*$. Moreover, let us assume that:

$$\sum_{m=0}^{+\infty} (m+2)! \int_I \int_I |t-s|^{2(H-1)} < f_m(.,t); f_m(.,s) >_{L^2(\mathbb{R}^m)} dtds < +\infty,$$

where $\phi_t = \sum_{m=0}^{+\infty} I_m(f_m(.,t))$. Thus, we have:

$$\begin{aligned}
&\mathbb{E}[(\int_I \phi_t \diamond \dot{X}_t^H dt)^2] = H(2H-1) \int_I \int_I |t-s|^{2(H-1)} \mathbb{E}[\phi_t \phi_s] dsdt \\
&+4\sqrt{\frac{H(2H-1)}{2}} \int_I \int_I |t-s|^{H-1} \mathbb{E}[D_{\sqrt{d(H)}\delta_s \circ I_+^{\frac{H}{2}}}(\phi_t) D_{\sqrt{d(H)}\delta_t \circ I_+^{\frac{H}{2}}}(\phi_s)] dsdt \\
&+ \int_I \int_I \mathbb{E}[(D_{\sqrt{d(H)}\delta_s \circ I_+^{\frac{H}{2}}})^2(\phi_t)(D_{\sqrt{d(H)}\delta_t \circ I_+^{\frac{H}{2}}})^2(\phi_s)] dsdt,
\end{aligned}$$

where $D_{\sqrt{d(H)}\delta_s \circ I_+^{\frac{H}{2}}}$ is the derivative operator in the direction $\sqrt{d(H)}\delta_s \circ I_+^{\frac{H}{2}}$.
Finally, in the last section of [42], we compare our approach to the one of [73]. More specifically, we prove that the stochastic integral with respect to the Rosenblatt process built using Malliavin calculus corresponds with the Rosenblatt noise integral when both of them exist.
Proposition: Let $\{\phi_t; t \in [0;T]\}$ be a stochastic process such that $\phi \in L^2(\Omega; \mathcal{H}) \cap L^2([0,T]; \mathbb{D}^{2,2})$ and $\mathbb{E}[\int_0^T \int_0^T ||D_{s_1,s_2}\phi||_{\mathcal{H}}^2] ds_1 ds_2 < \infty$ where

$$\mathcal{H} = \{f : [0;T] \to \mathbb{R}; \int_0^T \int_0^T f(s)f(t)|t-s|^{2H-2} dsdt < \infty\}.$$

Then, $\{\phi_t\}$ is Skorohod integrable and $(S)^*$-integrable with respect to the Rosenblatt process, $\{Z_t^H\}_{t \in [0;T]}$, and we have:

$$\int_0^T \phi_t \delta Z_t^H = \int_0^T \phi_t \diamond \dot{Z}_t^H dt$$

6.2. Sample path properties of multifractional Brownian motion

Participants: Paul Balança, Erick Herbin [supervision].

In [50], we have investigated the geometry of the sample paths of multifractional Brownian motion. Several representations of mBm exist, including the classic integral form:

$$\forall t \in \mathbf{R}; \quad X_t = \frac{1}{\Gamma\left(H(t) + \frac{1}{2}\right)} \int_{\mathbf{R}} \left[(t-u)_+^{H(t)-1/2} - (-u)_+^{H(t)-1/2}\right] \mathrm{d}W_u,$$

where $H : \mathbf{R} \mapsto (0,1)$ is a continuous function. Interestingly, we observe that geometric properties obtained in the probabilistic literature usually rely on a key assumption on the behaviour of the Hurst function:

$$H \text{ is a } \beta\text{-Hölder continuous function such that } \forall t \in \mathbf{R},\ \ H(t) < \beta. \qquad (\mathcal{H}_0) \tag{33}$$

Under the previous hypothesis, the local regularity of the mBm at t corresponds to the geometry of a fractional Brownian motion of parameter $H(t)$. Nevertheless, it has been shown in [15] that when this assumption does not hold, the sample path properties are not as simple and straightforward. More precisely, the latter has proved that the Hölder exponents satisfy at every $t \in \mathbf{R}$:

$$\alpha_{X,t} = H(t) \wedge \alpha_{H,t} \quad \text{ and } \quad \widetilde{\alpha}_{X,t} = H(t) \wedge \widetilde{\alpha}_{H,t} \quad \text{a.s.} \tag{34}$$

This result has been recently improved in [48], observing that the pointwise exponent can even be random under some assumptions on H.

Therefore, the main goal of this work was to obtain a more complete characterization of the geometry of the general mBm. We have first focused on the Hölder regularity of the sample paths, using for this purpose a deterministic representation of the fractional Brownian field:

$$B^{\pm}(t,H) = \frac{\pm 1}{\Gamma\left(H-\frac{1}{2}\right)} \int_{\mathbf{R}} B_u \left[(t-u)_{\pm}^{H-3/2} - (-u)_{\pm}^{H-3/2} \right] \mathrm{d}u, \tag{35}$$

where $H \geq 1/2$ and B is a continuous Brownian motion. Hence, observing that the mBm almost surely corresponds to the fractional integration of a Brownian motion, we have been able to use the 2-microlocal formalism and its interesting connections with fractional operators. As a consequence, we have proved that the pointwise exponent of the mBm almost surely satisfies:

$$\forall t \in \mathbf{R}; \quad \alpha_{X,t} = H(t) \wedge m_{t,H(t)} \alpha_{H,t}, \tag{36}$$

where $m_{t,H(t)}$ is defined as the multiplicity of the fractional Brownian field at (t,H), i.e.

$$m_{t,H} = \inf \left\{ k \in \mathbf{N} \setminus \{0\} : \partial_H^k B(t,H) \neq 0 \right\}.$$

We have also been able to obtain some uniform lower bounds on the 2-microlocal frontier, which are optimal under some mild assumptions on the Hurst function.

The second direction of our study has concerned the fractal dimension of the graph of the mBm. Interestingly, and on the contrary to fBm, we have to distinguish the Box and Hausdorff dimensions in our result. The first happens to be the easiest one to study and is closely related to the geometry of H itself. Therefore, with probability one,

$$\forall t \in \mathbf{R} \setminus \{0\}; \quad \dim_{\mathrm{B},t} \mathrm{Gr}(X) = (2 - H(t)) \vee \dim_{\mathrm{B},t} \mathrm{Gr}(H), \tag{37}$$

where $\dim_{\mathrm{B},t}$ denotes the localized Box dimension at t.

To study the Hausdorff dimension the graph, we need a slightly different approach which makes use of parabolic Hausdorff dimension. We first define for all $t \in \mathbf{R}$ a *parabolic metric*ϱ_H on $\mathbf{R}^2$, with $H > 0$: $\varrho_H\left((u,x);(v,y)\right) := \max\left(|u-v|^H, |x-y|\right)$. For any set $A \subset \mathbf{R}^2$, we denote by $\dim_{\mathcal{H}}(A;\varrho_H)$ the *parabolic Hausdorff dimension* of A. It is defined similarly to the classic Hausdorff dimension using covering balls relatively to the metric ϱ_H, i.e. it corresponds to the infimum of $s \geq 0$ for which

$$\liminf_{\delta \to 0}\left\{\sum_{i=0}^{\infty} \text{diam }\left(O_i\,;\varrho_H\right)^s : (O_i)_{i\in N} \text{ is a } \delta\text{-cover of } A\right\} < \infty$$

Studying the local Hausdorff dimension of the graph of the mBm, we have proved that with probability one

$$\forall t \in \mathbf{R} \setminus \{0\}; \quad \dim_{\mathrm{H},t} \mathrm{Gr}(X) = 1 + H(t)\left(\dim_{\mathrm{H},t}\left(\mathrm{Gr}(H)\,;\varrho_{H(t)}\right) - 1\right). \tag{38}$$

Even though this result might seem counter-intuitive, it can be checked that it induced the classic equality $\dim_{\mathrm{H},t}\mathrm{Gr}(X) = 2 - H(t)$ when the mBm satisfies the assumption $\mathcal{H}_0$. Interestingly, we observe that a similar expression has also emerged recently in the study [70] of the Hausdorff dimension of a fractional Brownian motion with variable drift. Finally, we also note this result can also been extended to images of fractal sets by the multifractional Brownian motion.

6.3. Large Deviations Inequalities

Participant: Xiequan Fan.

Let $(\xi_i)_{i=1,\dots,n}$ be a sequence of independent and centered random variables satisfying Bernstein's condition, for a constant $\varepsilon > 0$,

$$|\mathbb{E}\xi_i^k| \leq \frac{1}{2} k!\varepsilon^{k-2}\mathbb{E}\xi_i^2, \qquad \text{for all } k \geq 2 \text{ and all } i = 1, ..., n. \tag{39}$$

Denote by

$$S_n = \sum_{i=1}^{n} \xi_i \quad \text{and} \quad \sigma^2 = \sum_{i=1}^{n} \mathbb{E}\xi_i^2. \tag{40}$$

The well-known Bernstein inequality (1946) states that, for all $x > 0$,

$$\mathbb{P}(S_n > x\sigma) \;\leq\; \inf_{\lambda \geq 0} \mathbb{E}e^{\lambda(S_n - x\sigma)}. \tag{41}$$

In the i.i.d. case, Cramér (1938) has established a large deviation expansion under the condition $\mathbb{E}e^{|\xi_1|} < \infty$. For all $0 \leq x = o\left(\sqrt{n}\right)$, one has

$$\frac{\mathbb{P}(S_n > x\sigma)}{1 - \Phi(x)} = e^{\frac{x^3}{\sqrt{n}}\lambda\left(\frac{x}{\sqrt{n}}\right)}\left[1 + O\left(\frac{1+x}{\sqrt{n}}\right)\right], \qquad n \to \infty, \tag{42}$$

where $\lambda(\cdot) = c_1 + c_2\frac{x}{\sqrt{n}} + ...$ is the Cramér series and the values $c_1, c_2, ...$ depend on the distribution of ξ_1.

Bahadur-Rao (1960) proved the following sharp large deviations similar to (15). Assume Cramér's condition. Then, for given $y > 0$, there is a constant c_y depending on the distribution of ξ_1 and y such that

$$\mathbb{P}\left(\frac{S_n}{n} > y\right) = \frac{\inf_{\lambda \geq 0}\mathbb{E}e^{\lambda(S_n - yn)}}{\sigma_y t_y \sqrt{2\pi n}}\left[1 + O\left(\frac{c_y}{n}\right)\right], \qquad n \to \infty, \tag{43}$$

where t_y, σ_y and c_y depend on the distribution of ξ_1 and y.

We present an improvement on Bernstein's inequality. In particular, we establish a sharp large deviation expansion similar to the classical results of Cramér and Bahadur-Rao. The following theorem is our main result.

Theorem 0.1 Assume Bernstein's condition. Then, for all $0 \le x < \frac{1}{12}\frac{\sigma}{\varepsilon}$,

$$\mathbb{P}(S_n > x\sigma) = \inf_{\lambda \ge 0} \mathbb{E} e^{\lambda(S_n - x\sigma)} F\left(x, \frac{\varepsilon}{\sigma}\right), \tag{44}$$

where $\sqrt{2\pi}M(x)$ is the Mills ratio, the function

$$F\left(x, \frac{\varepsilon}{\sigma}\right) = M(x) + 28\,\theta R\left(4x\varepsilon/\sigma\right)\frac{\varepsilon}{\sigma} \tag{45}$$

with

$$R(t) = \frac{\left(1 - t + 6t^2\right)^3}{(1-3t)^{3/2}(1-t)^7}, \qquad 0 \le t < \frac{1}{3}, \tag{46}$$

and $|\theta| \le 1$. In particular, in the i.i.d. case, for all $0 \le x = o(\sqrt{n}), n \to \infty$,

$$\left|\mathbb{P}(S_n > x\sigma) - M(x)\inf_{\lambda \ge 0} \mathbb{E} e^{\lambda(S_n - x\sigma)}\right| = O\left(\frac{1}{\sqrt{n}}\inf_{\lambda \ge 0} \mathbb{E} e^{\lambda(S_n - x\sigma)}\right) \tag{47}$$

and thus

$$\frac{\mathbb{P}(S_n > x\sigma)}{M(x)\inf_{\lambda \ge 0} \mathbb{E} e^{\lambda(S_n - x\sigma)}} = 1 + o(1). \tag{48}$$

6.4. A fractional Brownian field indexed by L^2 and a varying Hurst parameter

Participant: Alexandre Richard.

Using structures of Abstract Wiener Spaces and their reproducing kernel Hilbert spaces, we define a fractional Brownian field indexed by a product space $(0, 1/2] \times L^2(T, m)$, where the first coordinate corresponds to the Hurst parameter of fractional Brownian motion. This field encompasses a large class of existing fractional Brownian processes, such as Lévy fractional Brownian motion and multiparameter fractional Brownian motion, and provides a setup for new ones. We prove that it has good incremental variance in both coordinates and derive certain continuity and Hölder regularity properties. Then, we apply these general results to multiparameter and set-indexed processes, which proves the existence of processes with prescribed local Hölder regularity on general indexing collections.

The family of fBm can be considered for the different Hurst parameters as a single Gaussian process indexed by $(h, t) \in (0, 1) \times \mathbb{R}_+$, which is the position we adopt. Besides, the "time" indexing is replaced by any separable L^2 space. We prove that there exists a Gaussian process indexed by $(0, 1/2] \times L^2(T, m)$, with the additional constraint that the variance of its increments is as well behaved as it was on $(0, 1) \times \mathbb{R}_+$, that is, for any compact of L^2, there is a constant $C > 0$ such that for any f in this compact, and any $h, h' \in (0, 1/2)$,

$$\mathbb{E}\left(B_f^h - B_f^{h'}\right)^2 \le C\ (h - h')^2. \tag{49}$$

When looking at the L^2–fBf with a fixed h, we have the following covariance: for each $h \in (0, 1/2]$,

$$k_h : (f,g) \in L^2 \times L^2 \mapsto \frac{1}{2}\left(m(f^2)^{2h} + m(g^2)^{2h} - m(|f-g|^2)^{2h}\right) . \tag{50}$$

An important subclass of these processes is formed by processes restricted to indicator functions of subsets of T. In particular, multiparameter when $(T,m) = (\mathbb{R}_+^d, \text{Leb.})$, and more largely set-indexed processes [62],[20] naturally appear and thus motivate generalization *b)*, besides the inherent interest of studying processes over an abstract space.

To define this field, we used fractional operators on the Wiener space W introduced in [56], and first expressed the fractional Brownian field (indexed by $(0,1/2] \times \mathbb{R}_+$) as a white noise integral over W:

$$\left\{ \int_W \langle \mathcal{K}_h R_h(\cdot,t), w\rangle \; \mathrm{d}\mathbb{B}_w, \;\; (h,t) \in (0,1/2] \times \mathbb{R}_+ \right\} ,$$

The advantage of this approach is to allow the transfer of techniques of calculus on the Wiener space to any other linearly isometric space with the same structure (those spaces are called Abstract Wiener Spaces). Using the separability and reproducing kernel property of the Cameron-Martin spaces built from the kernels $k_h, h \in (0,1/2]$, we prove the existence of a Brownian field $\{\mathbf{B}_{h,f}, \;\; h \in (0,1/2], f \in L^2(T,m)\}$ over some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Some Hilbert space analysis then provides the desired bound (22). Then, we used this to derive a sufficient condition for almost sure continuity of the fractional Brownian field, in terms of metric entropy.

For fixed h, we proved that the h-fractional Brownian motion has the strong local nondeterminism property, which allowed to compute a sharp estimate of its small deviations, that is, for a compact K of L^2:

$$\exp\left(-C \;\; N(K,d_h,\varepsilon)\right) \leq \mathbb{P}\left(\sup_{f \in K} |\mathbf{B}_f^h| \leq \varepsilon \right) \leq \exp\left(-C^{-1} \;\; N(K,d_h,\varepsilon)\right) \;\; ,$$

where $N(K,d_h,\varepsilon)$ is the metric entropy of K, i.e., the minimal number of balls necessary to cover K with d_h-balls (the metric induced by the h-fBm) of radius at most ε.

Finally, we looked at the Hölder regularity of the fBf, when the L^2 indexing collection is restricted to the indicator functions of the rectangles of $\mathbb{R}^d$ (multiparameter processes) or to some indexing collection (in the sense of [62]). This restriction permits to use local Hölder regularity exponents, in the flavour of what was done in [24]. When a regular path $\mathbf{h} : L^2 \to (0,1/2]$ is specified, this defines a multifractional Brownian field as $\mathbf{B}_f^{\mathbf{h}} = \mathbf{B}_{\mathbf{h}(f),f}$, whose Hölder regularity at each point is proved to equal $\mathbf{h}(f)$ almost surely.

6.5. Self-stabilizing processes

Participants: Xiequan Fan, Jacques Lévy Véhel.

In collaboration with K. Falconer, University of St Andrews.

Self-stabilizing processes are càdlàg processes whose local intensiy of jumpd depend on amplitude. We have investigated two paths to define such processes. The first one is based on a modification of the celebrated Lévy construction of Brownian motion.

The second one starts from a stochastic differentiel equation, and allows one to build Markov processes, a useful feature in applications such as financial modelling [40], [41].

6.6. Multifractal spectra of multistable Lévy motion

Participant: Jacques Lévy Véhel.

In collaboration with R. Le Guével, University of Rennes.

As a follow-up to the work in [34] we have computed the Hausdorf, large deviation, and Legendre multifractal spectra of multistable Lévy motion. It turns out that the shape of the Hausdorf multifractal spectrum is much more complex than could be expected considering the corresponding spectrum of plain Lévy motion. Also, the large deviation spectrum reveals more information on the fine structure of the process than the Hausdorf one, a situation reminiscent of what has already been observed for the model we have developped previously for TCP traffic [2],[39].

6.7. Self-regulating processes for the modelling of geophysical signals

Participant: Jacques Lévy Véhel.

In collaboration with A. Echelard and A. Philippe, University of Nantes.

We have shown that various geophysical signals, and in particular temperature records, can be modelled with self-regulating processes as introduced in [4]. For this purpose, we have used an estimator of the self-regulating function proposed in [44]. Such a modelling allows one to gain further insight on the fine structure of the evolution of temperatures.

6.8. Regularity-preserving signal denoising

Participant: Jacques Lévy Véhel.

In collaboration with A. Echelard.

We have proposed a new wavelet-based method for signal denoising, that allows one to recover the local Hölder regularity of the original signal under weak assumptions [43]. The algotihm is a modification of the well-known wavelet thresholding procedure, where "small" coefficients are not put to zero, but modified in a way governed by the behaviour of large scale coefficients. This will have applications in the frame of our Tandem project on the analysis of radar images.

REO Project-Team

6. New Results

6.1. Mathematical and numerical analysis of fluid-structure interaction problems

Participants: Muriel Boulakia, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Céline Grandmont, Mikel Landajuela Larma, Jimmy Mullaert, Marina Vidrascu.

- In [17] we analyze the performances of two types of Luenberger observers – namely, the so-called Direct Velocity Feedback and Schur Displacement Feedback procedures, originally devised for elasto-dynamics – 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. We first assess the observers using hemodynamics-inspired test problems with the complete model, including the Navier-Stokes equations in Arbitrary Lagrangian-Eulerian formulation, in particular. Then, in order to obtain more detailed insight we consider several well-chosen simplified models, each of which allowing a thorough analysis – emphasizing spectral considerations – while illustrating a major phenomenon of interest for the observer performance, namely, the added mass effect for the structure, the coupling with a lumped-parameter boundary condition model for the fluid flow, and the fluid dynamics effect per se. Whereas improvements can be sought for when additional measurements are available in the fluid domain, the present framework this establishes Luenberger observer methods as very attractive strategies – compared, e.g., to classical variational techniques – to perform state estimation, and more generally for uncertainty estimation since other observer procedures can be conveniently combined to estimate uncertain parameters.
- In [28] we introduce a class of incremental displacement-correction schemes for the explicit coupling of a thin-structure with an incompressible fluid. We provide a general stability and convergence analysis that covers both the incremental and the non-incremental variants. Their stability properties are independent of the added-mass effect. The superior accuracy of the incremental schemes (with respect to the original non-incremental variant) is highlighted by the error estimates, and then confirmed in a benchmark by numerical experiments.
- In [29], [62] we introduce a class of fully decoupled time-marching schemes (velocity-pressure-displacement splitting) for the coupling of an incompressible fluid with a thin-walled viscoelastic structure. A priori energy estimates guaranteeing unconditional stability are established for the variants without extrapolation and with first-order extrapolation. The accuracy and performance of the methods proposed are discussed in several numerical examples.
- In [30] we introduce a class of explicit coupling schemes for the numerical solution of fluid-structure interaction problems involving a viscous incompressible fluid and a general thin-walled structure (e.g., including damping and non-linear behavior). The fundamental ingredient in these methods is the (parameter free) explicit Robin interface condition for the fluid, which enables the fluid-solid splitting through appropriate extrapolations of the solid velocity and fluid stress on the interface. The resulting solution procedures are genuinely partitioned. Stability and error estimates are provided for all the variants (depending on the extrapolations), using energy arguments within a representative linear setting. In particular, we show that one of them yields added-mass free unconditional stability and optimal (firs-order) time accuracy. A comprehensive numerical study, involving different examples from the literature, supports the theory.
- In [63] we introduce a new class of explicit coupling schemes for the numerical solution of fluid-structure interaction problems involving a viscous incompressible fluid and an elastic structure. These methods generalize the arguments reported in [28], [30] to the case of the coupling with thick-walled structures. The basic idea lies in the derivation of an intrinsic interface Robin consistency at

the space semi-discrete level, using a lumped-mass approximation in the structure. The fluid-solid splitting is then performed through appropriate extrapolations of the solid velocity and stress on the interface. Based on these methods, a new, parameter-free, Robin-Neumann iterative procedure is also proposed for the partitioned solution of implicit coupling. A priori energy estimates, guaranteeing the (added-mass free) stability of the schemes and the convergence of the iterative procedure, are established. The accuracy and robustness of the methods are illustrated in several numerical examples.

- In [22] we discuss explicit coupling schemes for fluid-structure interaction problems where the added mass effect is important. We show the close relation between coupling schemes using Nitsche's method and a Robin-Robin type coupling. In the latter case the method may be implemented either using boundary integrals of the stresses or the more conventional discrete lifting operators. We also make the observation that these scheme are stable under a hyperbolic type CFL condition, but that optimal accuracy imposes a parabolic type CFL conditions due to the splitting error. Two strategies to enhance the accuracy of the coupling scheme under the hyperbolic CFL-condition are suggested, one using extrapolation and defect-correction and one using a penalty-free non-symmetric Nitsche method. Finally we illustrate the performance of the proposed schemes on some numerical examples in two and three space dimensions.
- In [60] we consider the extension of the Nitsche method to the case of fluid-structure interaction problems on unfitted meshes. We give a stability analysis for the space semi-discretized problem and show how this estimate may be used to derive optimal error estimates for smooth solutions, irrespective of the mesh/interface intersection. Some numerical examples illustrate the theoretical discussion.
- In [21] we are interested by the three-dimensional coupling between an incompressible fluid and a rigid body. The fluid is modeled by the Navier-Stokes equations, while the solid satisfies the Newton's laws. In the main result of the paper we prove that, with the help of a distributed control, we can drive the fluid and structure velocities to zero and the solid to a reference position provided that the initial velocities are small enough and the initial position of the structure is close to the reference position. This is done without any condition on the geometry of the rigid body.
- In the book chapter [57] we deal with some specific existence and numerical results applied to a 2D/1D fluid-structure coupled model, for an incompressible fluid and a thin elastic structure. We underline some of the mathematical and numerical difficulties that one may face when studying this kind of problems such as the geometrical nonlinearities or the added mass effect. In particular we underline the link between the strategies of proof of weak or strong solutions and the possible algorithms to discretize these types of coupled problems.

6.2. Numerical methods for fluid mechanics and application to blood flows

Participants: Grégory Arbia, Benoit Fabrèges, Jean-Frédéric Gerbeau, Sanjay Pant, Saverio Smaldone, Marc Thiriet, Irène Vignon-Clementel.

- In [61] we propose a new approach to the loosely coupled time-marching of a fluid-fluid interaction problems involving the incompressible Navier-Stokes equations. The methods combine a specific explicit Robin-Robin treatment of the interface coupling with a weakly consistent interface pressure stabilization in time. A priori energy estimates guaranteeing stability of the splitting are obtained for a total pressure formulation of the coupled problem. The performance of the proposed schemes is illustrated on several numerical experiments related to simulation of aortic blood flow.
- In [49] we present our strategy to meet the MICCAI Challenge 2013: the goal was to recover a measured (but unrevealed) pressure drop across a coarctation of the aorta through 3D simulation. A filtering-based strategy is devised to perform parameter estimation and subsequent multiscale CFD simulations of arterial blood flow. The method is applied to the patient-specific case in the two physiological states of rest and stress. In both cases, the method is shown to be effective in closely matching the available clinically measured data. Pressure drop across the coarctation is predicted

for both states. At the time of [47], these measurements were available: the computed pressure drop across the coarctation for the stress case appears to be very close to the measured one, while the one for the rest case is not as good. One should note that no participant of the challenge managed to recover the measured pressure drop for the rest case.

- In [35], we aim to reduce the complexity of patient-specific simulations by combining image analysis, computational fluid dynamics and model order reduction techniques. The proposed method makes use of a reference geometry estimated as an average of the population, within an efficient statistical framework based on the currents representation of shapes. Snapshots of blood flow simulations performed in the reference geometry are used to build a POD (Proper Orthogonal Decomposition) basis, which can then be mapped on new patients to perform reduced order blood flow simulations with patient specific boundary conditions. This approach is applied to a data-set of 17 tetralogy of Fallot patients to simulate blood flow through the pulmonary artery under normal (healthy or synthetic valves with almost no backflow) and pathological (leaky or absent valve with backflow) conditions to better understand the impact of regurgitated blood on pressure and velocity at the outflow tracts. The model reduction approach is further tested by performing patient simulations under exercise and varying degrees of pathophysiological conditions based on reduction of reference solutions (rest and medium backflow conditions respectively).
- In [16], we analyze two 3D-0D coupling approaches in which a fractional-step projection scheme is used in the fluid. Our analysis shows that explicit approaches might yield numerical instabilities, particularly in the case of realistic geometries with multiple outlets. We introduce and analyze an implicitly 3D-0D coupled formulation with enhanced stability properties and which requires a negligible additional computational cost. Furthermore, we also address the extension of these methods to fluid-structure interaction problems. The theoretical stability results are confirmed by meaningful numerical experiments in patient specific geometries coming from medical imaging.
- In [36], 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.
- 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 [43] 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.
- 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 [27], 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 [75]).

- In [26] we propose the first patient-specific predictive modeling of stage 2 palliation for congenital heart disease by using virtual surgery and closed-loop multi-scale modeling. We present a workflow to perform post-operative simulations from pre-operative clinical data. Two surgical options (bi-directional Glenn and hemi-Fontan operations) are virtually performed and coupled to the pre-operative LPM, with the hemodynamics of both options reported. Results are validated against postoperative clinical data.
- In [14] we study the influence of solvers and test case setup on the result of numerical simulations. The need for detailed construction of the numerical model depends on the precision needed to answer the biomedical question at hand and should be assessed for each problem on a combination of clinically relevant patient-specific geometry and physiological conditions. Methods and results between a commercial code and an in-house research code are illustrated on three congenital heart disease examples of increasing complexity. This publication is designed as a tool to better communicate with clinical researchers interested in simulations.

6.3. Numerical methods for cardiac electrophysiology

Participants: Muriel Boulakia, Jean-Frédéric Gerbeau, Annabelle Collin, Elisa Schenone.

- In [25], Computational electrophysiology is a very active field with tremendous potential in medical applications, albeit leads to highly intensive simulations. We here 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. Moreover, 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. Our proposed 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.
- In [45] we assess a previously-proposed surface-based electrophysiology model with detailed atrial simulations. This model - derived and substantiated by mathematical arguments - is specifically designed to address thin structures such as atria, and to take into account strong anisotropy effects related to fiber directions with possibly rapid variations across the wall thickness. The simulation results are in excellent adequacy with previous studies, and confirm the importance of anisotropy effects and variations thereof. Furthermore, this surface-based model provides dramatic computational benefits over 3D models with preserved accuracy.

6.4. Lung and respiration modeling

Participants: Grégory Arbia, Laurent Boudin, Muriel Boulakia, Benoit Fabrèges, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Céline Grandmont, Stéphane Liwarek, Jessica Oakes, Ayman Moussa, Irène Vignon-Clementel.

- In [66] we are interested in the mathematical modeling of the deformation of the human lung tissue, called the lung parenchyma, during the respiration process. The parenchyma is a foam-like elastic material containing millions of air-filled alveoli connected by a tree- shaped network of airways. In this study, the parenchyma is governed by the linearized elasticity equations and the air movement in the tree by the Poiseuille law in each airway. The geometric arrangement of the alveoli is assumed to be periodic with a small period. We use the two-scale convergence theory to study the asymptotic behavior as the period goes to zero. The effect of the network of airways is described by a nonlocal operator and we propose a simple geometrical setting for which we show that this operator converges. We identify in the limit the equations modeling the homogenized behavior under an abstract convergence condition on this nonlocal operator. We derive some mechanical properties of the limit material by studying the homogenized equations: the limit model is nonlocal both in space and time if the parenchyma material is considered compressible, but only in space if it is incompressible. Finally, we propose a numerical method to solve the homogenized equations and we study numerically a few properties of the homogenized parenchyma model.

- In [31] we present a calculation of the functioning of an acinus at exercise. We show that, given the geometry and the breathing dynamics of real acini, respiration can be correlated to a single equivalent parameter that we call the integrative permeability. We find that both V_{O2max} and PAO2 depend on this permeability in a non-linear manner.
- In [19], In this paper, we consider the Stokes equations and we are concerned with the inverse problem of identifying a Robin coefficient on some non accessible part of the boundary from available data on the other part of the boundary. We first study the identifiability of the Robin coefficient and then we establish a stability estimate of logarithm type thanks to a Carleman inequality due to A. L. Bukhgeim and 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 [18],In this work, 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 [39], In this paper we introduce a PDE system which aims at describing the dynamics of a dispersed phase of particles moving into an incompressible perfect fluid, in two space dimensions. The system couples a Vlasov-type equation and an Euler-type equation: the fluid acts on the dispersed phase through a gyroscopic force whereas the latter contributes to the vorticity of the former. First we give a Dobrushin type derivation of the system as a mean-field limit of a PDE system which describes the dynamics of a finite number of massive pointwise particles moving into an incompressible perfect fluid. This last system is itself inferred from a joint work of the second author with O. Glass and C. Lacave, where the system for one massive pointwise particle was derived as the limit of the motion of a solid body when the body shrinks to a point with fixed mass and circulation. Then we deal with the well-posedness issues including the existence of weak solutions. Next we exhibit the Hamiltonian structure of the system and finally, we study the behavior of the system in the limit where the mass of the particles vanishes.
- In [40] we solved for the airflow and aerosol particle distribution in healthy and emphysematous rat lungs. Following our preliminary work in [79], we first estimated the respiratory resistance and compliance parameters from pressure measurements taken during ventilation experiments performed in healthy and emphysematous rats. Next, the 3D Navier-Stokes equations were solved in a Magnetic Resonance derived airway geometry coupled to 0D models at the boundaries leading to the five rat lobes. The multiscale 3D-0D simulations enabled consistent pressure and airflow results, unlike what was found when a constant pressure was described at the boundaries. Aerosolized particles were tracked throughout inspiration and the effects of particle size and gravity were studied. Healthy, homogeneous and heterogeneous disease cases were assessed. Once available, these in-silico predictions may be compared to experimental deposition data.

6.5. Miscellaneous

Participants: Grégory Arbia, Laurent Boudin, Jean-Frédéric Gerbeau, Damiano Lombardi, Marina Vidrascu, Irène Vignon-Clementel.

- In [13] we analyse the solution of the linear advection equation on a uniform mesh by a non dissipative second order scheme for discontinuous initial condition. We focus on the case of advection of a step function by the leapfrog scheme. We derive closed form exact and approximate solutions for the scheme that accurately predict oscillations of the numerical scheme.
- In [41] The recent biomechanical theory of cancer growth considers solid tumors as liquid-like materials comprising elastic components. In this fluid mechanical view, the expansion ability of a solid tumor into a host tissue is mainly driven by either the cell diffusion constant or the cell division rate, the latter depending either on the local cell density (contact inhibition), on mechanical stress in the tumor, or both. For the two by two degenerate parabolic/elliptic reaction-diffusion system that results from this modeling, we prove there are always traveling 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 traveling waves are recovered numerically.
- In [68] This paper is devoted to the use of the entropy and duality methods for the existence theory of reaction-cross diffusion systems consisting of two equations, in any dimension of space. Those systems appear in population dynamics when the diffusion rates of individuals of two species depend on the concentration of individuals of the same species (self-diffusion), or of the other species (cross diffusion).
- In [65] We consider in this paper a spray constituted of an incompressible viscous gas and of small droplets which can breakup. This spray is modeled by the coupling (through a drag force term) of the incom- pressible Navier-Stokes equation and of the Vlasov-Boltzmann equation, together with a fragmentation kernel. We first show at the formal level that if the droplets are very small after the breakup, then the solutions of this system converge towards the solution of a simplified system in which the small droplets produced by the breakup are treated as part of the fluid. Then, existence of global weak solutions for this last system is shown to hold, thanks to the use of the DiPerna-Lions theory for singular transport equations.
- In [42],We propose a method of modelling sail type structures which captures the wrinkling behaviour 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 behaviour of various structures with spherical and cylindrical shapes, characteristic of downwind sails over a rather wide range of shape and 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, out of the scope of the present study.
- In [34],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 have 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 [48] We apply the domain decomposition method to linear elasticity problems for multi-materials

where the heterogeneities are concentrated in a thin internal layer. In the first case the heterogeneities are small, identical and periodically distributed on an internal surface and in the second one all the thin, curved internal layer is made of an elastic material much more strong than the surrounding one. In the first case the domain decomposition is used to efficiently solve the non-standard transmission problems obtained by the asymptotic expansion method. In the second case a non-standard membrane transmission problem originates from a surface shell like energy.

- In [32] : Our aim is to numerically validate the effectiveness of a matched asymptotic expansion formal method introduced in a pioneering paper by Nguetseng and Sànchez Palencia and extended in [76], [33]. Using this method a simplified model for the influence of small identical heterogeneities periodically distributed on an internal surface to the overall response of a linearly elastic body is derived. In order to validate this formal method a careful numerical study compares the solution obtained by a standard method on a fine mesh to the one obtained by asymptotic expansion. We compute both the zero and the first order terms in the expansion. To efficiently compute the first order term we introduce a suitable domain decomposition method.
- In [39] we introduce a PDE system which aims at describing the dynamics of a dispersed phase of particles moving into an incompressible perfect fluid, in two space dimensions. The system couples a Vlasov-type equation and an Euler-type equation: the fluid acts on the dispersed phase through a gyroscopic force whereas the latter contributes to the vorticity of the former. First we give a Dobrushin type derivation of the system as a mean-field limit of a PDE system which describes the dynamics of a finite number of massive pointwise particles moving into an incompressible perfect fluid. This last system is itself inferred from a joint work of the second author with O. Glass and C. Lacave, where the system for one massive pointwise particle was derived as the limit of the motion of a solid body when the body shrinks to a point with fixed mass and circulation. Then we deal with the well-posedness issues including the existence of weak solutions. Next we exhibit the Hamiltonian structure of the system and finally, we study the behavior of the system in the limit where the mass of the particles vanishes.
- In [65] we consider a spray constituted of an incompressible viscous gas and of small droplets which can breakup. This spray is modeled by the coupling (through a drag force term) of the incom- pressible Navier-Stokes equation and of the Vlasov-Boltzmann equation, together with a fragmentation kernel. We first show at the formal level that if the droplets are very small after the breakup, then the solutions of this system converge towards the solution of a simplified system in which the small droplets produced by the breakup are treated as part of the fluid. Then, existence of global weak solutions for this last system is shown to hold, thanks to the use of the DiPerna-Lions theory for singular transport equations.

REVES Project-Team

5. New Results

5.1. Plausible and Realistic Image Rendering

5.1.1. Depth Synthesis and Local Warps for Interactive Image-based Navigation

Participants: Gaurav Chaurasia, Sylvain Duchene, George Drettakis.

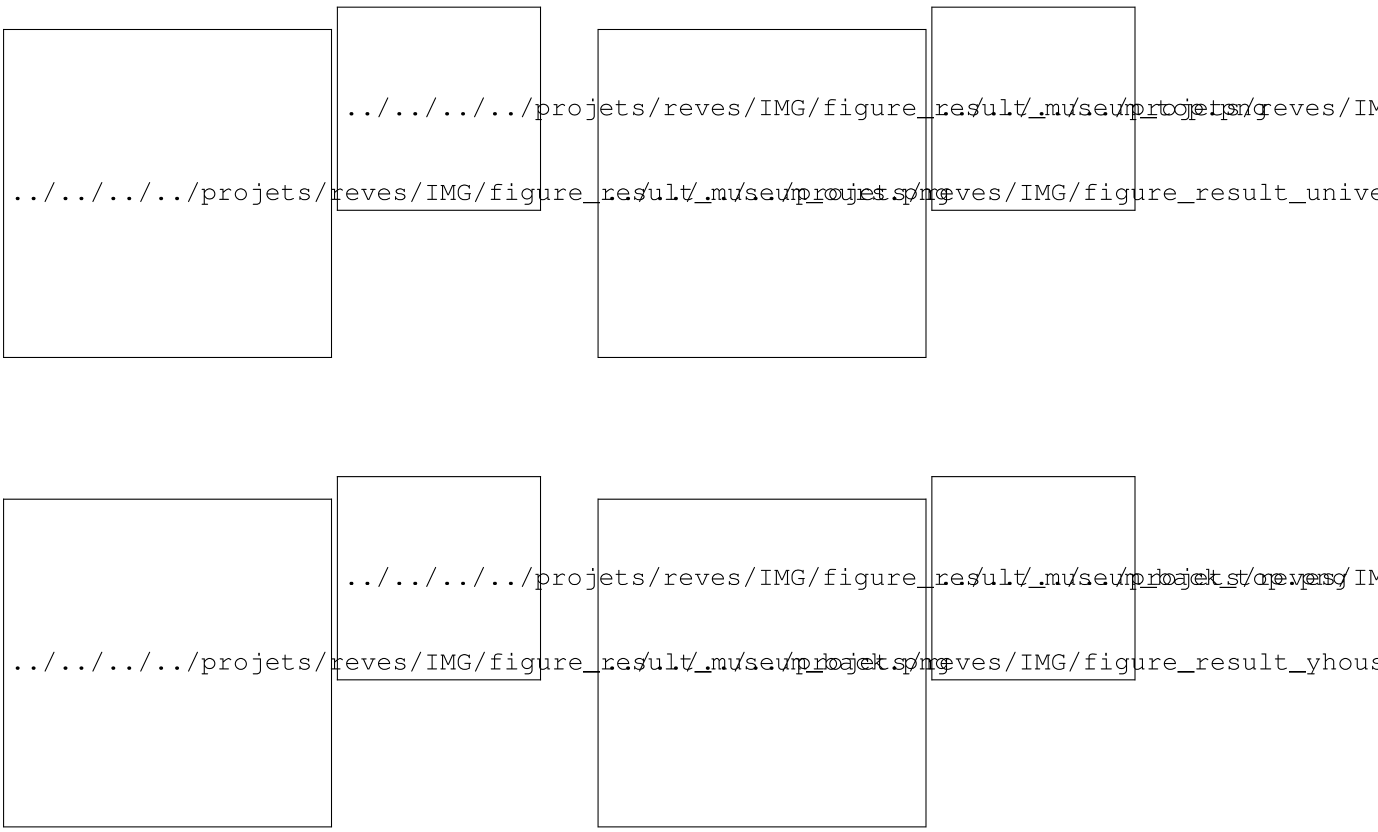


Figure 3. Novel views generated by the image-based rendering approach of [12] along with a visualization of novel camera position relative to the 3D scene and input cameras. This approach is among the first to handle very complex urban scenes such as those shown here and provide a stable solution for viewpoints that are far from the input cameras.

Modern camera calibration and multi-view stereo techniques enable users to smoothly navigate between different views of a scene captured using standard cameras. The underlying automatic 3D reconstruction methods work well for buildings and regular structures but often fail on vegetation, vehicles and other complex geometry present in everyday urban scenes. Consequently, missing depth information makes image-based rendering for such scenes very challenging. This paper introduces a new image-based rendering algorithm that is robust to missing or unreliable geometry, providing plausible novel views even in regions quite far from the input camera positions. The approach first oversegments the input images, creating superpixels of homogeneous color content which preserve depth discontinuities. It then introduces a *depth synthesis* step for poorly reconstructed regions. It defines a graph on the superpixels and uses *shortest walk* traversals to fill unreconstructed regions with approximate depth from regions that are well-reconstructed and similar in visual content. The superpixels augmented with synthesized depth allow a local shape-preserving warp which warps each superpixel of the input image to the novel view without incurring distortions and preserving the local visual content within the superpixel. This allows the approach to effectively compensate for missing photoconsistent depth, the lack of which is known to cause rendering artifacts. The final rendering algorithm blends the warped images, using heuristics to avoid ghosting artifacts. The results demonstrate novel view synthesis in real time for multiple challenging scenes with significant depth complexity (see Figure 3), providing a convincing immersive navigation experience. The paper presents comparisons with three of the state of the art image-based rendering techniques and demonstrate clear advantages.

This work was in collaboration with Olga Sorkine-Hornung at ETH Zurich. It has been published in ACM Transactions on Graphics 2013 [12] and presented at SIGGRAPH.

5.1.2. Megastereo: Constructing High-Resolution Stereo Panoramas

Participant: Christian Richardt.

There is currently a strong consumer interest in a more immersive experience of content, such as 3D photographs, television and cinema. A great way of capturing environmental content are panoramas (see Figure 4). We present a solution for generating high-quality stereo panoramas at megapixel resolutions. While previous approaches introduced the basic principles, we show that those techniques do not generalise well to today's high image resolutions and lead to disturbing visual artefacts. We describe the necessary correction steps and a compact representation for the input images in order to achieve a highly accurate approximation to the required ray space. In addition, we introduce a flow-based upsampling of the available input rays which effectively resolves known aliasing issues like stitching artefacts. The required rays are generated on the fly to perfectly match the desired output resolution, even for small numbers of input images. This upsampling is real-time and enables direct interactive control over the desired stereoscopic depth effect. In combination, our contributions allow the generation of stereoscopic panoramas at high output resolutions that are virtually free of artefacts such as seams, stereo discontinuities, vertical parallax and other mono-/stereoscopic shape distortions.

This work was carried out in collaboration with Yael Pritch, Henning Zimmer and Alexander Sorkine-Hornung at Disney Research Zurich. The paper has been published as an oral presentation at CVPR 2013 [20].

5.1.3. Probabilistic Connection Path Tracing

Participants: Stefan Popov, George Drettakis.

We propose an unbiased generalization of bi-directional path tracing (BPT) that significantly improves its rendering efficiency. Our main insight is that the set of paths traced by BPT contains a significant amount of statistical information, that is not exploited.

BPT repeatedly builds an eye and a light sub-paths, connects them, estimates the contribution to the corresponding pixel and then throws the path away. Instead, we propose to first trace all eye and light sub-paths, and then probabilistically connect each eye sub-path to one or more light sub-paths. From a Monte-Carlo perspective, this will connect each light to each eye sub-path, substantially increasing the number of paths used to estimate the solution. As a result, the convergence will be significantly increased as well.

This work is a collaboration with Frédo Durand from the Massachusetts Institute of Technology, Cambridge and Ravi Ramamoorthi from University of California, Berkeley in the context of the CRISP Associated Team.

5.1.4. Parallelization Strategies for Associative Image Processing Operators

Participants: Gaurav Chaurasia, George Drettakis.

../../../../projets/reves/IMG/megastereo-rooftop-anaglyph.jpg

Figure 4. A stereoscopic panorama corrected and stitched using our techniques, shown as red-cyan anaglyph image.

Basic image processing operations have been optimized on a case-by-case basis such as prefix sums and recursive filters. Moreover, these optimized algorithms are very complicated to program because parallelization involves non-trivial splitting of the input domain of the operator. The target of this is to generalize the optimization heuristics of a generic class of associative image processing operators by developing an algebraic understanding of the operator and parallelization options. The algebra can transform associative operations such as box filters, summed area table, recursive filters etc. by splitting their domain to smaller subsets of the input image that can be executed in parallel and recombine the intermediate result later. The ultimate target is to develop a compiler front-end based on the Halide language that implements this algebra and is capable of parallelizing associative operators of arbitrary footprints by a few lines of code, thereby relieving the programmer of the tedious task for programming the parallelized algorithms. Such a compiler would allow programmers to easily experiment with a plethora of parallelization strategies in a systematic manner.

This work is in collaboration with Jonathan Ragan-Kelley and Fredo Durand of MIT and Sylvain Paris (Adobe Research).

5.1.5. Lightfield Editing

Participant: Adrien Bousseau.

Lightfields capture multiple nearby views of a scene and are consolidating themselves as the successors of conventional photographs. As the field grows and evolves, the need for tools to process and manipulate lightfields arises. However, traditional image manipulation software such as Adobe Photoshop are designed to handle single views and their interfaces cannot cope with multiple views coherently. In this work we evaluate different user interface designs for lightfield editing. Our interfaces differ mainly in the way depth is presented to the user and build upon different depth perception cues.

This work is a collaboration with Adrian Jarabo, Belen Masia and Diego Gutierrez from Universidad de Zaragoza and Fabio Pellacini from Sapienza Universita di Roma.

5.2. Perception for Plausible Rendering

5.2.1. Perception of Perspective Distortions in Image-Based Rendering

Participants: Peter Vangorp, Christian Richardt, Gaurav Chaurasia, George Drettakis.

Image-based rendering (IBR) creates realistic images by enriching simple geometries with photographs, for example by mapping the photograph of a building façade onto a plane. However, as soon as the viewer moves away from the correct viewpoint, the image in the retina becomes distorted, sometimes leading to gross misperceptions of the original geometry. Two hypotheses from vision science state how viewers perceive such image distortions, one claiming that they can compensate for them (and therefore perceive scene geometry reasonably correctly), and one claiming that they cannot compensate (and therefore can perceive rather significant distortions). We modified the latter hypothesis so that it extends to street-level IBR. We then conducted a rigorous experiment that measured the magnitude of perceptual distortions that occur with IBR for façade viewing. We also conducted a rating experiment that assessed the acceptability of the distortions. The results of the two experiments were consistent with one another. They showed that viewers' percepts are indeed distorted, but not as severely as predicted by the modified vision science hypothesis. From our experimental results, we develop a predictive model of distortion for street-level IBR, which we use to provide guidelines for acceptability of virtual views and for capture camera density. We perform a confirmatory study to validate our predictions, and illustrate their use with an application that guides users in IBR navigation to stay in regions where virtual views yield acceptable perceptual distortions (see Figure 5).

This work is a collaboration with Emily Cooper and Marty Banks at UC Berkeley, within the associate team CRISP. The paper was accepted as a SIGGRAPH 2013 paper and published in the ACM Transactions on Graphics journal [18].

5.2.2. Gloss Perception in Painterly and Cartoon Rendering

Participant: Adrien Bousseau.

../../../../projets/reves/IMG/percepibr-navigation-application.png

Figure 5. This interactive navigation tool shows an inset (in the lower left) that predicts comfort ratings for all possible camera orientations as seen from the blue camera's viewpoint. The application also restricts the user's motion to regions with acceptable predicted quality (in blue and yellow).

Depictions with traditional media such as painting and drawing represent scene content in a stylized manner. It is unclear however how well stylized images depict scene properties like shape, material and lighting. In this project, we use non photorealistic rendering algorithms to evaluate how stylization alters the perception of gloss (see Figure 6). Our study reveals a compression of the range of representable gloss in stylized images so that shiny materials appear more diffuse in painterly rendering, while diffuse materials appear shinier in cartoon images.

From our measurements we estimate the function that maps realistic gloss parameters to their perception in a stylized rendering. This mapping allows users of NPR algorithms to predict the perception of gloss in their images. The inverse of this function exaggerates gloss properties to make the contrast between materials in a stylized image more faithful. We have conducted our experiment both in a lab and on a crowdsourcing website. While crowdsourcing allows us to quickly design our pilot study, a lab experiment provides more control on how subjects perform the task. We provide a detailed comparison of the results obtained with the two approaches and discuss their advantages and drawbacks for studies similar to ours.

../../../../projets/reves/IMG/npr_gloss.png

Figure 6. The experimental task used for studying gloss perception in stylized images.

This work is a collaboration with James O'Shea, Ravi Ramamoorthi and Maneesh Agrawala from UC Berkeley in the context of the Associate Team CRISP (see also Section 1) and Frédo Durand from MIT. It has been published in ACM Transactions on Graphics 2013 [11] and presented at SIGGRAPH.

5.2.3. A High-Level Visual Attention Model

Participant: George Drettakis.

The goal of this project is to develop a high-level attention model based on memory schemas and singleton theory in visual perception. We have developed an approach extending a Bayesian approach to attention, which incorporates these high level features and can be directly used in a game engine to improve scene design.

This project is in collaboration with the Tech. University of Crete in the context of the Ph.D. of George Koulieris, supervised by Prof. Katerina Mania and BTU Cottburg (D. Cunningham).

5.3. Interaction and Design for Virtual Environments

5.3.1. Diffusion Curves: A Vector Representation for Smooth-Shaded Images

Participant: Adrien Bousseau.

This paper was selected for presentation in the Communications of the ACM, as an important graphics research result of interest to the entire Computer Science community. We describe a new vector-based primitive for creating smooth-shaded images, called the diffusion curve. A diffusion curve partitions the space through which it is drawn, defining different colors on either side. These colors may vary smoothly along the curve. In addition, the sharpness of the color transition from one side of the curve to the other can be controlled. Given a set of diffusion curves, the final image is constructed by solving a Poisson equation whose constraints are specified by the set of gradients across all diffusion curves (Figure 7). Like all vector-based primitives, diffusion curves conveniently support a variety of operations, including geometry-based editing, keyframe animation, and ready stylization. Moreover, their representation is compact and inherently resolution independent. We describe a GPU-based implementation for rendering images defined by a set of diffusion curves in real time. We then demonstrate an interactive drawing system allowing artists to create artwork using diffusion curves, either by drawing the curves in a freehand style, or by tracing existing imagery. Furthermore, we describe a completely automatic conversion process for taking an image and turning it into a set of diffusion curves that closely approximate the original image content.

This work is a collaboration with Alexandrina Orzan, Pascal Barla (Inria / Manao), Holger Winnemöller (Adobe Systems), Joëlle Thollot (Inria / Maverick) and David Salesin (Adobe Systems). This work was originally published in ACM Transactions on Graphics (Proceeding of SIGGRAPH 2008) and was selected for publication in Communications of the ACM July 2013 [15].

5.3.2. Natural Gesture-based Interaction for Complex Tasks in an Immersive Cube

Participants: Emmanuelle Chapoulie, George Drettakis.

We present a solution for natural gesture interaction in an immersive cube in which users can manipulate objects with fingers of both hands in a close-to-natural manner for moderately complex, general purpose tasks. Our solution uses finger tracking coupled with a real-time physics engine, combined with a comprehensive approach for hand gestures, which is robust to tracker noise and simulation instabilities. To determine if our natural gestures are a feasible interface in an immersive cube, we perform an exploratory study for tasks involving the user walking in the cube while performing complex manipulations such as balancing objects. We compare gestures to a traditional 6-DOF Wand, and we also compare both gestures and Wand with the same task, faithfully reproduced in the real world. Users are also asked to perform a free task, allowing us to observe their perceived level of presence in the scene. Our results show that our robust approach provides a feasible natural gesture interface for immersive cube-like environments and is perceived by users as being closer to the real experience compared to the Wand.

../../../../projets/reves/IMG/diffusion_curves.jpg

Figure 7. A diffusion curve consists of a Bézier curve (a) enriched with color (b) and blur (c) control points. The final image (d) is obtained by diffusing the colors in the image domain.

This work is a collaboration with Jean-Christophe Lombardo of SED, with Evanthia Dimara and Maria Roussou from the University of Athens and with Maud Marchal from IRISA-INSA/Inria Rennes - Bretagne Atlantique. The work is under review in the journal Virtual Reality.

5.3.3. *Evaluation of Direct Manipulation in an Immersive Cube: a Controlled Study*

Participants: Emmanuelle Chapoulie, George Drettakis.

We are pursuing a study for interaction using finger tracking and traditional 6 degrees of freedom (DOF) flysticks in a virtual reality immersive cube. Our study aims at identifying which factors make one interface better than the other and which are the tradeoffs for the design of experiments, thus decomposing the movements into restricted DOF.

5.3.4. *The Drawing Assistant: Automated Drawing Guidance and Feedback from Photographs*

Participants: Emmanuel Iarussi, Adrien Bousseau.

Drawing is the earliest form of visual depiction and continues to enjoy great popularity with paint systems. However, drawing requires artistic skills that many people feel out of reach. We developed an interactive drawing tool that provides automated guidance over model photographs to help people practice traditional drawing-by-observation techniques. The drawing literature describes a number of techniques to help people gain consciousness of the shapes in a scene and their relationships. We compile these techniques and derive a set of construction lines that we automatically extract from a model photograph (see Figure 8). We then display these lines over the model to guide its manual reproduction by the user on the drawing canvas. Our pen-based interface also allows users to navigate between the techniques they wish to practice and to draw construction lines in dedicated layers. We use shape-matching to register the user's sketch with the model guides. We use this registration to provide corrective feedback to the user. We conducted two user studies to inform the design of our tool and evaluate our approach with a total of 20 users. Participants produced better drawings using the drawing assistant, with more accurate proportions and alignments. They also perceived that guidance and corrective feedback helped them better understand how to draw. Finally, some participants spontaneously applied the techniques when asked to draw without our tool after using it for about 30 minutes.

This work is a collaboration with Theophanis Tsandilas from the InSitu project team - Inria Saclay, in the context of the ANR DRAO project. It has been published at proceedings of UIST 2013 the 26th annual ACM symposium on User interface software and technology [19].

5.3.5. *Shape-Aware Sketch Editing with Covariant-Minimizing Cross Fields*

Participants: Emmanuel Iarussi, Adrien Bousseau.

Free-hand sketches are extensively used in product design for their ability to convey 3D surfaces with a handful of pen strokes. Skillful artists capture all surface information by strategically positioning strokes so that they depict the feature lines and curvature directions of surface patches. Viewers envision the intended 3D surface by mentally interpolating these lines to form a dense network representative of the curvature of the shape. Our goal is to mimic this interpolation process to estimate at each pixel of a sketch the projection of the two principal directions of the surface, or their extrapolation over umbilic regions. While the information we recover is purely 2D, it provides a vivid sense of the intended 3D surface and allows various shape-aware sketch editing applications, including normal estimation for shading, cross-hatching rendering and surface parameterization for texture mapping.

This work is a collaboration with David Bommes from the Titane project team, Sophia-Antipolis.

5.3.6. *Depicting Stylized Materials with Vector Shade Trees*

Participants: Jorge Lopez-Moreno, Stefan Popov, Adrien Bousseau, George Drettakis.

../../../../projets/reves/IMG/drawingAssistantFigure.png

Figure 8. Our drawing assistant provides guidance and feedback over a model photograph that the user reproduces on a virtual canvas (a). We use computer vision algorithms to extract visual guides that enhance the geometric structures in the image (b). In this example, the user first sketched the block-in construction lines (c, blue) before drawing the regions and adding details. This guidance helps users produce more accurate drawings.

../../../../projets/reves/IMG/teaser_v2.png

Figure 9. Depicting Stylized Materials with Vector Shade Trees.

Vector graphics represent images with compact, editable and scalable primitives. Skillful vector artists employ these primitives to produce vivid depictions of material appearance and lighting. However, such stylized imagery often requires building complex multi-layered combinations of colored fills and gradient meshes. We facilitate this task by introducing vector shade trees that bring to vector graphics the flexibility of modular shading representations as known in the 3D rendering community. In contrast to traditional shade trees that combine pixel and vertex shaders, our shade nodes encapsulate the creation and blending of vector primitives that vector artists routinely use. We propose a set of basic shade nodes that we design to respect the traditional guidelines on material depiction described in drawing books and tutorials. We integrate our representation as an Adobe Illustrator plug-in that allows even inexperienced users to take a line drawing, apply a few clicks and obtain a fully colored illustration. More experienced artists can easily refine the illustration, adding more details and visual features, while using all the vector drawing tools they are already familiar with. We demonstrate the power of our representation by quickly generating illustrations of complex objects and materials.

Figure 9 illustrates how our algorithm works. We use a combination of basic shade nodes composed of vector graphics primitives to describe Vector Shade Trees that represent stylized materials (a). Combining these nodes allows the depiction of a variety of materials while preserving traditional vector drawing style and practice. We integrate our vector shade trees in a vector drawing tool that allows users to apply stylized shading effects on vector line drawings (b,c).

This work is a collaboration with Maneesh Agrawala from University of California, Berkeley in the context of the CRISP Associated Team. The work was accepted as a SIGGRAPH 2013 paper and published in ACM Transactions on Graphics, volume 32, issue 4 [14].

5.3.7. Auditory-Visual Aversive Stimuli Modulate the Conscious Experience of Fear

Participants: Rachid Guerchouche, George Drettakis.

In a natural environment, affective information is perceived via multiple senses, mostly audition and vision. However, the impact of multisensory information on affect remains relatively undiscovered. In this study, we investigated whether the auditory-visual presentation of aversive stimuli influences the experience of fear. We used the advantages of virtual reality to manipulate multisensory presentation and to display potentially fearful dog stimuli embedded in a natural context. We manipulated the affective reactions evoked by the dog stimuli by recruiting two groups of participants: dog-fearful and non-fearful participants. The sensitivity to dog fear was assessed psychometrically by a questionnaire and also at behavioral and subjective levels using a Behavioral Avoidance Test (BAT). Participants navigated in virtual environments, in which they encountered virtual dog stimuli presented through the auditory channel, the visual channel or both. They were asked to report their fear using Subjective Units of Distress. We compared the fear for unimodal (visual or auditory) and bimodal (auditory-visual) dog stimuli. Dog-fearful participants as well as non-fearful participants reported more fear in response to bimodal audiovisual compared to unimodal presentation of dog stimuli. These results suggest that fear is more intense when the affective information is processed via multiple sensory pathways, which might be due to a cross-modal potentiation. Our findings have implications for the field of virtual reality-based therapy of phobias. Therapies could be refined and improved by implicating and manipulating the multisensory presentation of the feared situations.

This work is a collaboration with Marine Taffou and Isabelle Viaud-Delmon from CNRS-IRCAM, in the context of the European project VERVE. The work was published in the Multisensory Research journal 2013 [17].

5.3.8. Memory Motivation Virtual Experience

Participants: Emmanuelle Chapoulie, Rachid Guerchouche, George Drettakis.

Memory complaints are known to be one the first stages of Alzheimer's disease, for which -up to now, there is no known chemical treatment. In the context of the European project VERVE, and in collboartion with the Resources and Research Memory Centre of Nice Hospital (CM2R), we performed a study on the feasibility of treating memory complaints using realistic immersive virtual environments. Such environments are created

../../../../projets/reves/IMG/DogPhobia_Garden.jpg ../../../../projets/reves/IMG/DogPhobia_Hanga

Figure 10. Pictures of the auditory-visual VEs used to measure the participants' fear when encountering virtual dogs. On the left, the outdoor garden scene and on the right, the indoor hangar scene.

using Image-Based Rendering technique developed by REVES. It is possible to easily provide, realistic 3D environments of places familiar to the participants using only a few photograph, and investigate whether IBR virtual environments can convey familiarity.

This work is a collaboration with Pierre-David Petit and Pr. Philippe Robert from CM2R. The work will be presented in IEEE Virtual Reality conference 2014 and will be published in the conference proceedings.

5.3.9. Layered Image Vectorization

Participants: Christian Richardt, Adrien Bousseau, George Drettakis.

Vector graphics enjoy great popularity among graphic designers for their compactness, scalability and editability. The goal of *vectorization* algorithms is to facilitate the creation of vector graphics by converting bitmap images into vector primitives. However, while a vectorization algorithm should faithfully reproduce the appearance of a bitmap image, it should also generate vector primitives that are easily editable – a goal that existing methods have largely overlooked. We investigate layered representations which are more compact and editable, and hence better preserve the strengths of vector graphics. This work is in collaboration with Maneesh Agrawala in the context of the CRISP Associated Team and Jorge Lopez-Moreno, now a postdoc at the University of Madrid.

5.3.10. True2Form: Automatic 3D Concept Modeling from Design Sketches

Participants: Adrien Bousseau.

We developed a method to estimate smooth 3D shapes from design sketches. We do this by hypothesizing and perceptually validating a set of local geometric relationships between the curves in sketches. We then algorithmically reconstruct 3D curves from a single sketch by detecting their local geometric relationships and reconciling them globally across the 3D curve network.

This work is a collaboration with James McCrae and Karan Singh from the University of Toronto and Xu Baoxuan, Will Chang and Alla Sheffer from the University of British Columbia.

RMOD Project-Team

6. New Results

6.1. Tools for understanding applications: IDEs and Visualization

Performance Evolution Blueprint: Understanding the Impact of Software Evolution on Performance. Understanding the root of a performance drop or improvement requires analyzing different program executions at a fine grain level. Such an analysis involves dedicated profiling and representation techniques. JProfiler and YourKit, two recognized code profilers fail, on both providing adequate metrics and visual representations, conveying a false sense of the performance variation root. We propose performance evolution blueprint, a visual support to precisely compare multiple software executions. Our blueprint is offered by Rizel, a code profiler to efficiently explore performance of a set of benchmarks against multiple software revisions. [31]

Seamless Composition and Reuse of Customizable User Interfaces with Spec Implementing UIs is often a tedious task. To address this, UI Builders have been proposed to support the description of widgets, their location, and their logic. A missing aspect of UI Builders is however the ability to reuse and compose widget logic. In our experience, this leads to a significant amount of duplication in UI code. To address this issue, we built Spec: a UIBuilder for Pharo with a focus on reuse. With Spec, widget properties are defined declaratively and attached to specific classes known as composable classes. A composable class defines its own widget description as well as the model-widget bridge and widget interaction logic. Spec enables seamless reuse of widgets, its use in Pharo 2.0 has cut in half the amount of lines of code of six of its tools, mostly through reuse. This shows that Spec meets its goals of allowing reuse and composition of widget logic. [17]

Pragmatic Visualizations for Roassal: a Florilegium Software analysis and in particular reverse engineering often involves a large amount of structured data. This data should be presented in a meaningful form so that it can be used to improve software artefacts. The software analysis community has produced numerous visual tools to help understand different software elements. However, most of the visualization techniques, when applied to software elements, produce results that are difficult to interpret and comprehend. We present five graph layouts that are both expressive for polymetric views and agnostic to the visualization engine. These layouts favor spatial space reduction while emphasizing on clarity. Our layouts have been implemented in the Roassal visualization engine and are available under the MIT License. [23]

6.2. Software Quality: Bugs and Debuggers

BugMaps-Granger: A Tool for Causality Analysis between Source Code Metrics and Bugs. Despite the increasing number of bug analysis tools for exploring bugs in software systems, there are no tools supporting the investigation of causality relationships between internal quality metrics and bugs. We propose an extension of the BugMaps tool called BugMaps-Granger that allows the analysis of source code properties that caused bugs. For this purpose, we relied on Granger Causality Test to evaluate whether past changes to a given time series of source code metrics can be used to forecast changes in a time series of defects. Our tool extracts source code versions from version control platforms, generates source code metrics and defects time series, computes Granger, and provides interactive visualizations for causal analysis of bugs. We also provide a case study in order to evaluate the tool. [22]

Mining Architectural Patterns Using Association Rules

Software systems usually follow many programming rules prescribed in an architectural model. However, developers frequently violate these rules, introducing architectural drifts in the source code. We present a data mining approach for architecture conformance based on a combination of static and historical software analysis. For this purpose, the proposed approach relies on data mining techniques to extract structural and historical architectural patterns. In addition, we propose a methodology that uses the extracted patterns to detect both absences and divergences in source-code based architectures. We applied the proposed approach in an industrial strength system. As a result we detected 137 architectural violations, with an overall precision of 41.02%. [27]

Heuristics for Discovering Architectural Violations

Software architecture conformance is a key software quality control activity that aims to reveal the progressive gap normally observed between concrete and planned software architecture. We present ArchLint, a lightweight approach for architecture conformance based on a combination of static and historical source code analysis. For this purpose, ArchLint relies on four heuristics for detecting both absences and divergences in source code based architectures. We applied ArchLint in an industrial-strength system and as a result we detected 119 architectural violations, with an overall precision of 46.7% and a recall of 96.2%, for divergences. We also evaluated ArchLint with four open-source systems, used in an independent study on reflexion models. In this second study, ArchLint achieved precision results ranging from 57.1% to 89.4%. [26]

6.3. Software Quality: History and Changes

Representing Code History with Development Environment Events. Modern development environments handle information about the intent of the programmer: for example, they use abstract syntax trees for providing high-level code manipulation such as refactorings; nevertheless, they do not keep track of this information in a way that would simplify code sharing and change understanding. In most Smalltalk systems, source code modifications are immediately registered in a transaction log often called a ChangeSet. Such mechanism has proven reliability, but it has several limitations. We analyse such limitations and describe scenarios and requirements for tracking fine-grained code history with a semantic representation. We want to enrich code sharing with extra information from the IDE, which will help understanding the intention of the changes and let a new generation of tools act in consequence. [24]

Mining System Specific Rules from Change Patterns A significant percentage of warnings reported by tools to detect coding standard violations are false positives. Thus, there are some works dedicated to provide better rules by mining them from source code history, analyzing bug-fixes or changes between system releases. However, software evolves over time, and during development not only bugs are fixed, but also features are added, and code is refactored. In such cases, changes must be consistently applied in source code to avoid maintenance problems. We propose to extract system specific rules by mining systematic changes over source code history, i.e., not just from bug-fixes or system releases, to ensure that changes are consistently applied over source code. We focus on structural changes done to support API modification or evolution with the goal of providing better rules to developers. Also, rules are mined from predefined rule patterns that ensure their quality. In order to assess the precision of such specific rules to detect real violations, we compare them with generic rules provided by tools to detect coding standard violations on four real world systems covering two programming languages. The results show that specific rules are more precise in identifying real violations in source code than generic ones, and thus can complement them. [25]

6.4. Reconciling Dynamic Languages and Isolation

Virtual Smalltalk Images: Model and Applications. Reflective architectures are a powerful solution for code browsing, debugging or in-language process handling. However, these reflective architectures show some limitations in edge cases of self-modification and self-monitoring. Modifying the modifier process or monitoring the monitor process in a reflective system alters the system itself, leading to the impossibility to perform some of those tasks properly. We analyze the problems of reflective architectures in the context of image based object-oriented languages and solve them by providing a first-class representation of an image: a virtualized image. We present Oz, our virtual image solution. In Oz, a virtual image is represented by an object space. Through an object space, an image can manipulate the internal structure and control the execution of other images. An Oz object space allows one to introspect and modify execution information such as processes, contexts, existing classes and objects. We show how Oz solves the edge cases of reflective architectures by adding a third participant, and thus, removing the self modification and self-observation constraints. [30]

Bootstrapping Reflective Systems: The Case of Pharo. Bootstrapping is a technique commonly known by its usage in language definition by the introduction of a compiler written in the same language it compiles. This process is important to understand and modify the definition of a given language using the same language, taking benefit of the abstractions and expression power it provides. A bootstrap, then, supports the evolution of a language. However, the infrastructure of reflective systems like Smalltalk includes, in addition to a compiler, an environment with several self-references. A reflective system bootstrap should consider all its infrastructural components. We propose a definition of bootstrap for object-oriented reflective systems, we describe the architecture and components it should contain and we analyze the challenges it has to overcome. Finally, we present a reference bootstrap process for a reflective system and Hazelnut, its implementation for bootstrapping the Pharo Smalltalk-inspired system. [15]

Object Graph Isolation with Proxies More and more software systems are now made of multiple collaborating third-party components. Enabling fine-grained control over the communication between components becomes a major requirement. While software isolation has been studied for a long time in operating systems (OS), most programming languages lack support for isolation. In this context we explore the notion of proxy. A proxy is a surrogate for another object that controls access to this object. We are particularly interested in generic proxy implementations based on language-level reflection. We present an analysis that shows how these reflective proxies can propagate a security policy thanks to the transitive wrapping mechanism. We present a prototype implementation that supports transitive wrapping and allows a fine-grained control over an isolated object graph. [33]

6.5. Dynamic Languages: Compilers

Towards a flexible Pharo Compiler The Pharo Smalltalk-inspired language and environment started its development with a codebase that can be traced back to the original Smalltalk-80 release from 1983. Over the last years, Pharo has been used as the basis of many research projects. Often these experiments needed changes related to the compiler infrastructure. However, they did not use the existing compiler and instead implemented their own experimental solutions. This shows that despite being an impressive achievement considering its age of over 35 years, the compiler infrastructure needs to be improved. We identify three problems: (i) The architecture is not reusable, (ii) compiler can not be parametrized and (iii) the mapping between source code and bytecode is overly complex. Solving these problems will not only help researchers to develop new language features, but also the enhanced power of the infrastructure allows many tools and frameworks to be built that are important even for day-to-day development, such as debuggers and code transformation tools. [20]

Gradual Typing for Smalltalk Being able to combine static and dynamic typing within the same language has clear benefits in order to support the evolution of prototypes or scripts into mature robust programs. While being an emblematic dynamic object-oriented language, Smalltalk is lagging behind in this regard. We report on the design, implementation and application of Gradualtalk, a gradually-typed Smalltalk meant to enable incremental typing of existing programs. The main design goal of the type system is to support the features of the Smalltalk language, like metaclasses and blocks, live programming, and to accomodate the programming idioms used in practice. We studied a number of existing projects in order to determine the features to include in the type system. As a result, Gradualtalk is a practical approach to gradual types in Smalltalk, with a novel blend of type system features that accomodate most programming idioms. [13]

ROMA Team

6. New Results

6.1. Scheduling tree-shaped task graphs to minimize memory and makespan

In this work [37], we investigate the execution of tree-shaped task graphs using multiple processors. Each edge of such a tree represents a large IO file. A task can only be executed if all input and output files fit into memory, and a file can only be removed from memory after it has been consumed. Such trees arise, for instance, in the multifrontal method of sparse matrix factorization. The maximum amount of memory needed depends on the execution order of the tasks. With one processor the objective of the tree traversal is to minimize the required memory. This problem was well studied and optimal polynomial algorithms were proposed. Here, we extend the problem by considering multiple processors, which is of obvious interest in the application area of matrix factorization. With the multiple processors comes the additional objective to minimize the time needed to traverse the tree, i.e., to minimize the makespan. Not surprisingly, this problem proves to be much harder than the sequential one. We study the computational complexity of this problem and provide an inapproximability result even for unit weight trees. Several heuristics are proposed, each with a different optimization focus, and they are analyzed in an extensive experimental evaluation using realistic trees.

6.2. Model and complexity results for tree traversals on hybrid platforms

In this work [35], we study the complexity of traversing tree-shaped workflows whose tasks require large I/O files. We target a heterogeneous architecture with two resources of different types, where each resource has its own memory, such as a multicore node equipped with a dedicated accelerator (FPGA or GPU). Tasks in the workflow are tagged with the type of resource needed for their processing. Besides, a task can be processed on a given resource only if all its input files and output files can be stored in the corresponding memory. At a given execution step, the amount of data stored in each memory strongly depends upon the ordering in which the tasks are executed, and upon when communications between both memories are scheduled. The objective is to determine an efficient traversal that minimizes the maximum amount of memory of each type needed to traverse the whole tree. In this work, we establish the complexity of this two-memory scheduling problem, provide inapproximability results, and show how to determine the optimal depth-first traversal. Altogether, these results lay the foundations for memory-aware scheduling algorithms on heterogeneous platforms.

6.3. On the combination of silent error detection and checkpointing

In this work [19], we revisit traditional checkpointing and rollback recovery strategies, with a focus on silent data corruption errors. Contrarily to fail-stop failures, such latent errors cannot be detected immediately, and a mechanism to detect them must be provided. We consider two models: (i) errors are detected after some delays following a probability distribution (typically, an Exponential distribution); (ii) errors are detected through some verification mechanism. In both cases, we compute the optimal period in order to minimize the waste, i.e., the fraction of time where nodes do not perform useful computations. In practice, only a fixed number of checkpoints can be kept in memory, and the first model may lead to an irrecoverable failure. In this case, we compute the minimum period required for an acceptable risk. For the second model, there is no risk of irrecoverable failure, owing to the verification mechanism, but the corresponding overhead is included in the waste. Finally, both models are instantiated using realistic scenarios and application/architecture parameters.

6.4. Checkpointing algorithms and fault prediction

In this series of work [22], [49], we deal with the impact of fault prediction techniques on checkpointing strategies, when the fault-prediction system provides either prediction windows or exact predictions. We extend the classical first-order analysis of Young and Daly in the presence of a fault prediction system, characterized by its recall and its precision. In this framework, we provide optimal algorithms to decide whether and when to take predictions into account, and we derive the optimal value of the checkpointing period. These results allow us to analytically assess the key parameters that impact the performance of fault predictors at very large scale.

6.5. Mapping applications on volatile resources

In this series of work [12], [27], [28], we study the execution of iterative applications on volatile processors such as those found on desktop grids. We envision two models, one where all tasks are assumed to be independent, and another where all tasks are tightly coupled and keep exchanging information throughout the iteration. These two models cover the two extreme points of the parallelization spectrum. We develop master-worker scheduling schemes that attempt to achieve good trade-offs between worker speed and worker availability. Any iteration entails the execution of a fixed number of independent tasks or of tightly-coupled tasks. A key feature of our approach is that we consider a communication model where the bandwidth capacity of the master for sending application data to workers is limited. This limitation makes the scheduling problem more difficult both in a theoretical sense and in a practical sense. Furthermore, we consider that a processor can be in one of three states: available, down, or temporarily preempted by its owner. This preempted state also complicates the scheduling problem. In practical settings, e.g., desktop grids, master bandwidth is limited and processors are temporarily reclaimed. Consequently, addressing the aforementioned difficulties is necessary for successfully deploying master-worker applications on volatile platforms. Our first contribution is to determine the complexity of the scheduling problems in their offline versions, i.e., when processor availability behaviors are known in advance. Even with this knowledge, the problems are NP-hard. Our second contribution is an evaluation of the expectation of the time needed by a worker to complete a set of tasks. We obtain a close formula for independent tasks and an analytical approximation for tightly-coupled tasks. Those evaluations rely on a Markovian assumption for the temporal availability of processors, and are at the heart of some heuristics that aim at favoring “reliable” processors in a sensible manner. Our third contribution is a set of heuristics for both models, which we evaluate in simulation. Our results provide guidance to selecting the best strategy as a function of processor state availability versus average task duration.

6.6. Using group replication for resilience on exascale systems

High performance computing applications must be resilient to faults. The traditional fault-tolerance solution is checkpoint-recovery, by which application state is saved to and recovered from secondary storage throughout execution. It has been shown that, even when using an optimal checkpointing strategy, the checkpointing overhead precludes high parallel efficiency at large scale. Additional fault-tolerance mechanisms must thus be used. Such a mechanism is replication, i.e., multiple processors performing the same computation so that a processor failure does not necessarily imply an application failure. In spite of resource waste, replication can lead to higher parallel efficiency when compared to using only checkpoint-recovery at large scale. In this work [11], we propose to execute and checkpoint multiple application instances concurrently, an approach we term group replication. For Exponential failures we give an upper bound on the expected application execution time. This bound corresponds to a particular checkpointing period that we derive. For general failures, we propose a dynamic programming algorithm to determine non-periodic checkpoint dates as well as an empirical periodic checkpointing solution whose period is found via a numerical search. Using simulation we evaluate our proposed approaches, including comparison to the non-replication case, for both Exponential and Weibull failure distributions. Our broad finding is that group replication is useful in a range of realistic application and checkpointing overhead scenarios for future exascale platforms.

6.7. Unified model for assessing checkpointing protocols at extreme-scale

In this work [10], we present a unified model for several well-known checkpoint/restart protocols. The proposed model is generic enough to encompass both extremes of the checkpoint/restart space, from coordinated approaches to a variety of uncoordinated checkpoint strategies (with message logging). We identify a set of crucial parameters, instantiate them and compare the expected efficiency of the fault tolerant protocols, for a given application/platform pair. We then propose a detailed analysis of several scenarios, including some of the most powerful currently available HPC platforms, as well as anticipated Exascale designs. The results of this analytical comparison are corroborated by a comprehensive set of simulations. Altogether, they outline comparative behaviors of checkpoint strategies at very large scale, thereby providing insight that is hardly accessible to direct experimentation.

6.8. Revisiting the double checkpointing algorithm

In this work [33], we study fast checkpointing algorithms which require distributed access to stable storage. This work revisits the approach base upon double checkpointing, and compares the blocking algorithm of Zheng, Shi, and Kalé, with the non-blocking algorithm of Ni, Meneses, and Kalé in terms of both performance and risk. We also extend the model that they have proposed to assess the impact of the overhead associated to non-blocking communications. We then provide a new peer-to-peer checkpointing algorithm, called the triple checkpointing algorithm, that can work at constant memory, and achieves both higher efficiency and better risk handling than the double checkpointing algorithm. We provide performance and risk models for all the evaluated protocols, and compare them through comprehensive simulations.

6.9. Multi-criteria checkpointing strategies: Optimizing response-time versus resource utilization

Failures are increasingly threatening the efficiency of HPC systems, and current projections of Exascale platforms indicate that rollback recovery, the most convenient method for providing fault tolerance to general-purpose applications, reaches its own limits at such scales. One of the reasons explaining this unnerving situation comes from the focus that has been given to per-application completion time, rather than to platform efficiency. In this work [26], we discuss the case of uncoordinated rollback recovery where the idle time spent waiting recovering processors is used to progress a different, independent application from the system batch queue. We then propose an extended model of uncoordinated checkpointing that can discriminate between idle time and wasted computation. We instantiate this model in a simulator to demonstrate that, with this strategy, uncoordinated checkpointing per application completion time is unchanged, while it delivers near-perfect platform efficiency.

6.10. Optimal checkpointing period: Time vs. energy

In this work [18], we deal with parallel scientific applications using non-blocking and periodic coordinated checkpointing to enforce resilience. We provide a model and detailed formulas for total execution time and consumed energy. We characterize the optimal period for both objectives, and we assess the range of time/energy trade-offs to be made by instantiating the model with a set of realistic scenarios for Exascale systems. We give a particular emphasis to I/O transfers, because the relative cost of communication is expected to dramatically increase, both in terms of latency and consumed energy, for future Exascale platforms.

6.11. Energy-aware checkpointing of divisible tasks with soft or hard deadlines

In this work [20], we aim at minimizing the energy consumption when executing a divisible workload under a bound on the total execution time, while resilience is provided through checkpointing. We discuss several variants of this multi-criteria problem. Given the workload, we need to decide how many chunks to use, what are the sizes of these chunks, and at which speed each chunk is executed. Furthermore, since a failure may occur during the execution of a chunk, we also need to decide at which speed a chunk should be re-executed in the event of a failure. The goal is to minimize the expectation of the total energy consumption, while enforcing a deadline on the execution time, that should be met either in expectation (soft deadline), or in the worst case (hard deadline). For each problem instance, we propose either an exact solution, or a function that can be optimized numerically. The different models are then compared through an extensive set of experiments.

6.12. Assessing the performance of energy-aware mappings

In this work [8], we aim at mapping streaming applications that can be modeled by a series-parallel graph onto a 2-dimensional tiled chip multiprocessor (CMP) architecture. The objective of the mapping is to minimize the energy consumption, using dynamic voltage and frequency scaling (DVFS) techniques, while maintaining a given level of performance, reflected by the rate of processing the data streams. This mapping problem turns out to be NP-hard, and several heuristics are proposed. We assess their performance through comprehensive simulations using the StreamIt workflow suite and randomly generated series-parallel graphs, and various CMP grid sizes.

6.13. Computing the throughput of probabilistic and replicated streaming applications

In this work [7], we investigate how to compute the throughput of probabilistic and replicated streaming applications. We are given (i) a streaming application whose dependence graph is a linear chain; (ii) a one-to-many mapping of the application onto a fully heterogeneous target platform, where a processor is assigned at most one application stage, but where a stage can be replicated onto a set of processors; and (iii) a set of random variables modeling the computation and communication times in the mapping. We show how to compute the throughput of the application, i.e., the rate at which data sets can be processed, under two execution models, the Strict model where the actions of each processor are sequentialized, and the Overlap model where a processor can compute and communicate in parallel. The problem is easy when application stages are not replicated, i.e., assigned to a single processor: in that case the throughput is dictated by the critical hardware resource. However, when stages are replicated, i.e., assigned to several processors, the problem becomes surprisingly complicated: even in the deterministic case, the optimal throughput may be lower than the smallest internal resource throughput. The first contribution of this work is to provide a general method to compute the throughput when mapping parameters are constant or follow I.I.D. exponential laws. The second contribution is to provide bounds for the throughput when stage parameters (computation and communication times) form associated random sequences, and are N.B.U.E. (New Better than Used in Expectation) variables: the throughput is bounded from below by the exponential case and bounded from above by the deterministic case. An extensive set of simulation allows us to assess the quality of the model, and to observe the actual behavior of several distributions.

6.14. Reliability and performance optimization of pipelined real-time systems

In this work [6], we consider pipelined real-time systems that consist of a chain of tasks executing on a distributed platform. The processing of the tasks is pipelined: each processor executes only one interval of consecutive tasks. We are interested in minimizing both the input-output latency and the period of application mapping. For dependability reasons, we are also interested in maximizing the reliability of the system. We therefore assign several processors to each interval of tasks, so as to increase the reliability of the system. Both processors and communication links are unreliable and subject to transient failures. We assume that the arrival of the failures follows a constant parameter Poisson law, and that the failures are statistically independent events. We study several variants of this multiprocessor mapping problem, with several hypotheses on the target platform (homogeneous/heterogeneous speeds and/or failure rates). We provide NP-hardness complexity results, and optimal mapping algorithms for polynomial problem instances. Efficient heuristics are presented to solve the general case, and experimental results are provided.

6.15. Scheduling linear chain streaming applications on heterogeneous systems with failures

In this work [5], we study the problem of optimizing the throughput of streaming applications for heterogeneous platforms subject to failures. Applications are linear graphs of tasks (pipelines), with a type associated to each task. The challenge is to map each task onto one machine of a target platform, each machine having to be specialized to process only one task type, given that every machine is able to process all the types before being specialized in order to avoid costly setups. The objective is to maximize the throughput, i.e., the rate at which jobs can be processed when accounting for failures. Each instance can thus be performed by any machine specialized in its type and the workload of the system can be shared among a set of specialized machines. For identical machines, we prove that an optimal solution can be computed in polynomial time. However, the problem becomes NP-hard when two machines may compute the same task type at different speeds. Several polynomial time heuristics are designed for the most realistic specialized settings. Simulation results assess their efficiency, showing that the best heuristics obtain a good throughput, much better than the throughput obtained with a random mapping. Moreover, the throughput is close to the optimal solution in the particular cases where the optimal throughput can be computed.

6.16. A survey of pipelined workflow scheduling: Models and algorithms

In this survey [4], we consider a large class of applications that need to execute the same workflow on different data sets of identical size. Efficient execution of such applications necessitates intelligent distribution of the application components and tasks on a parallel machine, and the execution can be orchestrated by utilizing task-, data-, pipelined-, and/or replicated-parallelism. The scheduling problem that encompasses all of these techniques is called pipelined workflow scheduling, and it has been widely studied in the last decade. Multiple models and algorithms have flourished to tackle various programming paradigms, constraints, machine behaviors or optimization goals. This work surveys the field by summing up and structuring known results and approaches.

6.17. Reclaiming the energy of a schedule: Models and algorithms

In this work [1], we consider a task graph to be executed on a set of processors. We assume that the mapping is given, say by an ordered list of tasks to execute on each processor, and we aim at optimizing the energy consumption while enforcing a prescribed bound on the execution time. Although it is not possible to change the allocation of a task, it is possible to change its execution speed. Rather than using a local approach such as backfilling, we consider the problem as a whole and study the impact of several speed variation models on its complexity. For continuous speeds, we give a closed-form formula for trees and series-parallel graphs, and we cast the problem into a geometric programming problem for general directed acyclic graphs. We show that the classical dynamic voltage and frequency scaling (DVFS) model with discrete modes leads to an NP-complete problem, even if the modes are regularly distributed (an important particular case in practice, which we analyze as the incremental model). On the contrary, the Vdd-hopping model that allows to switch between different supply voltages (VDD) while executing a task leads to a polynomial solution. Finally, we provide an approximation algorithm for the incremental model, which we extend for the general DVFS model.

6.18. Non-clairvoyant reduction algorithms for heterogeneous platforms

In this work [24], we revisit the classical problem of the reduction collective operation in a heterogeneous environment. We discuss and evaluate four algorithms that are non-clairvoyant, i.e., they do not know in advance the computation and communication costs. On the one hand, Binomial-stat and Fibonacci-stat are static algorithms that decide in advance which operations will be reduced, without adapting to the environment; they were originally defined for homogeneous settings. On the other hand, Tree-dyn and Non-Commut-Tree-dyn are fully dynamic algorithms, for commutative or non-commutative reductions. With identical computation costs, we show that these algorithms are approximation algorithms with constant or asymptotic ratios. When costs are exponentially distributed, we perform an analysis of Tree-dyn based on Markov chains.

Finally, we assess the relative performance of all four non-clairvoyant algorithms with heterogeneous costs through a set of simulations.

6.19. Non-linear divisible loads: There is no free lunch

Divisible Load Theory (DLT) has received a lot of attention in the past decade. A divisible load is a perfect parallel task, that can be split arbitrarily and executed in parallel on a set of possibly heterogeneous resources. The success of DLT is strongly related to the existence of many optimal resource allocation and scheduling algorithms, what strongly differs from general scheduling theory. Moreover, recently, close relationships have been underlined between DLT, that provides a fruitful theoretical framework for scheduling jobs on heterogeneous platforms, and MapReduce, that provides a simple and efficient programming framework to deploy applications on large scale distributed platforms.

The success of both have suggested to extend their framework to non-linear complexity tasks. In this work [23], we show that both DLT and MapReduce are better suited to workloads with linear complexity. In particular, we prove that divisible load theory cannot directly be applied to quadratic workloads, such as it has been proposed recently. We precisely state the limits for classical DLT studies and we review and propose solutions based on a careful preparation of the dataset and clever data partitioning algorithms. In particular, through simulations, we show the possible impact of this approach on the volume of communications generated by MapReduce, in the context of Matrix Multiplication and Outer Product algorithms.

6.20. Direct solvers for sparse linear systems

This work is closely related to the MUMPS solver (see Section 5.1) and was performed in close collaboration with INPT (Toulouse). First, we have pursued the study of low-rank representations to speed-up sparse direct solvers using the so called BLR (Block Low Rank) format [44]. This work was done in collaboration with LSTC (Livermore Software Technology Corp., USA) and in the context of a contract with EDF which funded the PhD thesis of Clément Weisbecker at INPT. We also worked on shared-memory parallelism [61] in the context of the PhD thesis of Wissam M. Sid-Lakhdar. Concerning low-rank approximations, they were experimented on geophysics applications [38] (Helmholtz equations) in the context of a collaboration with members of the ISTerre and Geoazur laboratories. The impact of both low-rank compression and shared-memory parallelism was also studied on electromagnetism problems [17], in collaboration with University of Padova (Italy) and CEDRAT.

We have started the design and implementation of a distributed-memory low-rank multifrontal solver. When computations are faster (thanks to low-rank compression or multithreading within each node), we observed that communications become critical; we are therefore currently studying the limits of the communication schemes from the MUMPS approach and their possible improvements.

On numerical and industrial aspects, we worked on rank detection and null space basis computations (in collaboration with CERFACS and Total/Hutchinson) as well as on improved parallel pivoting strategies for symmetric indefinite systems, in collaboration with ESI-Group (see Section 7.1).

6.21. Push-relabel based algorithms for the maximum transversal problem

In this work [14], we investigate the push-relabel algorithm for solving the problem of finding a maximum cardinality matching in a bipartite graph in the context of the maximum transversal problem. We describe in detail an optimized yet easy-to-implement version of the algorithm and fine-tune its parameters. We also introduce new performance-enhancing techniques. On a wide range of real-world instances, we compare the push-relabel algorithm with state-of-the-art algorithms based on augmenting paths and pseudoflows. We conclude that a carefully tuned push-relabel algorithm is competitive with all known augmenting path-based algorithms, and superior to the pseudoflow-based ones.

6.22. Constructing elimination trees for sparse unsymmetric matrices

The elimination tree model for sparse unsymmetric matrices and an algorithm for constructing it have been recently proposed [82], [83]. The construction algorithm has a worst-case time complexity of $\Theta(mn)$ for an $n \times n$ unsymmetric matrix having m off-diagonal nonzeros. In this work [15], we propose another algorithm that has a worst-case time complexity of $\mathcal{O}(m \log n)$. We compare the two algorithms experimentally and show that both algorithms are efficient in general. The algorithm of Eisenstat and Liu is faster in many practical cases, yet there are instances in which there is a significant difference between the running time of the two algorithms in favor of the proposed one.

6.23. Semi-matching algorithms for scheduling parallel tasks under resource constraints

In this work [25], we study the problem of minimum makespan scheduling when tasks are restricted to subsets of the processors (resource constraints), and require either one or multiple distinct processors to be executed (parallel tasks). This problem is related to the minimum makespan scheduling problem on unrelated machines, as well as to the concurrent job shop problem, and it amounts to finding a semi-matching in bipartite graphs or hypergraphs. The problem is known to be NP-complete for bipartite graphs with general vertex (task) weights, and solvable in polynomial time for unweighted graphs with unit weights (i.e., unit-weight tasks). We prove that the problem is NP-complete for hypergraphs even in the unweighted case. We design several greedy algorithms of low complexity to solve two versions of the problem, and assess their performance through a set of exhaustive simulations. Even though there is no approximation guarantee for these low-complexity algorithms, they return solutions close to the optimal (or a known lower bound) in average.

6.24. Maximum cardinality bipartite matching algorithms on GPUs

In two studies [30], [31], we propose, develop, and evaluate maximum cardinality matching algorithms from two different families (called push-relabel and augmenting-path based) on GPUs. The problem of finding a maximum cardinality matching in bipartite graphs has applications in computer science, scientific computing, bioinformatics, and other areas. To the best of our knowledge, the proposed algorithms are the first investigation of the push-relabel and augmenting-path based on GPUs/ We compare the proposed algorithms with serial and multicore implementations from the literature on a large set of real-life problems where in majority of the cases one of our GPU-accelerated algorithms is demonstrated to be faster than both the sequential and multicore implementations.

6.25. Analysis of partitioning models and metrics in parallel sparse matrix-vector multiplication

Graph/hypergraph partitioning models and methods have been successfully used to minimize the communication among processors in several parallel computing applications. Parallel sparse matrix-vector multiplication (SpMxV) is one of the representative applications that renders these models and methods indispensable in many scientific computing contexts. In this work [36], [55], we investigate the interplay of the partitioning metrics and execution times of SpMxV implementations in three libraries: Trilinos, PETSc, and an in-house one. We carry out experiments with up to 512 processors and investigate the results with regression analysis. Our experiments show that the partitioning metrics influence the performance greatly in a distributed memory setting. The regression analyses demonstrate which metric is the most influential for the execution time of the libraries.

6.26. On partitioning and reordering problems in a hierarchically parallel hybrid linear solver

PDSLin is a general-purpose algebraic parallel hybrid (direct/iterative) linear solver based on the Schur complement method. The most challenging step of the solver is the computation of a preconditioner based on the global Schur complement. Efficient parallel computation of the preconditioner gives rise to partitioning problems with sophisticated constraints and objectives. In this work [39], we identify two such problems and propose hypergraph partitioning methods to address them. The first problem is to balance the workloads associated with different subdomains to compute the preconditioner. We first formulate an objective function and a set of constraints to model the preconditioner computation time. Then, to address these complex constraints, we propose a recursive hypergraph bisection method. The second problem is to improve the data locality during the parallel solution of a sparse triangular system with multiple sparse right-hand sides. We carefully analyze the objective function and show that it can be well approximated by a standard hypergraph partitioning method. Moreover, an ordering compatible with a post ordering of the subdomain elimination tree is shown to be very effective in preserving locality. To evaluate the two proposed methods in practice, we present experimental results using linear systems arising from some applications of our interest. First, we show that in comparison to a commonly-used nested graph dissection method, the proposed recursive hypergraph partitioning method reduces the preconditioner construction time, especially when the number of subdomains is moderate. This is the desired result since PDSLin is based on a two-level parallelization to keep the number of subdomains small by assigning multiple processors to each subdomain. We also show that our second proposed hypergraph method improves the data locality during the sparse triangular solution and reduces the solution time. Moreover, we show that partitioning time can be greatly reduced while maintaining its quality by removing quasi-dense rows from the solution vectors.

6.27. UMPA: A Multi-objective, multi-level partitioner for communication minimization

In this work [42], we propose a directed hypergraph model and a refinement heuristic to distribute communicating tasks among the processing units in a distributed memory setting. The aim is to achieve load balance and minimize the maximum data sent by a processing unit. We also take two other communication metrics into account with a tie-breaking scheme. With this approach, task distributions causing an excessive use of network or a bottleneck processor which participates to almost all of the communication are avoided. We show on a large number of problem instances that our model improves the maximum data sent by a processor up to 34% for parallel environments with 4, 16, 64, and 256 processing units compared to the state of the art which only minimizes the total communication volume.

6.28. A Partitioning-based divisive clustering technique for maximizing the modularity

In this work [43], we present a new graph clustering algorithm aimed at obtaining clusterings of high modularity. The algorithm pursues a divisive clustering approach and uses established graph partitioning algorithms and techniques to compute recursive bipartitions of the input as well as to refine clusters. Experimental evaluation shows that the modularity scores obtained compare favorably to many previous approaches. In the majority of test cases, the algorithm outperformed the best known alternatives. In particular, among 13 problem instances common in the literature, the proposed algorithm improves the best known modularity in 9 cases.

6.29. Randomized matching heuristics with quality guarantees on shared memory parallel computers

In this work [56], we propose two heuristics for the bipartite matching problem that are amenable to shared-memory parallelization. The first heuristic is very intriguing from parallelization perspective. It has no significant algorithmic synchronization overhead and no conflict resolution is needed across threads. We show that this heuristic has an approximation ratio of around 0.632. The second heuristic is designed to obtain a larger matching by employing the well-known Karp-Sipser heuristic on a judiciously chosen subgraph of the original graph. We show that the Karp-Sipser heuristic always finds a maximum cardinality matching in the chosen subgraph. Although the Karp-Sipser heuristic is hard to parallelize for general graphs, we exploit the structure of the selected subgraphs to propose a specialized implementation which demonstrates a very good scalability. Based on our experiments and theoretical evidence, we conjecture that this second heuristic obtains matchings with cardinality of at least 0.866 of the maximum cardinality. We discuss parallel implementations of the proposed heuristics on shared memory systems. Experimental results, for demonstrating speed-ups and verifying the theoretical results in practice, are provided.

6.30. On the minimum edge cover and vertex partition by quasi-cliques problems

A γ-quasi-clique in a simple undirected graph is a set of vertices which induces a subgraph with the edge density of at least γ for $0 < \gamma < 1$. A cover of a graph by γ-quasi-cliques is a set of γ-quasi-cliques where each edge of the graph is contained in at least one quasi-clique. The minimum cover by γ-quasi-cliques problem asks for a γ-quasi-clique cover with the minimum number of quasi-cliques. A partition of a graph by γ-quasi-cliques is a set of γ-quasi-cliques where each vertex of the graph belongs to exactly one quasi-clique. The minimum partition by γ-quasi-cliques problem asks for a vertex partition by γ-quasi-cliques with the minimum number of quasi-cliques. In this work [60], we show that the decision versions of the minimum cover and partition by γ-quasi-cliques problems are NP-complete for any fixed γ satisfying $0 < \gamma < 1$.

RUNTIME Project-Team

6. New Results

6.1. SIMD Analysis Support in MAQAO

Either on ARM and x86 architectures, compilers and tools are needed for automatic and efficient vectorization. Although commercial compilers (e.g. IBM xlc, Intel icc, PGI pgcc) have made significant advances in auto-vectorization, a lot of source codes still remain too complicated for a compiler to vectorize, particularly when complex data structures are involved, or because of the lack of information at compile time. However, when vectorization fails, compilers leave the user with little clues about the cause of the failure, even though in certain cases moderate modifications could be applied on the source code to enable the compiler to vectorize.

Thus, the main objective for this work was to analyse SIMD vectorization potentials through loop detection. Parallelism detection is done through the instrumentation of the binary codes, capturing all memory streams in target loops and computing memory dependences using MAQAO. When combined to a static analysis for register dependences, this technique ensures that parallel slices of computation will be detected.

From a practical point of view, this work consists in the capture of the trace and its processing to extract memory reference patterns. To do so, we made use of the current state of the art MAQAO for instrumentation and trace capture on Intel architectures. We then implemented the dependence analysis on memory traces for performing loop pattern recognition. Finally, using this mechanism for loop pattern recognition, we can conclude about the vectorization potential of computation intensive loop nests. The dependence analysis does not depend on the target architecture, hence results computed for x86 architectures are valuable for ARM target as well.

6.2. NUMA-aware fine grain parallelization for multi-core architecture

Today, popular frameworks like Intel TBB or OpenMP offer a task based programming interface that allows to easily parallelize algorithms in shared memory. We have proposed some improvements to these task-based parallelization frameworks in order to cope with the problem of expressing an algorithm with a suitable task grain size and with the problem of Non Uniform Memory Accesses that degrades performance. In its current prototype state, our framework does not fully automate the selection of an optimal grain size. However, it significantly helps the programmer by proposing a simple interface to deal with DAG coarsening.

We have shown the benefits of this work on the parallelization of a sparse ILU preconditioner which is a challenging application with respect to task grain tuning and NUMA effect to an Intel TBB implementation. To improve even more the NUMA aspects, we are working on improving the task scheduler with cache-aware hierarchical scheduling support using a similar approach as the one implemented in the Bubblesched thread scheduler.

6.3. Task scheduling over heterogeneous architectures

We continued our work on extending STARPU to master exploitation of Heterogeneous Platforms through dynamic task scheduling, leading to the release of STARPU 1.1. We have extended our lightweight DSM to support out-of-core scheduling over disks. We have finished integrating STARPU with SIMGRID and obtained very accurate simulated times, which allows to experiment scheduling heuristics without having to actually execute the application on the target platform, thus tremendously reducing experimentation time and resource consumption.

We have modularized the scheduling part of STARPU, which permits to create complex schedulers by assembling simple scheduling components. This will allow theoreticians to work on writing the simple scheduling components without having to deal with the technical parts of the scheduling, performed in other scheduling components.

We have also collaborated with various research project to leverage the potential of STARPU: for instance, the PaStiX sparse matrix solver was ported over STARPU, so that we improved the dynamic task and management for applications with such fine-grain task size. This resulted with fair-enough performance on CPUs, compared to the hand-optimized static scheduler of PaStiX, and very promising performance on CPUs + GPUs. EADS ported its sparse hmatrix solver over STARPU, and we collaborated to work on adding STARPU support for communicating sparse data over MPI.

6.4. Task Size Control with XcalableMP/StarPU

On the work sharing among GPUs and CPU cores on GPU equipped clusters, it is a critical issue to select the task computational weights suited to these heterogeneous computing resources. We have been developing a solution for this problem, based on the cooperation of a PGAS language named XcalableMP (developed at the University of Tsukuba) together with a runtime sytem named XMP-dev/StarPU building on the work of the University of Tsukuba and on the StarPU platform developed by the Inria Runtime Team. Through the development, we found the necessity of adaptive task weight control for the GPU/CPU work sharing to achieve the best performance for various application codes. In particular, the language was extended to add a new feature allowing to alter the task size to be assigned to these heterogeneous resources dynamically during application execution. As a result of performance evaluation on several benchmarks, we confirmed the proposed feature correctly works and perform well even for relatively small size of problems.

6.5. Scheduling contexts for StarPU

Scheduling context is an extension of STARPU that allows multiple parallel codes to run concurrently with minimal interference. A scheduling context encapsulates an instance of the runtime system, and runs on top of a subset of the available processing units (i.e. regular cores or GPU accelerators). In order to maximize the overall efficiency of applications, contexts can be dynamically shrunk or expanded by a *hypervisor* that periodically gathers performance statistics inside each context (e.g. resource utilization, computation progress) and tries to determine how resources should be assigned to contexts so as to minimize the overall execution time. We have demonstrated the relevance of this approach using benchmarks invoking multiple high performance linear algebra kernels simultaneously on top of heterogeneous multicore machines. We have shown that our mechanism can dramatically improve the overall application run time (-34%), most notably by reducing the average cache miss ratio (-50%).

6.6. Load-balancing with TreeMatch

In the context of the Joint Laboratory for Petascale Computing (JLPC) included Inria and the University of Illinois at Urbana, we developed two load balancers for Charm++.

The two load-balancers we wrote take into account both the computing power and the hierarchical topology depending on the fact that the application is compute-bound or communication-bound. This work is based on our TREEMATCH library that computes process placement in order to reduce an application communication costs based on the hardware topology. Compared to some other solutions based on weighted topologies (latency, bandwidth, ...), ours is fully dynamic because we use only a qualitative approach for our representation of the hardware architecture.

The first load balancer is designed for compute-bound applications as it favours the leveling of CPU loads. The second load balancer focuses on communication-bound applications as it first reduces the congestion on the upper links in the topology tree.

These two load balancers gave us improvements for some applications up to 10% of the execution time.

6.7. List scheduling in embedded systems taking into account memroy constraints

Video decoding and image processing in embedded systems are subject to strong resource constraints, particularly in terms of memory. List-scheduling heuristics with static priorities (HEFT, SDC, etc.) being the often-cited solution due to both their good performance and their low complexity, we propose a method aimed at introducing the notion of memory into them. Moreover, we show that through appropriate adjustment of task priorities and judicious resort to insertion-based policy, speedups up to 20% can be achieved. Lastly, we show that our technique allows to prevent deadlock and to substantially reduce the required memory footprint compared to classic list-scheduling heuristics.

6.8. NewMadeleine generic multi-threading

The PIOMan progression engine utilized in NewMadeleine used to rely on the Marcel specific multi-threading library, with dedicated hooks and close co-operation between libraries. It restricted the target platforms and applications, and was considered as a constraint by users. We have designed mechanisms to make communication progress without hooks in the thread scheduler, able to run on any system with a `pthread` library. We have re-written PIOMan from the ground up to implement these mechanisms, and based on lock-free structures, with scalability in mind. A proof-of-concept port to the Intel Xeon Phi has been implemented in cooperation with the University of Tokyo, using the DCFA (Direct Communication Facility for manycore-based Accelerators) library to access InfiniBand boards from the Xeon Phi.

SAGE Project-Team

6. New Results

6.1. Parallel numerical algorithms

6.1.1. Parallel Adaptive GMRES with deflated restarting

Participant: Jocelyne Erhel.

Grants and projects: C2S@EXA 8.2.3 , JLPC 8.4.4

Software: DGMRES, AGMRES, GPREMS.

Publications: [17], [26].

Abstract: The GMRES iterative method is widely used as a Krylov subspace technique for solving sparse linear systems when the coefficient matrix is nonsymmetric and indefinite. The Newton basis implementation has been proposed on distributed memory computers as an alternative to the classical approach with the Arnoldi process. The aim of our work here is to introduce a modification based on deflation techniques. This approach builds an augmented subspace in an adaptive way to accelerate the convergence of the restarted formulation. In our numerical experiments, we show the benefits of using this implementation with hybrid direct/iterative methods to solve large linear systems.

6.1.2. Hybrid algebraic solvers for CFD problems

Participant: Jocelyne Erhel.

Grants and projects: C2S@EXA 8.2.3 , JLPC 8.4.4

Software: DGMRES, AGMRES, GPREMS.

Publications: [18].

Abstract: Sparse linear systems arise from design optimization in computational fluid dynamics. In this approach, a linearization of the discretized compressible Navier-Stokes equations is built, in order to evaluate the sensitivity of the entire flow with respect to each design parameter. The goal is to reduce the memory requirements and indirectly, the computational cost at different steps of this scheme. Numerical results are presented with industrial test cases to show the benefits of our methodology.

6.1.3. Algebraic multilevel preconditioning

Participant: Thomas Dufaud.

Grants: C2S@EXA 8.2.3

Publications: [51], [23], [24].

Conferences: [37], [24].

Abstract: The Schwarz domain decomposition method is a very attractive numerical method for parallel computing as it needs only to update the boundary conditions on the artificial interfaces generated by domain decomposition. Thus only local communications between the neighbouring sub-domains are required. We review the use of Aitken's acceleration applied to the Schwarz domain decomposition method.

6.1.4. Counting eigenvalues in domains of the complex field

Participant: Bernard Philippe.

Grants: momappli 8.4.2

Publications: [15], [28].

Conferences: [47], [48], [22].

Abstract: A procedure for counting the number of eigenvalues of a matrix in a region surrounded by a closed curve is presented. It is based on the application of the residual theorem. The quadrature is performed by evaluating the principal argument of the logarithm of a function. A strategy is proposed for selecting a path length that insures that the same branch of the logarithm is followed during the integration. Numerical tests are reported for matrices obtained from conventional matrix test sets.

6.1.5. *Sliced-time computation method*

Participant: Jocelyne Erhel.

Grants: MODNUM 8.4.5

Publications: [16], [25].

Abstract: We consider the mathematical framework of a sliced-time computation method for explosive solutions to systems of ordinary differential equations. We also derive an Adaptive Parallel-in-Time Method with application to a membrane problem.

6.1.6. *Interacting particles systems*

Participant: Lionel Lenôtre.

Grants: H2MNO4 8.2.1

Conferences: [31]

Abstract: We consider a variance reduction method for simulations with particles.

6.2. Numerical models and simulations applied to physics

6.2.1. *Heat and mass transfer modeling in porous media*

Participants: Édouard Canot, Salwa Mansour.

Grants: MODNUM 8.4.5 , HYDRINV 8.4.7

Conferences: [33], [35],[44]

Abstract: The effective thermal conductivity is a key parameter for obtaining good simulations of heat transfer in wet porous media. It is very sensitive to the presence of liquid water, even in very small quantity. Moreover, during the evaporation of water, some changes of geometric configuration of the liquid meniscus lead to hysteresis behaviors. Micro-scale studies help us in understanding the global properties, via numerical simulations.

6.2.2. *Heat transfer in soils applied to archaeological fires*

Participants: Édouard Canot, Salwa Mansour.

Grants: MODNUM 8.4.5 , ARPHYMAT 8.4.6

Conferences: [34], [36]

Abstract: In order to be validated, the numerical simulations of heat transfer at the surface of the soil are compared to experimental results, because of the complexity of the phenomenon and the great number of physical mechanisms involved. It appears that making good experiments is hard, not to mention the limitations and lacks of the Laloy and Massard method used to obtained the effective thermal conductivity of the granular material. The Laloy and Massard method have been slightly improved; besides a different, new experimental method, based on the mathematical properties of heat transfer, has been proposed.

6.2.3. *Granular materials*

Participant: Édouard Canot.

Publications: [19].

Abstract: Using the $\mu(I)$ continuum model recently proposed for dense granular flows, we study theoretically steady and fully developed granular flows in two configurations: a plane shear cell and a channel made of two parallel plates (Poiseuille configuration).

6.2.4. Geodesy

Participants: Amine Abdelmoula, Bernard Philippe.

Grants: LIRIMA-EPIC 8.4.3 , joint Ph-D 8.4.9 .

Publications: [12].

Thesis: Ph-D of Amine Abdelmoula, University of Rennes 1 and Tunis, defended in December 2013.

Abstract: We solve a geodetic inverse problem for the determination of a distribution of point masses (characterized by their intensities and positions), such that the potential generated by them best approximates a given potential field.

6.3. Models and simulations for flow and transport in porous media

6.3.1. Flow and transport in highly heterogeneous porous medium

Participants: Jean-Raynald de Dreuzy, Jocelyne Erhel, Géraldine Pichot.

Grants: H2MN04 8.2.1 , H2OGuilde 8.2.4 , HEMERA 8.2.2

Software: PARADIS, H2OLab

Publications: [13]

Abstract: Models of hydrogeology must deal with both heterogeneity and lack of data. We consider a flow and transport model for an inert solute. The conductivity is a random field following a stationary log normal distribution with an exponential or Gaussian covariance function, with a very small correlation length. The quantities of interest studied here are the expectation of the spatial mean velocity, the equivalent permeability and the macro spreading. In particular, the asymptotic behavior of the plume is characterized, leading to large simulation times, consequently to large physical domains. Uncertainty is dealt with a classical Monte Carlo method, which turns out to be very efficient, thanks to the ergodicity of the conductivity field and to the very large domain. These large scale simulations are achieved by means of high performance computing algorithms and tools.

6.3.2. Diffusion processes in porous media

Participants: Lionel Lenôtre, Géraldine Pichot.

Grants: H2MN04 8.2.1

Software: SBM 5.1.7 , PALMTREE 5.2.1

Publications: [21]

Conferences: [41], [43], [42]

Abstract: We present some recent results about Monte Carlo simulations in media with interfaces. By nature, porous media are extremely heterogeneous. We consider a one-dimensional advection-diffusion equation with piecewise constant coefficients. Without drift term, the Skew Brownian Motion permits to develop several exact algorithms with constant time step. We aim at adding the drift term and dealing with higher dimensional problems.

6.3.3. Adaptive stochastic model for flow and transport with random data

Participants: Jocelyne Erhel, Mestapha Oumouni.

Grants: HYDRINV 8.4.7 , joint Ph-D 8.4.8

Publications: [27].

Conferences: [46].

Thesis:[11].

Abstract: This work presents a development and an analysis of an effective approach for partial differential equation with random coefficients and data. We are interesting in the steady flow equation with stochastic input data.

A projection method in the one-dimensional case is presented to compute efficiently the average of the solution.

An anisotropic sparse grid collocation method is also used to solve the flow problem. First, we introduce an indicator of the error satisfying an upper bound of the error, it allows us to compute the anisotropy weights of the method. We demonstrate an improvement of the error estimation of the method which confirms the efficiency of the method compared with Monte Carlo and will be used to accelerate this method by the Richardson extrapolation technique.

We also present a numerical analysis of a probabilistic method to quantify the migration of a contaminant in random media. We consider the previous flow problem coupled with the advection-diffusion equation, where we are interested in the computation of the mean extension and the mean dispersion of the solute. The flow model is discretized by a mixed finite elements method and the concentration of the solute is the density of the solution of a stochastic differential equation, which is discretized by an Euler scheme. We present an explicit formula of the dispersion and optimal a priori error estimates.

6.3.4. Reactive transport

Participants: Édouard Canot, Jocelyne Erhel, Souhila Sabit.

Grants: H2MN04 8.2.1 , ANDRA 7.1 , MOMAS 8.2.7 , C2SEXA 8.2.3

Software: GRT3D.

Publications: [52],[30].

Conferences: [20],[40],[50], [32].

Abstract: Numerical simulations are essential for studying the fate of contaminants in aquifers, for risk assessment and resources management. In this study, we deal with reactive transport models and show how a Newton method can be used efficiently. Numerical experiments illustrate the efficiency of a substitution technique. Moreover, it appears that using logarithms in the chemistry equations lead to ill conditioned matrices and increase the computational cost.

6.4. Models and simulations for flow in porous fractured media

6.4.1. Synthetic benchmark for modeling flow in 3D fractured media

Participants: Jean-Raynald de Dreuzy, Jocelyne Erhel, Géraldine Pichot.

Grants: GEOFRAC 8.2.5 , FRACINI 8.1.1

Software: MPFRAC

Publications: [14]

Abstract: Intensity and localization of flows in fractured media have promoted the development of a large range of different modeling approaches including Discrete Fracture Networks, pipe networks and equivalent continuous media. While benchmarked usually within site studies, we propose an alternative numerical benchmark based on highly-resolved Discrete Fracture Networks (DFNs) and on a stochastic approach. Test cases are built on fractures of different lengths, orientations, aspect ratios and hydraulic apertures, issuing the broad ranges of topological structures and hydraulic properties classically observed. We present 18 DFN cases, with 10 random simulations by case.

6.4.2. Robust numerical methods for solving flow in stochastic fracture networks

Participants: Jean-Raynald de Dreuzy, Jocelyne Erhel, Géraldine Pichot.

Grants: GEOFRAC 8.2.5 , FRACINI 8.1.1

Software: MPFRAC, H2OLab.

Publications: [29].

Conferences: [49].

Abstract: In this work, flow in Discrete Fracture Networks (DFN) is solved using a Mortar Mixed Hybrid Finite Element Method. To solve large linear systems derived from a nonconforming discretization of stochastic fractured networks, a Balancing Domain Decomposition is used. Tests on three stochastically generated DFN are proposed to show the ability of the iterative solver SIDNUR to solve the flow problem.

6.4.3. Flow in complex 3D geological fractured porous media

Participants: Jean-Raynald de Dreuzy, Thomas Dufaud, Jocelyne Erhel, Géraldine Pichot.

Grants: GEOFRAC 8.2.5 , FRACINI 8.1.1

Software: MPFRAC, H2OLab

Conferences: [38], [39]

Abstract: Taking into account water and solute exchanges between porous and fractured media is of great interest in geological applications. The coupled porous-fractured flow equations and their discretization by a Mixed Hybrid Finite Element Method are presented as well as the derived linear system. An appropriate mesh generation is proposed to deal with the complexity involved by randomly generated fracture networks. Numerical experiments are shown, that provide flow fields for forthcoming transport simulations.

SCIPORT Team

6. New Results

6.1. Automatic Differentiation and parallel codes

Participants: Valérie Pascual, Laurent Hascoet, Jean Utke [Argonne National Lab. (Illinois, USA)], Michel Schanen [RWTH Aachen University (Germany)].

Together with colleagues in Argonne National Lab. and RWTH Aachen, we are studying how AD tools can handle MPI-parallel codes, especially in adjoint mode. Results are progressively incorporated into a library (AMPI, for Adjoinable-MPI) that is designed to provide efficient tangent and adjoint differentiation for MPI-parallel codes, independently of the AD tool used (AdolC, dco, OpenAD, TAPENADE ...). Primitives from the AMPI library dynamically orchestrate, at run-time, the MPI calls that are needed to compute the derivatives.

This year we studied issues raised by the collective reduction operations of MPI, and by the one-sided communications (i.e. remote memory access) offered by MPI-II.

The participants met on two occasions, two weeks in march in Sophia-Antipolis, and two weeks in october in Argonne.

This work was presented in particular at the meeting of the Inria-Illinois joint lab in june in Lyon. An article is in preparation.

6.2. Automatic Differentiation and Dynamic Memory

Participants: Valérie Pascual, Laurent Hascoet, Jean Utke [Argonne National Lab. (Illinois, USA)].

Adjoint differentiated code obtained by source transformation (OpenAD, TAPENADE...) consists of a forward sweep that essentially copies the original code, and a backward sweep that computes the derivatives, These two sweeps must have the same control flow shape, only reversed. The allocation and deallocation of some dynamic memory inside the forward sweep requires a similar pattern in the backward sweep. However, allocations do not always return the same memory chunk, and therefore all memory addresses must be updated to preserve their consistency in the backward sweep.

This problem can only be solved dynamically, at run-time. A compile-time analysis would have to be conservative, implying many overapproximations and in the end an unreasonably inefficient adjoint code. Our approach is thus to design a library that encapsulates all calls to memory allocation primitives (`malloc`, `free`...) in order to register the allocated addresses and to be able to restore consistency of pointers during the backward sweep. This strategy is similar to the one we use for MPI calls, *cf* 6.1 , and is actually needed in our AMPI strategy. All we can hope from a static analysis is to detect the simple cases where addresses could be recomputed instead of stored and updated. This may apply to a significant portion of memory manipulalions, and may thus reduce the overhead due to the dynamic updating.

We started developing this library, called ADMM for Adjoinable Dynamic Memory Management. TAPENADE will eventually produce adjoint code that calls these primitives instead of the standard memory management primitives.

6.3. Automatic Differentiation and iterative processes

Participants: Laurent Hascoet, Ala Taftaf.

Adjoint codes naturally propagate partial gradients backwards from the result of the simulation. However, this uses the data flow of the simulation in reverse order, at a cost that increases with the length of the simulation. AD research looks for strategies to reduce this cost, taking advantage of the structures of the given program. One such structure is iterative fixed point loops, commonplace in numerical computation. They occur at the topmost level of steady-state simulations, as well as in unsteady simulations. They may also occur deeper in the simulation, for instance in linear solvers.

It clear that the first iterations of a fixed-point search operate on a meaningless state vector, and that reversing the corresponding data-flow is wasted effort. An adapted adjoint strategy for the iterative process should consider only the last or the few last iterations. Furthermore, there is a discrete component to an iterative algorithm, namely the number of iterations, and this makes differentiability questionnable. For these reasons we are looking for a specific strategy for the adjoint, that reverses only the necessary data-flow, and that restores confidence in the validity of the derivative.

We seek inspiration in the strategies proposed by two authors [19], [27] to design one strategy that is amenable to implementation in a source-transformation AD tool such as TAPENADE. This will be triggered by user-given differentiation directives. We are also selecting example codes (a steady-state flow solver and a Newton solver) to benchmark and experiment.

Ala Taftaf presented her preliminary results at Queen Mary University in september, and at the 13th EuroAD workshop in Oxford, december 9-10. She attended a training on the CFD code OpenFOAM at Queen Mary, september 3-6.

6.4. Differentiation of third-party codes

Participants: Valérie Pascual, Laurent Hascoet, Alain Dervieux.

This year, we have differentiated two applications brought to us by academic colleagues. This is an important activity as it points us to problems that should be solved or interfaces that should be improved in TAPENADE.

- Striation simulates ionospheric plasma. It was developed in the University of Lille, then Toulouse. The Fortran90 source is relatively compact (10,000 lines). We obtained and validated the tangent and adjoint derivatives that were needed to solve an inverse problem i.e., identify the initial condition that causes an observed instability in plasma density. This work uncovered important AD issues when dynamic memory is used intensively, *cf* 6.2 .
- Mascaret is a hydrodynamic simulation code developed and used by EDF to study river flows. Mascaret consists of 120,000 lines of Fortran90. In this first experiment, we differentiated only one of the three kernel solvers. We obtained validated tangent and adjoint derivatives. A further collaboration with EDF and CERFACS is planned next year.

In addition, Automatic Differentiation of the CFD code AIRONUM (*cf* 5.1) will continue in cooperation with the partners of the FP7 project UMRIDA.

6.5. Resolution of linearised systems

Participants: Hubert Alcin [Inria Bordeaux-Sud-Ouest], Olivier Allain [Lemma], Marianna Braza [IMF-Toulouse], Alexandre Carabias, Alain Dervieux, Bruno Koobus [Université Montpellier 2], Carine Moussaed [Université Montpellier 2], Stephen Wornom [Lemma].

The work of Hubert Alcin for the ANR ECINADS on scalable parallel solvers based on coarse grids has been continued by Carine Moussaed and Bruno Koobus. This results in scalable computations up to 2048 processors.

Bruno Koobus and Carine Moussaed presented their results on “Un modèle VMS-LES dynamique pour la simulation d’écoulements autour d’obstacles” at CANUM congress, Super-Besse, France, May 21-25.

6.6. Control of approximation errors

Participants: Alexandre Carabias, Gautier Brethes, Alain Dervieux, Adrien Loseille [Gamma3 team, Inria-Rocquencourt], Frédéric Alauzet [Gamma3 team, Inria-Rocquencourt], Estelle Mbinky [Gamma3 team, Inria-Rocquencourt], Stephen Wornom [Lemma], Olivier Allain [Lemma], Anca Belme [university of Paris 6].

Third-order mesh adaptation was the main topic of the year in error control. Two PhD have been completed this year on third-order mesh adaptation:

- In team Gamma3, Estelle Mbinky has studied a method from Bernard Mourrain for transforming trilinear Taylor terms of the approximation error into a power of a bilinear term. Estelle Mbinky defended her thesis at Paris 6 on december 20.
- In our team, Alexandre Carabias (who spent most of this year with team Gamma3) has developed a 2D third-order scheme for the Euler model. The scheme is based on the ENO finite-volume formulation with quadratic reconstruction. Some effort was devoted to improve the performance of the scheme. The scheme is much less dissipative than an usual quadratic ENO scheme and of smaller cost. Implementation of a 3D version in AIRONUM (*cf* 5.1) is now starting. The 2D scheme has been the basis of an investigation of third order anisotropic mesh adaptation. Alexandre Carabias defended his thesis in Sophia-Antipolis on december 12.

A. Carabias and E. Mbinky presented their work on “A priori-based mesh adaptation for third-order accurate Euler simulation” at HONOM 2013, Bordeaux, France, March 18-22. We further studied mesh adaptation for viscous flows and we are preparing a journal article in collaboration with Gamma3 and University of Paris 6.

This year’s new topic is the combination of Multi-Grid and anisotropic mesh adaption, with the starting PhD of Gautier Brèthes. The study involves several problematics, and in particular stopping criteria and construction of correctors. This was supported by the ANR project ECINADS, ended in november, but continues with the ANR project MAIDESC (started in october, coordinated by our team) following on mesh adaption and in particular meshes for interfaces, third-order accuracy, meshes for boundary layers, and curved meshes.

SCORE Team

5. New Results

5.1. Evaluation and Design of Collaborative Editing Algorithms

Participants: Mehdi Ahmed-Nacer, Luc André, Claudia-Lavinia Ignat, Stéphane Martin, Gérald Oster, Pascal Urso.

Since the Web 2.0 era, the Internet is a huge content editing place in which users contribute to the content they browse. Users do not just edit the content but they collaborate on this content. Such shared content can be edited by thousands of people. However, current consistency maintenance algorithms seem not to be adapted to massive collaborative updating involving large amount of contributors and a high velocity of changes. This year we continued our work on the evaluation of existing collaborative editing approaches and on the design of new algorithms that overcome limitations of state of the art ones. Moreover, we started to work on experimental user studies for understanding the real-time requirements for collaborative editing and grounding a theory for the effect of real-time constraints in collaborative work [26].

We also run experiments to compare the merge automatically obtained by collaborative editing algorithms – CRDTs, OTs and the world-wide used diff3 – to the merge validated by the user. We obtain automatically such results exploiting the massively available distributed version control systems histories of open-source software. We use these results to improve an existing collaborative editing algorithm and obtain result statistically better than the existing ones (including diff3 used in major DVCS systems) [9].

In existing collaborative editing algorithms shared data is usually fragmented into fixed granularity atomic elements that can only be added or removed. Coarse-grained data leads to the possibility of conflicting updates while fine-grained data requires more metadata. In [11] we offer a solution for handling an adaptable granularity for shared data that overcomes the limitations of fixed-grained data approaches. Our solution relies on a novel commutative replicated data type (CRDT) for sequences of text that assigns unique identifiers to substrings of variable length contrary to existing CRDTs that assign unique identifiers to fixed size elements of the text (i.e. characters or lines). This offers the possibility to define coarse grained elements when they are created and refine them when needed. This greatly reduces the memory consumption since a smaller memory overhead is needed to store metadata (identifiers). Moreover, we show using simulations that overall performances of our algorithms are superior to existing ones.

We proposed a new concurrency control algorithm, based on conflict-free data types. It is built on the ideas previously developed for synchronous collaboration, extending them to support asynchronous collaboration. Our solution also includes the necessary information for providing comprehensive awareness information to users. The evaluation of our algorithm shows that comparing our solution with traditional solutions in collaborative editing, the conflict resolution strategy proposed in this paper leads to results closer to the ones expected by users [10].

5.2. Decentralized monitoring of orchestration execution

Participants: Mohamed Aymen Baouab, Olivier Perrin, Claude Godart.

Cross-organizational service-based processes are increasingly adopted by different companies when they cannot achieve goals on their own. The dynamic nature of these processes poses various challenges to their successful execution. In order to guarantee that all involved partners are informed about errors that may happen in the collaboration, it is necessary to monitor the execution process by continuously observing and checking message exchanges during runtime. This allows a global process tracking and evaluation of process metrics. Complex event processing can address this concern by analysing and evaluating message exchange events, to the aim of checking if the actual behaviour of the interacting entities effectively adheres to the modelled business constraints. In our recent work (Aymen Baouab thesis [1]), we presented an approach for decentralized monitoring of cross-organizational choreographies. We have defined a hierarchical propagation model for exchanging external notifications between the collaborating parties. We also proposed a runtime event-based approach to deal with the problem of monitoring conformance of interaction sequences. Our approach allows for an automatic and optimized generation of rules. After parsing the choreography graph into a hierarchy of canonical blocks, tagging each event by its block ascendancy, an optimized set of monitoring queries is generated. We evaluate the concepts based on a scenario showing how much the number of queries can be significantly reduced [12].

5.3. Optimization and security of business processes in SaaS contexts

Participants: Claude Godart, Elio Goettelmann, Samir Youcef.

Globalization and the increase of competitive pressures created the need for agility in business processes, including the ability to outsource, offshore, or otherwise distribute its once-centralized business processes or parts thereof. While hampered thus far by limited infrastructure capabilities, the increase in bandwidth and connectivity and decrease in communication cost have removed these limits. This is even more true with the advent of cloud, particularly in its "Service as a software" dimension. To adapt to such a context, there is a growing need for the ability to fragment one's business processes in an agile manner, and be able to distribute and wire these fragments so that their combined execution recreates the function of the original process. Our work focuses on solving some of the core challenges resulting from the need to dynamically restructure enterprise interactions. Restructuring such interactions corresponds to the fragmentation of intra and inter enterprise business process models. It describes how to identify, create, and execute process fragments without loosing the operational semantics of the original process models. In addition, this fragmentation is complicated by the constraints of quality of service, in particular the execution time and the cost, and of security, especially privacy. During the year, we consider this problem at two levels: the design of privacy-aware process models, and the process scheduling optimization. We developed a methodology to integrate privacy concerns in the design of a business process before distribution in the cloud. Based on a risk analysis, the result of the design is a set of process (re-)modelling actions, a set of constraints on process fragments assignments to clouds, and a set of constraints for cloud selection based on cloud properties [19]. We developed bi-criteria strategies for business processes scheduling in cloud environments with execution time and cost constraints, augmented with fairness metrics, and taking into account the availability of human resources, a critical point in business processes [14], [15], [3].

5.4. Large Scale Coordination of Crowdsourcing Activities

Participants: François Charoy, Karim Benouaret, Raman Valliyur-Ramalingam, Alexandre Roux d Anzi.

As a follow-up of our work on coordination of large scale processes that we have investigated in the domain of crisis management [4], [5], we have studied a new application domain for BPM, crowdsourcing. In order to make cities smarter, it would be interesting to design a platform where citizens are given an opportunity to be effectively connected to the governing bodies in their location and to contribute to the general well being. We have developed CrowdSC, a crowdsourcing framework designed for smarter cities. We have shown that it is possible to combine data collection, data selection and data assessment crowdsourcing activities in a crowdsourcing process to achieve sophisticated goals in a predefined context. Depending on the executing strategy of this process, different kinds of outcomes can be produced. We have conducted an experimental study that evaluates these process outcomes depending on different execution strategies [2], [13].

SECRET Project-Team

5. New Results

5.1. Symmetric cryptosystems

Participants: Anne Canteaut, Pascale Charpin, Virginie Lallemand, Gaëtan Leurent, María Naya-Plasencia, Joëlle Roué, Valentin Suder.

From outside, it might appear that symmetric techniques become obsolete after the invention of public-key cryptography in the mid 1970's. However, they are still widely used because they are the only ones that can achieve some major features like high-speed or low-cost encryption, fast authentication, and efficient hashing. Today, we find symmetric algorithms in GSM mobile phones, in credit cards, in WLAN connections. Symmetric cryptology is a very active research area which is stimulated by a pressing industrial demand for low-cost implementations (in terms of power consumption, gate complexity...). These extremely restricted implementation requirements are crucial when designing secure symmetric primitives and they might be at the origin of some weaknesses. Actually, these constraints seem quite incompatible with the rather complex mathematical tools needed for constructing a provably secure system.

The specificity of our research work is that it considers all aspects 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 new SHA-3 standard.

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 [11].
- Study of a new technique for attacking symmetric primitives based on the existence of linear relations between some input and output bits of the Sbox. This method has been used for improving the best known attack against the SHA-3 candidate Hamsi [36], [58].

5.1.2. Block ciphers

Even if the security of the current block cipher standard, AES, is not threatened when it is used in a classical context, there is still a need for the design of improved attacks, and for the determination of design criteria which guarantee that the existing attacks do not apply. This notably requires a deep understanding of all previously proposed attacks. Moreover, there is a high demand from the industry of lightweight block ciphers for some constrained environments. Several such algorithms have been proposed in the last few years and their security should be carefully analysed. Most of our work in this area is related to an ANR Project named BLOC.

Recent results:

- Cryptanalysis of several recently proposed lightweight block ciphers. This includes an attack against the full cipher KLEIN-64 [66], [49], and an attack against 8 rounds (out of 12) of PRINCE [37].
- Analysis of the resistance of AES-like permutations to improved rebound attacks. Most notably, this improved technique leads to a distinguisher on 10 rounds of the internal permutation of the SHA-3 candidate Grøstl [14].
- Proposal of a new family of distinguishers against AES-based permutations, named *limited-birthday distinguishers*; these distinguishers exploit some some improved rebound techniques. They have been successfully applied to various AES-based primitives including AES, ECHO, Grøstl, LED, PHOTON and Whirlpool [42].
- Design of an improved variant of Meet-in-the-Middle attacks, named *Sieve-in-the-Middle*: instead of selecting the key candidates by searching for a collision in an intermediate state which can be computed forwards and backwards, we here look for the existence of valid transitions through some middle Sbox. In the same paper, an improved technique is also proposed to build bicliques without needing any additional data (on the contrary to classical biclique attacks). These new methods have been exploited to break 8 rounds (out of 12) of the lightweight block cipher PRINCE [37], [59], [30].
- Analysis of the differential properties of the AES Superbox [48].
- Design of a new block cipher, named ZORRO, for which physical security is considered as an optimisation criterion [41].
- Design and study of a new construction for low-latency block ciphers, named *reflection ciphers*, which generalises the so-called α-reflection property exploited in PRINCE. This construction aims at reducing the implementation overhead of decryption on top of encryption [24].

5.1.3. Cryptographic properties and construction of appropriate building blocks

The construction of building blocks which guarantee a high resistance against the known attacks is a major topic within our project-team, for stream ciphers, block ciphers and hash functions. The use of such optimal objects actually leads to some mathematical structures which may be at the origin of new attacks. This work involves fundamental aspects related to discrete mathematics, cryptanalysis and implementation aspects. Actually, 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 [16], [51].
- Definition of a new criterion for Sboxes and link with some recent algebraic attacks on the hash function Hamsi [36], [58].
- Definition of some extended criterion for estimating the resistance of a block cipher to differential attacks. Most notably, this new criterion points out the fact that affinely equivalent Sboxes may not provide the same security level regarding differential cryptanalysis. This work emphasizes the role played by the affine permutation of the set of 8-bit words which follows the inverse function in the AES [21], [48].
- A new sufficient (and simpler) condition for checking that a mapping is APN has been established [62].
- Surveys of PN and APN mappings [55], [54].

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 factorisation problem (RSA, Rabin), or on the discrete logarithm problem (Diffie-Hellman, El Gamal, DSA). These systems have evolved and today instead of the classical groups ($\mathbf{Z}/n\mathbf{Z}$) we may use groups on elliptic curves. They allow a shorter block and key size for the same level of security. An intensive effort of the research community has been and is still being conducted to investigate the main aspects of these systems: implementation, theoretical and practical security. It must be noted that these systems all rely on algorithmic number theory. As they are used in most, if not all, applications of public-key cryptography today (and it will probably remain so in the near future), cryptographic applications are thus vulnerable to a single breakthrough in algorithmics or in hardware (a quantum computer can break all those schemes).

Diversity is a way to dilute that risk, and it is the duty of the cryptographic research community to prepare and propose alternatives to the number theoretic based systems. The most serious tracks today are lattice-based cryptography (NTRU,...), multivariate cryptography (HFE,...) and code-based cryptography (McEliece encryption scheme,...). All these alternatives are referred to as *post-quantum cryptosystems*, since they rely on difficult algorithmic problems which would not be solved by the coming-up of the quantum computer.

The code-based primitives have been investigated in details within the project-team. The first cryptosystem based on error-correcting codes was a public-key encryption scheme proposed by McEliece in 1978; a dual variant was proposed in 1986 by Niederreiter. We proposed the first (and only) digital signature scheme in 2001. Those systems enjoy very interesting features (fast encryption/decryption, short signature, good security reduction) but also have their drawbacks (large public key, encryption overhead, expensive signature generation). Some of the main issues in this field are

- security analysis , implementation and practicality of existing solutions,
- reducing the key size, *e.g.*, by using rank metric instead of Hamming metric, or by using particular families of codes,
- addressing new functionalities, like hashing or symmetric encryption.

Recent results:

- Design of a new variant of McEliece using Moderate Density Parity Check (MDPC) codes [45];
- Cryptanalysis of McEliece system based on Wild Goppa codes from a quadratic finite field extension. This polynomial-time structural attack relies on some filtration of nested subcodes which will reveal the secret algebraic description of the underlying secret code [39], [63].
- Cryptanalysis of a variant of the McEliece cryptosystem based on Reed-Solomon codes [38].

- Cryptanalysis of a variant of the McEliece cryptosystem based on convolutional codes proposed by Löndahl and Johansson in 2012 [43].
- Design of the first algorithm for distinguishing between Goppa codes (or alternant codes) over any field and random codes. Provided that the codes have sufficiently large rates, this technique can solve in polynomial-time the Goppa-Code-Distinguishing problem, which is an assumption in the security proof of McEliece cryptosystem [12].
- Study of the hardness of the code equivalence problem over $\mathbf{F}_q$. This problem has been extensively studied for permutation-equivalence (which covers all cases for $q = 2$). For $q \in \{3, 4\}$, we have generalised the support-splitting algorithm, and we have shown that the problem seems intractable for most instances when $q \geq 5$ [46]. This property has been exploited in an improvement version of an identification protocol due to Girault [47].

5.3. Reverse engineering of communication systems

Participants: Marion Bellard, Nicolas Sendrier, Jean-Pierre Tillich, Audrey Tixier.

To assess the quality of a cryptographic algorithm, it is usually assumed that its specifications are public, as, in accordance with Kerckhoffs principle [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, are observed out of a noisy channel. To access the information, the whole communication system has first to be disassembled and every constituent reconstructed. Our activity within this domain, whose first aim is to establish the scientific and technical foundations of a discipline which does not exist yet at an academic level, has been supported by some industrial contracts driven by the DGA.

Recent results:

- Reconstruction of the constellation labelling (i.e. used in the modulator of a communication system) in the presence of errors and when the underlying code is convolutional (Marion Bellard's PhD).

5.4. Quantum information theory

Participants: André Chailloux, Anthony Leverrier, Denise Maurice, Jean-Pierre Tillich.

The field of Quantum Information and Computation aims at exploiting the laws of quantum physics to manipulate information in radically novel ways. Two main applications come to mind: quantum computers, that offer the promise of solving some problems intractable with classical computers (for instance, factorization); and quantum cryptography, which provides new ways to exchange data in a provably secure fashion.

The main obstacle towards the development of quantum computing is decoherence, a consequence of the interaction of the computer with a noisy environment. We investigate approaches to quantum error-correction as a way to fight against this effect, and we study more particularly some families of quantum error-correcting codes which generalise the best classical codes available today.

Our research also covers quantum cryptography where we study the security of efficient protocols for key distribution, in collaboration with experimental groups. More generally, we investigate how quantum theory severely constrains the action of honest and malicious parties in cryptographic scenarios.

5.4.1. Quantum codes

Protecting quantum information from external noise is an issue of paramount importance for building a quantum computer. It also worthwhile to notice that all quantum error-correcting code schemes proposed up to now suffer from the very same problem that the first (classical) error-correcting codes had: there are constructions of good quantum codes, but for the best of them it is not known how to decode them in polynomial time. Our approach for overcoming this problem has been to study whether or not the family of turbo-codes and LDPC codes (and the associated iterative decoding algorithms) have a quantum counterpart.

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

Recent results:

- Construction of quantum codes combining an improved version of a family of spatially coupled quantum LDPC codes with a family of error reducing turbo-codes [44];
- construction of quantum LDPC codes with fixed non-zero rate and a minimum distance which grows proportionally to the square root of the block-length. This greatly improves the previously best known construction whose minimum distance was logarithmic in the block-length [19].
- Mamdouh Abbara's PhD thesis [9]

5.4.2. Quantum cryptography

A recent approach to cryptography takes into account that all interactions occur in a physical world described by the laws of quantum physics. These laws put severe constraints on what an adversary can achieve, and allow for instance to design provably secure key distribution protocols. We study such protocols as well as more general cryptographic primitives with security properties based on quantum theory.

Recent results:

- Experimental demonstration of quantum key distribution with continuous variables over 80 km [15], greatly improving over previous records around 25 km.
- Security proof of continuous-variable quantum key distribution protocols against general attacks [17], [29].
- Security proof of device-independent quantum key distribution in the bounded storage model [18].
- Study of BosonSampling, a recently introduced problem where quantum computers offer a provable speedup over classical computers [67], [28].
- Introduction and study of "Local Orthogonality", an information-theoretical principle for quantum correlations [13], [68].
- Introduction of a general formalism for the study of contextuality and non locality in quantum theory, based on the combinatorics of hypergraphs [65], [27].

SECSI Project-Team

6. New Results

6.1. Dishonest keys (Objective 2)

Participants: Hubert Comon-Lundh, Guillaume Scerri.

One of the main issues in the formal verification of the security protocols is the validity (and scope) of the formal model. Otherwise, it may happen that a protocol is proved and later someone finds an attack. This paradoxical situation may happen when the formal model used in the proof is too abstract.

A main stream of research therefore consists in proving full abstraction results (also called *soundness*): if the protocol is secure in the (symbolic) model, then an attack can only occur with negligible probability in a computational model. Such results have two main drawbacks: first they are very complicated, and have to be completed again and again for each combination of security primitives. Second, they require strong hypotheses on the primitives, some of which are not realistic. For instance, it is assumed that the attacker cannot forge his own keys (or that all keys come with their certificates, even for symmetric encryption keys).

Hubert Comon-Lundh, Véronique Cortier and Guillaume Scerri had proposed an extension of the symbolic model in 2012, and proved it computationally sound, without this restriction on the dishonest keys.

6.2. Deciding trace equivalence

Participants: David Baelde, Stéphanie Delaune, Rémy Chrétien, Lucca Hirschi.

Most existing results focus on trace properties like secrecy or authentication. There are however several security properties, which cannot be defined (or cannot be naturally defined) as trace properties and require the notion of indistinguishably. Typical examples are anonymity, privacy related properties or statements closer to security properties used in cryptography.

In the framework of the applied pi-calculus as in similar languages based on equational logics, indistinguishably corresponds to a relation called trace equivalence. Roughly, two processes are trace equivalent when an observer cannot see any difference between the two processes. Static equivalence applies only to observations on finite sets of messages, and do not take into account the dynamic behavior of a process whereas the notion of trace equivalence is more general and takes into account this aspect.

6.2.1. Static equivalence.

As explained above, static equivalence is a cornerstone to provide decision procedures for observational equivalence.

Stéphanie Delaune, in collaboration with Mathieu Baudet and Véronique Cortier, has designed a generic procedure for deducibility and static equivalence that takes as input any convergent rewrite system [12]. They have shown that their algorithm covers most of the existing decision procedures for convergent theories. They also provide an efficient implementation. This paper is a journal version of the work presented at RTA'09.

6.2.2. Trace equivalence.

When the processes under study do not contain replication, trace equivalence can be reduced to the problem of deciding symbolic equivalence [13]. Thanks to this reduction and relying on a result first proved by M. Baudet, this yields the first decidability result of observational equivalence for a general class of equational theories (for processes without else branches and without replication). Moreover, based on another decidability result for deciding equivalence between sets of constraint systems, we get decidability of trace equivalence for processes with else branch for standard primitives.

Even though there are some implementations of the procedures described above, this does not suffice to obtain practical tools. Current prototypes suffer from a classical combinatorial explosion problem caused by the exploration of many interleavings in the behaviour of processes. David Baelde, Stéphanie Delaune, and Lucca Hirschi revisit a work due to Mödersheim et al., generalize it and adapt it for equivalence checking. They obtain an optimization in the form of a reduced symbolic semantics that eliminates redundant interleavings on the fly. This work will be published as:

- D. Baelde, S. Delaune, and L. Hirschi. A Reduced Semantics for Deciding Trace Equivalence using Constraint Systems. In *Proc. 3rd Conference on Principles of Security and Trust (POST 2014)*, Grenoble, April 2014, France.

When processes under study contain replication, the approach relying on symbolic equivalence does not work anymore. Moreover, since it is well-known that deciding reachability properties is undecidable under various restrictions, there is actually no hope to do better for equivalence-based properties. Rémy Chrétien, Véronique Cortier, and Stéphanie Delaune provide the first results of (un)decidability for certain classes of protocols for the equivalence problem. They consider a class of protocols shown to be decidable for reachability properties, and establish a first undecidability result. Then, they restrained the class of protocols a step further by making the protocols deterministic in some sense and preventing it from disclosing secret keys. This tighter class of protocols was then shown to be decidable after reduction to an equivalence between deterministic pushdown automata. This work has been published at ICALP'13 [14].

To deal with replication, another approach has been studied by Vincent Cheval in collaboration with Bruno Blanchet. They propose an extension of the automatic protocol verifier ProVerif. 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, they extend the language of terms by defining more functions (destructors) by rewrite rules. These extensions have been implemented in ProVerif and allow one to automatically prove anonymity in the private authentication protocol by Abadi and Fournet. This work is part of Vincent Cheval's PhD thesis, and was published as:

- V. Cheval, B. Blanchet. Proving More Observational Equivalences with ProVerif. In *2nd Conference on Principles of Security and Trust (POST 2013)*. David Basin, John Mitchell, eds. Springer Verlag, Lecture Notes in Computer Science 7796, 2013.

6.3. Mobile ad-hoc networks

Participants: Rémy Chrétien, Stéphanie Delaune.

Mobile ad hoc networks consist of mobile wireless devices which autonomously organize their communication infrastructure: each node provides the function of a router and relays packets on paths to other nodes. Finding these paths in an a priori unknown and constantly changing network topology is a crucial functionality of any ad hoc network. Specific protocols, called *routing protocols*, are designed to ensure this functionality known as *route discovery*. Secured versions of routing protocols have been proposed to provide more guarantees on the resulting routes, and some of them have been designed to protect the privacy of the users.

Rémy Chrétien and Stéphanie Delaune propose a framework for analysing privacy-type properties for routing protocols. They use the notion of equivalence between traces to formalise three security properties related to privacy, namely indistinguishability, unlinkability, and anonymity. They study the relationship between these definitions and we illustrate them using two versions of the ANODR routing protocol. This work was published as:

- R. Chrétien, S. Delaune. Formal Analysis of Privacy for Routing Protocols in Mobile Ad Hoc Networks. *Principles of Security and Trust - Second International Conference, POST 2013*, held as Part of the *European Joint Conferences on Theory and Practice of Software, ETAPS 2013*, Rome, Italy, March 16-24, 2013. Proceedings. Springer 2013. Lecture Notes in Computer Science. ISBN 978-3-642-36829-5. Pages 1-20.

6.4. Composition results

Participant: Stéphanie Delaune.

Formal methods have proved their usefulness for analysing the security of protocols. However, protocols are often analysed in isolation, and this is well-known to be not sufficient as soon as the protocols share some keys.

Stéphanie Delaune, in collaboration with Céline Chevalier, Steve Kremer, and Mark Ryan, study whether password protocols can be safely composed, even when a same password is reused. More precisely, they present a transformation which maps a password protocol that is secure for a single protocol session (a decidable problem) to a protocol that is secure for an unbounded number of sessions. Their result provides an effective strategy to design secure password protocols: (i) design a protocol intended to be secure for one protocol session; (ii) apply their transformation and obtain a protocol which is secure for an unbounded number of sessions. Their technique also applies to compose different password protocols allowing one to obtain both inter-protocol and inter-session composition. This work was published as:

- C. Chevalier, S. Delaune, S. Kremer and M. Ryan. Composition of Password-based Protocols. *Formal Methods in System Design* 43(3), pages 369-413, 2013.

6.5. Unconditional Soundness (Objective 2)

Participants: Hubert Comon-Lundh, Guillaume Scerri.

Hubert Comon-Lundh, Véronique Cortier and Guillaume Scerri had shown in a 2012 CCS paper how one could drop one of the assumptions of computational soundness results. However, the proofs remain very complicated and there are still assumptions such as the absence of key cycles, or no dynamic corruption... that are still necessary for all these results.

Gergei Bana and Hubert Comon-Lundh investigated a completely different approach to formal security proofs in a 2012 POST paper, which does not make any such assumptions. The idea can be stated in a nutshell: whereas all existing formal models state the attacker's abilities, they propose to formally state what the attacker *cannot* do.

This makes a big difference, since the soundness need only to be proved formula by formula and only the very necessary assumptions are used for such formulas (for instance, no absence of key cycles is needed). This does not need to be proved again when a primitive is added.

Once the general setting is fixed, the question was how practical is the method. We studied the complexity of the consistency proofs in this setting and showed that we can complete such proofs in Polynomial Time for a wide class of axioms in

- H. Comon-Lundh, V. Cortier and G. Scerri. Tractable inference systems: an extension with a deducibility predicate. In CADE'13, LNAI 7898, pages 91-108. Springer, 2013

The development of a prototype implementation is under development. We expect to complete experiments on a number of protocols.

6.6. Static Analysis of Programs with Imprecise Probabilities

Participant: Jean Goubault-Larrecq [correspondant].

Static analyses allows one to obtain guarantees about the behavior of programs, without running them. Programs that handle numerical data such as feedback control loops pose a challenge in this area. This gets even harder when one considers programs that read numerical data from sensors, and write to actuators, as these data are imprecise, and are governed by probability distributions that may themselves be unknown, and only know to fall into some interval of distributions.

As part of the ANR projet blanc CPP, an efficient static analysis framework that deals with this kind of programs was proposed in 2011 by J. Goubault-Larrecq, O. Bouissou, E. Goubault, Sylvie Putot, based on P-boxes and Dempster-Shafer structures to handle imprecise probabilities.

The semantic foundations were made clearer, a new, improved algorithm was proposed, and new applications were examined in:

- A. Adjé, O. Bouissou, J. Goubault-Larrecq, E. Goubault and S. Putot. Static Analysis of Programs with Imprecise Probabilistic Inputs. In *VSTTE'13*, LNCS. Springer, 2013.

SELECT Project-Team

6. New Results

6.1. Model selection in Regression and Classification

Participants: Gilles Celeux, Serge Cohen, Jairo Cugliari, Tim Van Erwen, Clément Levrard, Erwan Le Pennec, Pascal Massart, Nelo Molter Magalhaes, Lucie Montuelle, Mohammed Sedki.

Erwan Le Pennec is still working with Serge Cohen (IPANEMA Soleil) on hyperspectral image segmentation based on a spatialized Gaussian Mixture Model. Their scheme is supported by some theoretical investigation and have been applied in pratice with an efficient minimization algorithm combining EM algorithm, dynamic programming and model selection implemented with MIXMOD. Lucie Montuelle is studying extensions of this model that comprise parametric logistic weights and regression mixtures.

Unsupervised segmentation is an issue similar to unsupervised classification with an added spatial aspect. Functional data is acquired on points in a spatial domain and the goal is to segment the domain in homogeneous domain. The range of applications includes hyperspectral images in conservation sciences, fMRi data and all spatialized functional data. Erwan Le Pennec and Lucie Montuelle are focusing on the questions of the way to handle the spatial component from both the theoretical and the practical point of views. They study in particular the choice of the number of clusters. Furthermore, as functional data require heavy computation, they are required to propose numerically efficient algorithms. They have also extend the model to regression mixture.

Lucie Montuelle focused on conditional density estimation by Gaussian mixtures with logistic weights. Using maximum likelihood estimators, a model selection procedure has been applied, supported by a theoretical guarantee. Numerical experiments have been conducted for regression mixtures with parametric logistic weights, using EM and Newton algorithms. This work is available in the research report and a submitted article.

In collaboration with Lucien Birgé (Université Paris 6), Pascal Massart and Nelo Molter Magalhaes define for the algorithm selection problem a new general cross validation procedure based on robust tests, which is an extension of the hold-out defined by Birgé. They get an original procedure based on the Hellinger distance. This procedure is the unique procedure which does not use any contrast function since it does not estimate the risk. They provide theoretical results showing that, under some weak assumptions on the considered statistical methods, the selected estimator satisfies an oracle type inequality. And, they prove that their robust method can be implemented with a sub-quadratic complexity. Simulations show that their estimator performs generally well for estimating a density with different sample sizes and can handle well-known problems, such as histogram or bandwidth selection.

In collaboration with Gérard Biau (Université Paris 6), Clément Levrard and Pascal Massart provide intuitive conditions have been derived for the k-means clustering algorithm to achieve its optimal rate of convergence. They can be thought of as margin conditions such as ones introduced by Mammen and Tsybakov in the statistical learning framework. These conditions can be checked in many cases, such as Gaussian mixtures with a known number of components and do not require the underlying distribution to have a density, on the contrary to the previous fast rates conditions introduced in this domain. Moreover, It allows to derive non-asymptotic bounds on the mean squared distortion of the k-mean estimator, emphasizing the role played by several other parameters of the quantization issue, such as the smallest distance between optimal codepoints or the excess risk of local minimizers. The influence of these parameters is still in discussion, but some previous results show that some of them are crucial for the minimax results obtained in quantization theory.

Tim van Erven is studying model selection for the long term. When a model selection procedure forms an integrated part of a company's day-to-day activities, its performance should be measured not on a single day, but on average over a longer period, like for example a year. Taking this long-term perspective, it is possible to aggregate model predictions optimally even when the data probability distribution is so irregular that no statistical guarantees can be given for any individual day separately. He studies the relation between model selection for individual days and for the long term, and how the geometry of the models affects both. This work has potential applications in model aggregation for the forecasting of electrical load consumption at EDF. Together with Jairo Cugliari it has also been applied to improve regional forecasts of electrical load consumption using the fact that the consumption of all regions together must add up to the total consumption over the whole country.

The well-documented and consistent variable selection procedure in model-based cluster analysis and classification, that Cathy Maugis (INSA Toulouse) has designed during her PhD. thesis in SELECT, makes use of stepwise algorithms which are painfully slow in high dimensions. In order to circumvent this drawback, Gilles Celeux and Mohammed Sedki, in collaboration with Cathy Maugis, proposed to sort the variables using a lasso-like penalization adapted to the Gaussian mixture model context. Using this rank to select the variables they avoid the combinatory problem of stepwise procedures. Their algorithm is now tested on several challenging simulated and real data sets, showing encouraging performances.

In collaboration with Jean-Michel Marin (Université de Montpellier) and Olivier Gascuel (LIRMM), Gilles Celeux has started a research aiming to select a short list of models rather a single model. This short list of models is declared to be compatible with the data using a p-value derived from the Kullback-Leibler distance between the model and the empirical distribution. And, the Kullback-Leibler distances at hand are estimated trough parametric bootstrap procedures.

6.2. Statistical learning methodology and theory

Participants: Vincent Brault, Gilles Celeux, Christine Keribin, Erwan Le Pennec, Lucie Montuelle, Mesrob Ohannessian, Michel Prenat, Solenne Thivin.

Gilles Celeux, Christine Keribin and the Ph D. student Vincent Brault continued their study on the Latent Block Model (LBM), and worked more especially on categorical data. They further investigated a Gibbs algorithm to avoid solutions with empty clusters on synthetic as well as real data (Congressional Voting Records and genomic data) [STCO13].They detailed the link between the information criteria ICL and BIC, compared them on synthetic and real data, and conjectured that these criteria are both consistent for LBM, which is not a standard behavior. ICL has been proved to be preferred for LBM.

V. Brault applied the Large Gaps algorithm and compared it with other existing algorithms [Aussois13]. He also derived a CEM algorithm for categorical LBM [Agroselect13]. In partnership with the Inria- MODAL team, he implemented the algorithms and information criteria in the R package blockcluster.

C. Keribin has started a collaboration with Tristan Mary-Huard (AgroParisTech) by the supervision of an internship (Master 2) on the use of LBM with truncated Poisson data.

Erwan Le Pennec is supervising Solenne Thivin in her CIFRE with Michel Prenat and Thales Optronique. The aim is target detection on complex background such as clouds or sea. Their approach is a local approach based on test decision theory. They have obtained theoretical and numerical results on a segmentation based approach in which a simple Markov field testing procedure is used in each cell of a data driven partition.

Erwan Le Pennec and Michel Prenat have also collaborated on a cloud texture modeling using a non-parametric approach. Such a modeling coud be used to better calibrate the detection procedure: it can lead to more examples than the one acquired and it could be the basis of an ensemble method.

Mesrob Ohannessian joined SELECT through an ERCIM Alain Bensoussan fellowship. During his stay, his work focused on two different aspects of statistics: large datasets and data scarcity. In collaboration with researchers in ETH Zurich (Prof. Andreas Krause), he studied the possibility of trading off statistical performance and computational speed in the context of k-means clustering, using the notion of coresets. In collaboration with researchers in Paris 11 (Prof. Elisabeth Gassiat) and Paris 7 (Prof. Stéphane Boucheron), he worked on adaptive universal compression when the alphabet is very large, meaning that some symbol observations are scarce.

6.3. Reliability

Participants: Yves Auffray, Gilles Celeux, Rémy Fouchereau, Patrick Pamphile.

Since 2011, in the framework of a CIFRE convention with Snecma-SAFRAN Rémy Fouchereau has started a thesis on the modeling of fatigue lifetime supervised by Gilles Celeux and Patrick Pamphile. In aircraft, space and nuclear industry, fatigue test is the main basic tool for analyzing fatigue lifetime of a given material, component, or structure. A sample of the material is subjected to cyclic loading S (stress, force, strain, etc.), by a testing machine which counts N, the number of cycles to failure. Fatigue test results are plotted on a SN-curve. A probabilistic model for the construction of SN-curve is proposed. In general, fatigue test results are widely scattered for High Cycle Fatigue region and "duplex" SN-curves appears for Very High Cycle region. That is why classic models from mechanic of rupture theory on one hand, probability theory on the other hand, do not fit SN-curve on the whole range of cycles. We have proposed a probabilistic model, based on a fracture mechanic approach: few parameters are required and they are easily interpreted by mechanic or material engineers.This model has been applied to both simulated and real fatigue test data sets. The SN-curves have been well fitted on the whole range of cycles. The parameters have been estimated using the EM algorithm, combining Newton-Raphson optimisation method and Monte Carlo integral estimations. Recently, the model has been improved taking into account production process information, thanks to a clustering approach. Thus, we have provided engineers with a probabilistic tool for reliability design of mechanical parts, but also with a diagnostic tool for material elaboration.

Since 2013, Gilles Celeux and Patrick Pamphile supervise, in the framework of a collaboration with CEA not yet finalized, a thesis on the modeling of battery State Of Charge for electrical vehicles. Electrical battery is an electrochemical device that converts stored chemical energy into electrical energy. This conversion is reversible and can be repeated during charge/discharge cycles. In an electric vehicle, the battery State Of Charge (SOC) gives the driver indication of how long he can drive without recharging the battery. Unfortunately the complex nature of electrochemical reactions does not allow to measure the SOC directly. Different methods of estimation exist, but they are not robust to various environment conditions (temperature, vehicle driving,...) and to the battery ageing. We propose to estimate the SOC from an *Markov-switching model* : the measurement equation specifies how the SOC depends of an unobservable Markov chain and physical data (temperature, voltage and current intensity,...). Moreover, the SOC estimation is included in the Battery Management System, and therefore estimations must be done online, i.e. with minimum information.

A collaboration has started in 2013 with Dassault Aviation on modal analysis of mechanical structures, which aims at identifying the vibration behavior of structures under dynamic excitations. From algorithmic view point, modal analysis amounts to estimation in parametric models on the basis of measured excitations and structural responses data. As it appears from literature and existing implementations, the model selection problem attached to this estimation is currently treated by a rather heavy and very heuristic procedure. The model selection via penalization tools are intended to be tested on this model selection problem.

6.4. Statistical analysis of genomic data

Participants: Vincent Brault, Gilles Celeux, Christine Keribin.

In collaboration with Florence Jaffrezic and Andrea Rau (INRA, animal genetic department), Mélina Gallopin has started a thesis under the supervision of Gilles Celeux. This thesis is concerned with building statistical networks of genes in animal grenetic. In animal genetic, datasets have a large number of genes and low number of statistical units. For this reason, standard network inference techniques work poorly in this case. At first, this team has developed a data-based method to filter replicated RNA-seq experiments. The method, implemented in the Bioconductor R package `HTSFilter`, removes low expressed genes by optimizing the Jaccard index and reduce the dimension of the dataset. Now, they are studying a clustering model on their expression profiles measured by RNAseq data using Poisson mixture models. External biological knowledge, such as Gene Ontology annotations are taken into account in the model selection step, based on a approximation of the completed log-likelihood given the annotations.

In collaboration with Marie-Laure Martin-Magniette (URGV), GIlles Celeux and Christine Keribin has started a research concerning the buliding statistical networks of transcription factors (TF) with Gaussian Graphical Models (GGM) in the frawork of the intership of Yann Vasseur (Université Paris-sud) who is starting a PhD. thesis on the same subject at the end of 2013. Since the number of TF is greater than the number of statistical units, a lasso-like procedure is used. Moreover the edges of the network are interpreted using the Latent Block Model studied by Vincent Brault in his thesis. An open issue to be solved is the choice of the regularization parameter in the lasso procedure. It is also important to develop this statistical inference for data with good biological control and knowledge to assess the biological relevance of the proposed models.

6.5. Curves classification, denoising and forecasting

Participants: Jairo Cugliari, Émilie Devijver, Pascal Massart, Jean-Michel Poggi, Vincent Thouvenot.

In collaboration with Farouk Mhamdi and Meriem Jaidane (ENIT, Tunis, Tunisia), Jean-Michel Poggi proposed a method for trend extraction from seasonal time series through the Empirical Mode Decomposition (EMD). Experimental comparison of trend extraction based on EMD, X11, X12 and Hodrick Prescott filter are conducted. First results show the eligibility of the blind EMD trend extraction method. Tunisian real peak load is also used to illustrate the extraction of the intrinsic trend.

Jean-Michel Poggi was the supervisor (with A. Antoniadis) of the PhD Thesis of Jairo Cugliari-Duhalde which takes place in a CIFRE convention with EDF. It was strongly related to the use of wavelets together with curves clustering in order to perform accurate load consumption forecasting. The thesis contains methodological and applied aspects linked to the electrical context as well as theoretical ones by introducing external variables in the context of nonparametric forecasting time series. See http://hal.archives-ouvertes.fr/docs/00/78/82/49/PDF/cugliari-jma.pdf and http://hal.inria.fr/docs/00/55/99/39/PDF/RR-7515.pdf The industrial post-doc of Jairo Cugliari, funded by EDF, explores three aspects of this model that complement the original methodology: first, the construction of a confidence interval for the predictor function, second, the flexibility and simplicity of the model to provide, without extra effort, forecasts horizons further and further away and finally, and third: study of the ability to provide good predictions in the presence of subtle signal nonstationarities induced by loss of customers coming from various scenarios, see http://hal.archives-ouvertes.fr/docs/00/81/49/24/PDF/kwf-suite.pdf

Jean-Michel Poggi, co-supervising with Anestis Antoniadis (Université Joseph Fourier Grenoble) the PhD thesis of Vincent Thouvenot, funded by a CIFRE with EDF. The industrial motivation of this work is the recent development of new technologies for measuring power consumption by EDF to acquire consumption data for different mesh network. The thesis will focus on the development of new statistical methods for predicting power consumption by exploiting the different levels of aggregation of network data collection. From the mathematical point of view, the work is to develop generalized additive models for this type of kind of aggregated data for the modeling of functional data, associating closely nonparametric estimation and variable selection using various penalization methods.

Jean-Michel Poggi and Pascal Massart are the co-advisors of the PhD thesis of Émilie Devijver, strongly motivated by the same kind of industrial forecasting problems in electricity, is dedicated to curves clustering for the prediction. A natural framework to explore this question is mixture of regression models for functional data. The theoretical subject of the thesis is to extend to functional data the recent work by Bühlmann et al. dealing with the simultaneous estimation of mixture regression models in the scalar case using Lasso type methods. Of course, it will be based on the technical tools of the work of Caroline Meynet (which completes her thesis Orsay under the direction of P. Massart), which deals with the clustering of functional data using Lasso methods choosing simultaneously number of clusters and selecting significant wavelet coefficients.

6.6. Neuroimaging, Statistical analysis of fMRI data

Participants: Gilles Celeux, Christine Keribin.

This research takes place as part of a collaboration with Neurospin on brain functional Magnetic Resonance Imaging (fMRI) data. (http://www.math.u-psud.fr/select/reunions/neurospin/Welcome.html). and concerns essentially regularisation in a supervised clustering methodology that includes spatial information in the prediction framework, and yields clustered weighted maps. C. Keribin examined the PhD defence of Virgile Fritsch High-dimensional statistical methods for inter-subjects studies in neuroimaging (Inria, Parietal team).

SÉMAGRAMME Project-Team

6. New Results

6.1. Syntax-Semantics Interface

6.1.1. TAG, Dependency Grammars, and ACG

Aleksandre Maskharashvili and Sylvain Pogodalla gave an ACG account of [41]'s process of transformation of the derivation trees of Tree Adjoining Grammar (TAG) into dependency trees. They made explicit how the requirement of keeping a direct interpretation of dependency trees into strings results into lexical ambiguity. Since the ACG framework has already been used to provide a logical semantics from TAG derivation trees, it results in a unified picture where derivation trees and dependency trees are related but independent equivalent ways to account for the same surface–meaning relation. This result has been published in [15].

6.1.2. Semantics of Neg-Raising Predicates in TAG

Laurence Danlos, Philippe de Groote, and Sylvain Pogodalla proposed a lexical semantic interpretation of Neg-Raising (NR) predicates that heavily relies on a Montague-like semantics for TAG and on higher-order types. NR verbs form a class of verbs with a clausal complement that show the following behavior: when a negation syntactically attaches to the ma- trix predicate, it can semantically attach to the embedded predicate, as the implication of (2) by (1) shows. This corresponds to the NR reading of this predicate.

- *Marie ne pense pas que Pierre partira.*
- *Marie pense que Pierre ne partira pas.*

As a base case, the approach lexically provides both NR and non-NR readings to NR predicates. The proposal is implemented in the ACG framework as it offers a fairly standard interface to logical formal semantics for TAG. This result has been published in [13].

6.1.3. Intensionalization

Makoto Kanazawa and Philippe de Groote have defined a general intensionalization procedure that turns an extensional semantics for a language into an intensionalized one that is capable of accommodating truly intensional lexical items without changing the compositional semantic rules [10]. They have proved some formal properties of this procedure and have clarified its relation to the procedure implicit in Montague's PTQ.

6.2. Lexical Disambiguation

Guy Perrier adapted the methods of lexical disambiguation presented in Mathieu Morey's PhD thesis [49] to the formalism of Tree Adjoining Grammar (TAG) in a common work with Claire Gardent, Yannick Parmentier and Sylvain Schmitz [24].

More precisely, the algorithm of lexical disambiguation for TAG uses the one-to-one relations between substitution nodes and roots of elementary tress in the parsing process and it takes also into account the position of the subsitution nodes with respect to the anchors in elementary trees, to discard lexical selections that do not respect some constraints. These constraints are implemented through a polarization of the elementary trees and for sake of efficiency, the lexical selections are represented in a compact way with automata.

A major default of the methods of lexical disambiguation presented in Mathieu Morey's PhD thesis is that they ignore local contexts. To overcome this default, Guy Perrier proposed an algorithm to foresee the elementary structures of the grammar that can be inserted between two words that will interact in the parsing process [20]. This algorithm applies to lexicalized grammars, in which the elementary structures are trees.

6.3. Linguistic Resources

6.3.1. Large Scale Grammatical Resources

Guy Perrier and Bruno Guillaume continued to develop FRIGRAM [2] a French grammar with a large coverage, written in the formalism of Interaction Grammars [16].

A major challenge in this task is to guarantee and to maintain the consistency of the grammar while aiming at the largest coverage. For this, they resorted an original property coming from the polarization of the elementary structures of an interaction grammar : the *companion property*. It is possible to determine all elementary structures (the *companions*) that are able to interact with a given elementary structure, in a static computation on the whole non anchored grammar, using the systeme of polarities. The knowledge of the companions of every elementary structure is very useful to check the linguistic consistency of a grammar.

Guy Perrier wrote a detailed documentation on FRIGRAM illustrated with a lot of examples [26].

6.3.2. Deep Syntax Annotation of the Sequoia French Treebank

Marie Candito, Guy Perrier, Bruno Guillaume, Corentin Ribeyre, Karën Fort, Djamé Seddah and Eric de la Clergerie started a project of annotating the Sequoia French Treebank with deep syntax dependencies.

The Sequoia French Treebank [33] is a 3 200 sentence treebank covering several domains (news, medical, europarl and fr-wikipedia). It is freely available and has already been annotated with surface dependency representations.

The participants in the project have defined a deep syntactic representation scheme for French, which abstracts away from surface syntactic variation and diathesis alternations. The goal is to obtain a freely available corpus, which will be useful for corpus linguistics studies and for training deep analyzers to prepare semantic analysis.

The different steps of the annotation process were conducted in a collaborative way. As the members of the project are located in two different French towns (Paris and Nancy), they decided to produce a complete annotation of the TreeBank in both towns and to collaboratively adjudicate the two results. In Nancy, Line Heckler, Mathilde Huguin and Alice Kneip produced a double annotation of the corpus and Guy Perrier was in charge of the adjudication.

At the beginning of the project, a mini reference was selected randomly, composed of 250 sentences from the Sequoia Corpus. Its annotation was conducted in parallel to the production of the annotation guide, in order to get feedback for the guide. Each team separately produced an initial annotated version of the mini reference. The final version, resulting from several iterations and adjudications, is already available [3].

The full version of the Sequoia French Treebank with deep syntax dependencies and its annotation guide will be released during Spring 2014.

6.3.3. Agile Annotation

In [19], Bruno Guillaume and Karën Fort present a methodology, inspired from the agile development paradigm, that helps preparing an annotation campaign. The idea behind the methodology is to formalize as much as possible the instructions given in the guidelines, in order to automatically check the consistency of the corpus being annotated with the guidelines, as they are being written. To formalize the guidelines, the authors use a graph rewriting tool, that allows to use a rich language to describe the instructions. This formalization allows to spot the rightfully annotated constructions and, by contrast, those that are not consistent with the guidelines. In case of inconsistency, an expert can either correct the annotation or update the guidelines and rerun the process.

[2] http://wikilligramme.loria.fr/doku.php?id=frigram:frigram
[3] http://talc2.loria.fr/mini_sequoia/

6.3.4. Integration of Multiple Constraints in ACG

In [14], Jiri Marsik and Maxime Amblard present a first step toward the integration of multiple constraints in ACG. However, all of the known treatments only consider tiny fragments of languages. We are interested in building a wide-coverage grammar which integrates and reconciles the existing formal treatments of discourse and allows us to study their interactions and to build discourse representations automatically.

This proposal is a first step towards a wide-coverage Abstract Categorial Grammar (ACG) that could be used to automatically build discourse-level representations. We focus on the challenge of integrating the treatment of disparate linguistic constraints in a single ACG and propose a generalization of the formalism: Graphical Abstract Categorial Grammars.

6.4. Graph Rewriting

Guillaume Bonfante and Bruno Guillaume studied formal properties of the Graph Rewriting in [12]. It is well-known that some linguistic phenomena do not cope properly with trees as the core mathematical structure to represent linguistic informations. In a former paper, the authors showed the benefit of encoding linguistic structures by graphs and of using graph rewriting rules to compute on those structures.

The Graph Rewriting formalism they consider is a formalization of the system which is implemented in the Grew software. Justified by some linguistic considerations, this Graph Rewriting formalization is characterized by two features: first, there is no node creation along computations and second, there are non-local edge modifications. Under these hypotheses, the article shows that uniform termination is undecidable and that non-uniform termination is decidable. Two termination techniques based on weights are described and a complexity bound on the derivation length for these rewriting systems is given.

6.5. Discourse in Pathological context

Maxime Amblard, Manuel Rebuschi and Michel Musiol continue to analyze in fine details pathological dialogues from the SLAM project. They present all theses results in [22] [21] and [11]. Schizophrenia is well-known among mental illnesses for the severity of the thought disorders it involves, and for their widespread and spectacular manifestations: from deviant social behavior to delusion, not to mention affective and sensitive distortions. The goal of our interdisciplinary work is to (i) analyze linguistic troubles in conversational contexts in which one of the speakers is schizophrenic, (ii) construe how the concept of rationality and logicality may apply to them, and (iii) propose a formal representation about this specific manifestation.

SequeL Project-Team

6. New Results

6.1. Decision-making Under Uncertainty

6.1.1. Reinforcement Learning

Minimax PAC bounds on the sample complexity of reinforcement learning with a generative model [2]

We consider the problem of learning the optimal action-value function in discounted-reward Markov decision processes (MDPs). We prove new PAC bounds on the sample-complexity of two well-known model-based reinforcement learning (RL) algorithms in the presence of a generative model of the MDP: value iteration and policy iteration. The first result indicates that for an MDP with N state-action pairs and the discount factor γin[0, 1) only O(N log(N/δ)/ [(1 - γ)3 ϵ2]) state-transition samples are required to find an ϵ-optimal estimation of the action-value function with the probability (w.p.) 1-δ. Further, we prove that, for small values of ϵ, an order of O(N log(N/δ)/ [(1 - γ)3 ϵ2]) samples is required to find an ϵ-optimal policy w.p. 1-δ. We also prove a matching lower bound of Ω(N log(N/δ)/ [(1 - γ)3 ϵ2]) on the sample complexity of estimating the optimal action-value function. To the best of our knowledge, this is the first minimax result on the sample complexity of RL: The upper bound matches the lower bound interms of N , ϵ, δ and 1/(1 -γ) up to a constant factor. Also, both our lower bound and upper bound improve on the state-of-the-art in terms of their dependence on 1/(1-γ).

Regret Bounds for Reinforcement Learning with Policy Advice [13]

In some reinforcement learning problems an agent may be provided with a set of input policies, perhaps learned from prior experience or provided by advisors. We present a reinforcement learning with policy advice (RLPA) algorithm which leverages this input set and learns to use the best policy in the set for the reinforcement learning task at hand. We prove that RLPA has a sub-linear regret of $\tilde{O}(\sqrt{T})$ relative to the best input policy, and that both this regret and its computational complexity are independent of the size of the state and action space. Our empirical simulations support our theoretical analysis. This suggests RLPA may offer significant advantages in large domains where some prior good policies are provided.

Optimistic planning for belief-augmented Markov decision processes [11]

This paper presents the Bayesian Optimistic Planning (BOP) algorithm, a novel model-based Bayesian reinforcement learning approach. BOP extends the planning approach of the Optimistic Planning for Markov Decision Processes (OP-MDP) algorithm [10], [9] to contexts where the transition model of the MDP is initially unknown and progressively learned through interactions within the environment. The knowledge about the unknown MDP is represented with a probability distribution over all possible transition models using Dirichlet distributions, and the BOP algorithm plans in the belief-augmented state space constructed by concatenating the original state vector with the current posterior distribution over transition models. We show that BOP becomes Bayesian optimal when the budget parameter increases to infinity. Preliminary empirical validations show promising performance.

Aggregating optimistic planning trees for solving markov decision processes [16]

This paper addresses the problem of online planning in Markov decision processes using a generative model and under a budget constraint. We propose a new algorithm, ASOP, which is based on the construction of a forest of single successor state planning trees, where each tree corresponds to a random realization of the stochastic environment. The trees are explored using a "safe" optimistic planning strategy which combines the optimistic principle (in order to explore the most promising part of the search space first) and a safety principle (which guarantees a certain amount of uniform exploration). In the decision-making step of the algorithm, the individual trees are aggregated and an immediate action is recommended. We provide a finite-sample analysis and discuss the trade-off between the principles of optimism and safety. We report numerical results on a benchmark problem showing that ASOP performs as well as state-of-the-art optimistic planning algorithms.

Optimal Regret Bounds for Selecting the State Representation in Reinforcement Learning [20]

We consider an agent interacting with an environment in a single stream of actions, observations, and rewards, with no reset. This process is not assumed to be a Markov Decision Process (MDP). Rather, the agent has several representations (mapping histories of past interactions to a discrete state space) of the environment with unknown dynamics, only some of which result in an MDP. The goal is to minimize the average regret criterion against an agent who knows an MDP representation giving the highest optimal reward, and acts optimally in it. Recent regret bounds for this setting are of order $O(T^{2/3})$ with an additive term constant yet exponential in some characteristics of the optimal MDP. We propose an algorithm whose regret after T time steps is $O(\sqrt{T})$, with all constants reasonably small. This is optimal in T since $O(\sqrt{T})$ is the optimal regret in the setting of learning in a (single discrete) MDP.

Competing with an Infinite Set of Models in Reinforcement Learning [21]

We consider a reinforcement learning setting where the learner also has to deal with the problem of finding a suitable state-representation function from a given set of models. This has to be done while interacting with the environment in an online fashion (no resets), and the goal is to have small regret with respect to any Markov model in the set. For this setting, recently the BLBãlgorithm has been proposed, which achieves regret of order $T^{2/3}$, provided that the given set of models is finite. Our first contribution is to extend this result to a countably infinite set of models. Moreover, the BLBr˜egret bound suffers from an additive term that can be exponential in the diameter of the MDP involved, since the diameter has to be guessed. The algorithm we propose avoids guessing the diameter, thus improving the regret bound.

A review of optimistic planning in Markov decision processes [30]

We review a class of online planning algorithms for deterministic and stochastic optimal control problems, modeled as Markov decision processes. At each discrete time step, these algorithms maximize the predicted value of planning policies from the current state, and apply the first action of the best policy found. An overall receding-horizon algorithm results, which can also be seen as a type of model-predictive control. The space of planning policies is explored optimistically, focusing on areas with largest upper bounds on the value - or upper confidence bounds, in the stochastic case. The resulting optimistic planning framework integrates several types of optimism previously used in planning, optimization, and reinforcement learning, in order to obtain several intuitive algorithms with good performance guarantees. We describe in detail three recent such algorithms, outline the theoretical guarantees on their performance, and illustrate their behavior in a numerical example.

6.1.2. Multi-arm Bandit Theory

Automatic motor task selection via a bandit algorithm for a brain-controlled button [4]

Objective. Brain-computer interfaces (BCIs) based on sensorimotor rhythms use a variety of motor tasks, such as imagining moving the right or left hand, the feet or the tongue. Finding the tasks that yield best performance, specifically to each user, is a time-consuming preliminary phase to a BCI experiment. This study presents a new adaptive procedure to automatically select (online) the most promising motor task for an asynchronous brain-controlled button. Approach. We develop for this purpose an adaptive algorithm UCB-classif based on the stochastic bandit theory and design an EEG experiment to test our method. We compare (offline) the adaptive algorithm to a naïve selection strategy which uses uniformly distributed samples from each task. We also run the adaptive algorithm online to fully validate the approach. Main results. By not wasting time on inefficient tasks, and focusing on the most promising ones, this algorithm results in a faster task selection and a more efficient use of the BCI training session. More precisely, the offline analysis reveals that the use of this algorithm can reduce the time needed to select the most appropriate task by almost half without loss in precision, or alternatively, allow us to investigate twice the number of tasks within a similar time span. Online tests confirm that the method leads to an optimal task selection. Significance. This study is the first one to optimize the task selection phase by an adaptive procedure. By increasing the number of tasks that can be tested in a given time span, the proposed method could contribute to reducing 'BCI illiteracy'.

Kullback-Leibler Upper Confidence Bounds for Optimal Sequential Allocation [3]

We consider optimal sequential allocation in the context of the so-called stochastic multi-armed bandit model. We describe a generic index policy, in the sense of Gittins (1979), based on upper confidence bounds of the arm payoffs computed using the Kullback-Leibler divergence. We consider two classes of distributions for which instances of this general idea are analyzed: The kl-UCB algorithm is designed for one-parameter exponential families and the empirical KL-UCB algorithm for bounded and finitely supported distributions. Our main contribution is a unified finite-time analysis of the regret of these algorithms that asymptotically matches the lower bounds of Lai and Robbins (1985) and Burnetas and Katehakis (1996), respectively. We also investigate the behavior of these algorithms when used with general bounded rewards, showing in particular that they provide significant improvements over the state-of-the-art.

Sequential Transfer in Multi-armed Bandit with Finite Set of Models [14]

Learning from prior tasks and transferring that experience to improve future performance is critical for building lifelong learning agents. Although results in supervised and reinforcement learning show that transfer may significantly improve the learning performance, most of the literature on transfer is focused on batch learning tasks. In this paper we study the problem of *sequential transfer in online learning*, notably in the multi–armed bandit framework, where the objective is to minimize the total regret over a sequence of tasks by transferring knowledge from prior tasks. Under the assumption that the tasks are drawn from a stationary distribution over a finite set of models, we define a novel bandit algorithm based on a method-of-moments approach for the estimation of the possible tasks and derive regret bounds for it. We introduce a novel bandit algorithm based on a method-of-moments approach for estimating the possible tasks and derive regret bounds for it. Finally, we report preliminary empirical results confirming the theoretical findings.

Optimizing P300-speller sequences by RIP-ping groups apart [25]

So far P300-speller design has put very little emphasis on the design of optimized flash patterns, a surprising fact given the importance of the sequence of flashes on the selection outcome. Previous work in this domain has consisted in studying consecutive flashes, to prevent the same letter or its neighbors from flashing consecutively. To this effect, the flashing letters form more random groups than the original row-column sequences for the P300 paradigm, but the groups remain fixed across repetitions. This has several important consequences, among which a lack of discrepancy between the scores of the different letters. The new approach proposed in this paper accumulates evidence for individual elements, and optimizes the sequences by relaxing the constraint that letters should belong to fixed groups across repetitions. The method is inspired by the theory of Restricted Isometry Property matrices in Compressed Sensing, and it can be applied to any display grid size, and for any target flash frequency. This leads to P300 sequences which are shown here to perform significantly better than the state of the art, in simulations and online tests.

Stochastic Simultaneous Optimistic Optimization [26]

We study the problem of global maximization of a function f given a finite number of evaluations perturbed by noise. We consider a very weak assumption on the function, namely that it is locally smooth (in some precise sense) with respect to some semi-metric, around one of its global maxima. Compared to previous works on bandits in general spaces (Kleinberg et al., 2008; Bubeck et al., 2011a) our algorithm does not require the knowledge of this semi-metric. Our algorithm, StoSOO, follows an optimistic strategy to iteratively construct upper confidence bounds over the hierarchical partitions of the function domain to decide which point to sample next. A finite-time analysis of StoSOO shows that it performs almost as well as the best specifically-tuned algorithms even though the local smoothness of the function is not known.

Toward optimal stratification for stratified monte-carlo integration [9]

We consider the problem of adaptive stratified sampling for Monte Carlo integration of a noisy function, given a finite budget n of noisy evaluations to the function. We tackle in this paper the problem of adapting to the function at the same time the number of samples into each stratum and the partition itself. More precisely, it is interesting to refine the partition of the domain in area where the noise to the function, or where the variations of the function, are very heterogeneous. On the other hand, having a (too) refined stratification is not optimal. Indeed, the more refined the stratification, the more difficult it is to adjust the allocation of the samples to the stratification, i.e. sample more points where the noise or variations of the function are larger. We provide in this paper an algorithm that selects online, among a large class of partitions, the partition that provides the optimal trade-off, and allocates the samples almost optimally on this partition

Thompson sampling for one-dimensional exponential family bandits [18]

Thompson Sampling has been demonstrated in many complex bandit models, however the theoretical guarantees available for the parametric multi-armed bandit are still limited to the Bernoulli case. Here we extend them by proving asymptotic optimality of the algorithm using the Jeffreys prior for 1-dimensional exponential family bandits. Our proof builds on previous work, but also makes extensive use of closed forms for Kullback-Leibler divergence and Fisher information (and thus Jeffreys prior) available in an exponential family. This allow us to give a finite time exponential concentration inequality for posterior distributions on exponential families that may be of interest in its own right. Moreover our analysis covers some distributions for which no optimistic algorithm has yet been proposed, including heavy-tailed exponential families.

Finite-Time Analysis of Kernelised Contextual Bandits [27]

We tackle the problem of online reward maximisation over a large finite set of actions described by their contexts. We focus on the case when the number of actions is too big to sample all of them even once. However we assume that we have access to the similarities between actions' contexts and that the expected reward is an arbitrary linear function of the contexts' images in the related reproducing kernel Hilbert space (RKHS). We propose KernelUCB, a kernelised UCB algorithm, and give a cumulative regret bound through a frequentist analysis. For contextual bandits, the related algorithm GP-UCB turns out to be a special case of our algorithm, and our finite-time analysis improves the regret bound of GP-UCB for the agnostic case, both in the terms of the kernel-dependent quantity and the RKHS norm of the reward function. Moreover, for the linear kernel, our regret bound matches the lower bound for contextual linear bandits.

From Bandits to Monte-Carlo Tree Search: The Optimistic Principle Applied to Optimization and Planning [33]

This work covers several aspects of the optimism in the face of uncertainty principle applied to large scale optimization problems under finite numerical budget. The initial motivation for the research reported here originated from the empirical success of the so-called Monte-Carlo Tree Search method popularized in computer-go and further extended to many other games as well as optimization and planning problems. Our objective is to contribute to the development of theoretical foundations of the field by characterizing the complexity of the underlying optimization problems and designing efficient algorithms with performance guarantees. The main idea presented here is that it is possible to decompose a complex decision making problem (such as an optimization problem in a large search space) into a sequence of elementary decisions, where each decision of the sequence is solved using a (stochastic) multi-armed bandit (simple mathematical model for decision making in stochastic environments). This so-called hierarchical bandit approach (where the reward observed by a bandit in the hierarchy is itself the return of another bandit at a deeper level) possesses the nice feature of starting the exploration by a quasi-uniform sampling of the space and then focusing progressively on the most promising area, at different scales, according to the evaluations observed so far, and eventually performing a local search around the global optima of the function. The performance of the method is assessed in terms of the optimality of the returned solution as a function of the number of function evaluations. Our main contribution to the field of function optimization is a class of hierarchical optimistic algorithms designed for general search spaces (such as metric spaces, trees, graphs, Euclidean spaces, ...) with different algorithmic instantiations depending on whether the evaluations are noisy or noiseless and whether some measure of the ”smoothness” of the function is known or unknown. The performance of the algorithms depend on the local behavior of the function around its global optima expressed in terms of the quantity of near-optimal states measured with some metric. If this local smoothness of the function is known then one can design very efficient optimization algorithms (with convergence rate independent of the space dimension), and when it is not known, we can build adaptive techniques that can, in some cases, perform almost as well as when it is known.

6.2. Statistical analysis of time series

6.2.1. Change Point Analysis

Nonparametric multiple change point estimation in highly dependent time series [17]

Given a heterogeneous time-series sample, it is required to find the points in time (called change points) where the probability distribution generating the data has changed. The data is assumed to have been generated by arbitrary, unknown, stationary ergodic distributions. No modeling, independence or mixing are made. A novel, computationally efficient, nonparametric method is proposed, and is shown to be asymptotically consistent in this general framework; the theoretical results are complemented with experimental evaluations.

6.2.2. Clustering Time Series, Online and Offline

A Binary-Classification-Based Metric between Time-Series Distributions and Its Use in Statistical and Learning Problems [6]

A metric between time-series distributions is proposed that can be evaluated using binary classification methods, which were originally developed to work on i.i.d. data. It is shown how this metric can be used for solving statistical problems that are seemingly unrelated to classification and concern highly dependent time series. Specifically, the problems of time-series clustering, homogeneity testing and the three-sample problem are addressed. Universal consistency of the resulting algorithms is proven under most general assumptions. The theoretical results are illustrated with experiments on synthetic and real-world data.

6.2.3. Semi-Supervised and Unsupervised Learning

Learning from a Single Labeled Face and a Stream of Unlabeled Data [19]

Face recognition from a single image per person is a challenging problem because the training sample is extremely small. We consider a variation of this problem. In our problem, we recognize only one person, and there are no labeled data for any other person. This setting naturally arises in authentication on personal computers and mobile devices, and poses additional challenges because it lacks negative examples. We formalize our problem as one-class classification, and propose and analyze an algorithm that learns a non-parametric model of the face from a single labeled image and a stream of unlabeled data. In many domains, for instance when a person interacts with a computer with a camera, unlabeled data are abundant and easy to utilize. This is the first paper that investigates how these data can help in learning better models in the single-image-per-person setting. Our method is evaluated on a dataset of 43 people and we show that these people can be recognized 90% of time at nearly zero false positives. This recall is 25+% higher than the recall of our best performing baseline. Finally, we conduct a comprehensive sensitivity analysis of our algorithm and provide a guideline for setting its parameters in practice.

Unsupervised model-free representation learning [23]

Numerous control and learning problems face the situation where sequences of high-dimensional highly dependent data are available, but no or little feedback is provided to the learner. In such situations it may be useful to find a concise representation of the input signal, that would preserve as much as possible of the relevant information. In this work we are interested in the problems where the relevant information is in the time-series dependence. Thus, the problem can be formalized as follows. Given a series of observations $X\_0, \cdots, X\_n$ coming from a large (high-dimensional) space $\mathcal{X}$, find a representation function f mapping $\mathcal{X}$ to a finite space $\mathcal{Y}$ such that the series $f(X\_0), \cdots, f(X\_n)$ preserve as much information as possible about the original time-series dependence in $X\_0, \cdots, X\_n$. For stationary time series, the function f can be selected as the one maximizing the time-series information $I\_\infty(f) = h\_0(f(X)) - h\_\infty(f(X))$ where $h\_0(f(X))$ is the Shannon entropy of $f(X\_0)$ and $h\_\infty(f(X))$ is the entropy rate of the time series

$f(X\_0), \cdots, f(X\_n), \cdots$. In this paper we study the functional $I\_\infty(f)$ from the learning-theoretic point of view. Specifically, we provide some uniform approximation results, and study the behaviour of $I\_\infty(f)$ in the problem of optimal control.

Time-series information and learning [22]

Given a time series $X\_1, \cdots, X\_n, \cdots$ taking values in a large (high-dimensional) space $\mathcal{X}$, we would like to find a function f from $\mathcal{X}$ to a small (low-dimensional or finite) space $\mathcal{Y}$ such that the time series $f(X\_1), \cdots, f(X\_n), \cdots$ retains all the information about the time-series dependence in the original sequence, or as much as possible thereof. This goal is formalized in this work, and it is shown that the target function f can be found as the one that maximizes a certain quantity that can be expressed in terms of entropies of the series $(f(X\_i))\_i \in \mathbb{N}$. This quantity can be estimated empirically, and does not involve estimating the distribution on the original time series $(X\_i)\_i \in \mathbb{N}$.

6.3. Statistical Learning and Bayesian Analysis

6.3.1. Dictionary learning

Learning a common dictionary over a sensor network [10]

We consider the problem of distributed dictionary learning, where a set of nodes is required to collectively learn a common dictionary from noisy measurements. This approach may be useful in several contexts including sensor networks. Diffusion cooperation schemes have been proposed to solve the distributed linear regression problem. In this work we focus on a diffusion-based adaptive dictionary learning strategy: each node records independent observations and cooperates with its neighbors by sharing its local dictionary. The resulting algorithm corresponds to a distributed alternate optimization. Beyond dictionary learning, this strategy could be adapted to many matrix factorization problems in various settings. We illustrate its efficiency on some numerical experiments.

Distributed dictionary learning over a sensor network [29]

We consider the problem of distributed dictionary learning, where a set of nodes is required to collec- tively learn a common dictionary from noisy measure- ments. This approach may be useful in several con- texts including sensor networks. Diffusion cooperation schemes have been proposed to solve the distributed linear regression problem. In this work we focus on a diffusion-based adaptive dictionary learning strategy: each node records observations and cooperates with its neighbors by sharing its local dictionary. The resulting algorithm corresponds to a distributed block coordi- nate descent (alternate optimization). Beyond dictio- nary learning, this strategy could be adapted to many matrix factorization problems and generalized to var- ious settings. This article presents our approach and illustrates its efficiency on some numerical examples.

6.4. Applications

6.4.1. Medical Applications

Outlier detection for patient monitoring and alerting [5]

We develop and evaluate a data-driven approach for detecting unusual (anomalous) patient-management decisions using past patient cases stored in electronic health records (EHRs). Our hypothesis is that a patient-management decision that is unusual with respect to past patient care may be due to an error and that it is worthwhile to generate an alert if such a decision is encountered. We evaluate this hypothesis using data obtained from EHRs of 4486 post-cardiac surgical patients and a subset of 222 alerts generated from the data. We base the evaluation on the opinions of a panel of experts. The results of the study support our hypothesis that the outlier-based alerting can lead to promising true alert rates. We observed true alert rates that ranged from 25% to 66% for a variety of patient-management actions, with 66% corresponding to the strongest outliers.

6.5. Miscellaneous

6.5.1. Miscellaneous

A confidence-set approach to signal denoising [7]

The problem of filtering of finite-alphabet stationary ergodic time series is considered. A method for constructing a confidence set for the (unknown) signal is proposed, such that the resulting set has the following properties. First, it includes the unknown signal with probability γ, where γ is a parameter supplied to the filter. Second, the size of the confidence sets grows exponentially with a rate that is asymptotically equal to the conditional entropy of the signal given the data. Moreover, it is shown that this rate is optimal. We also show that the described construction of the confidence set can be applied to the case where the signal is corrupted by an erasure channel with unknown statistics.

Quantification adaptative pour la stéganalyse d'images texturées [28]

Nous cherchons à améliorer les performances d'un schéma de stéganalyse (i.e. la détection de messages cachées) pour des images texturées. Le schéma de stéganographie étudié consiste à modifier certains pixels de l'image par une perturbation +/-1, et le schéma de stéganalyse utilise les caractéristiques construites à partir de la probabilité conditionnelle empirique de différences de 4 pixels voisins. Dans sa version originale, la stéganalyse n'est pas trés efficace sur des images texturées et ce travail vise àâ€ explorer plusieurs techniques de quantification en utilisant d'abord un pas de quantification plus important puis une quantification adaptative scalaire ou vectorielle. Les cellules de la quantification adaptative sont générées en utilisant un K-means ou un K-means ”équilibré” de manière à ce chaque cellule quantifie approximativement le même nombre d'échantillon. Nous obtenons un gain maximal de classification de 3% pour un pas de quantification uniforme de 3. En utilisant l'algorithme K-means équilibré sur [-18,18], le gain par rapport à la version de base est de 4.7%.

Cost-sensitive Multiclass Classification Risk Bounds [8]

A commonly used approach to multiclass classification is to replace the 0-1 loss with a convex surrogate so as to make empirical risk minimization computationally tractable. Previous work has uncovered sufficient and necessary conditions for the consistency of the resulting procedures. In this paper, we strengthen these results by showing how the 0-1 excess loss of a predictor can be upper bounded as a function of the excess loss of the predictor measured using the convex surrogate. The bound is developed for the case of cost-sensitive multiclass classification and a convex surrogate loss that goes back to the work of Lee, Lin and Wahba. The bounds are as easy to calculate as in binary classification. Furthermore, we also show that our analysis extends to the analysis of the recently introduced "Simplex Coding" scheme.

Approximate Dynamic Programming Finally Performs Well in the Game of Tetris [12]

Tetris is a video game that has been widely used as a benchmark for various optimization techniques including approximate dynamic programming (ADP) algorithms. A look at the literature of this game shows that while ADP algorithms that have been (almost) entirely based on approximating the value function (value function based) have performed poorly in Tetris, the methods that search directly in the space of policies by learning the policy parameters using an optimization black box, such as the cross entropy (CE) method, have achieved the best reported results. This makes us conjecture that Tetris is a game in which good policies are easier to represent, and thus, learn than their corresponding value functions. So, in order to obtain a good performance with ADP, we should use ADP algorithms that search in a policy space, instead of the more traditional ones that search in a value function space. In this paper, we put our conjecture to test by applying such an ADP algorithm, called classification-based modified policy iteration (CBMPI), to the game of Tetris. Our experimental results show that for the first time an ADP algorithm, namely CBMPI, obtains the best results reported in the literature for Tetris in both small 10×10 and large 10×20 boards. Although the CBMPI's results are similar to those of the CE method in the large board, CBMPI uses considerably fewer (almost 1/6) samples (calls to the generative model) than CE.

A Generalized Kernel Approach to Structured Output Learning [15]

We study the problem of structured output learning from a regression perspective. We first provide a general formulation of the kernel dependency estimation (KDE) problem using operator-valued kernels. We show that some of the existing formulations of this problem are special cases of our framework. We then propose a covariance-based operator-valued kernel that allows us to take into account the structure of the kernel feature space. This kernel operates on the output space and encodes the interactions between the outputs without any reference to the input space. To address this issue, we introduce a variant of our KDE method based on the conditional covariance operator that in addition to the correlation between the outputs takes into account the effects of the input variables. Finally, we evaluate the performance of our KDE approach using both covariance and conditional covariance kernels on two structured output problems, and compare it to the state-of-the-art kernel-based structured output regression methods.

Gossip-based distributed stochastic bandit algorithms [24]

The multi-armed bandit problem has attracted remarkable attention in the machine learning community and many efficient algorithms have been proposed to handle the so-called exploitation-exploration dilemma in various bandit setups. At the same time, significantly less effort has been devoted to adapting bandit algorithms to particular architectures, such as sensor networks, multi-core machines, or peer-to-peer (P2P) environments, which could potentially speed up their convergence. Our goal is to adapt stochastic bandit algorithms to P2P networks. In our setup, the same set of arms is available in each peer. In every iteration each peer can pull one arm independently of the other peers, and then some limited communication is possible with a few random other peers. As our main result, we show that our adaptation achieves a linear speedup in terms of the number of peers participating in the network. More precisely, we show that the probability of playing a suboptimal arm at a peer in iteration $t=\Omega(\log N)$ is proportional to $1/(Nt)$ where N denotes the number of peers. The theoretical results are supported by simulation experiments showing that our algorithm scales gracefully with the size of network.

Sur quelques problèmes non-supervisés impliquant des séries temporelles hautement dèpendantes [1]

Cette thèse est consacrée à l'analyse théorique de problèmes non supervisés impliquant des séries temporelles hautement dépendantes. Plus particulièrement, nous abordons les deux problèmes fondamentaux que sont le problème d'estimation des points de rupture et le partitionnement de séries temporelles. Ces problèmes sont abordés dans un cadre extrêmement général oùles données sont générées par des processus stochastiques ergodiques stationnaires. Il s'agit de l'une des hypothèses les plus faibles en statistiques, comprenant non seulement, les hypothèses de modèles et les hypothèses paramétriques habituelles dans la littérature scientifique, mais aussi des hypothèses classiques d'indépendance, de contraintes sur l'espace mémoire ou encore des hypothèses de mélange. En particulier, aucune restriction n'est faite sur la forme ou la nature des dépendances, de telles sortes que les échantillons peuvent être arbitrairement dépendants. Pour chaque problème abordé, nous proposons de nouvelles méthodes non paramétriques et nous prouvons de plus qu'elles sont, dans ce cadre, asymptotiquement consistantes. Pour l'estimation de points de rupture, la consistance asymptotique se rapporte à la capacité de l'algorithme à produire des estimations des points de rupture qui sont asymptotiquement arbitrairement proches des vrais points de rupture. D'autre part, un algorithme de partitionnement est asymptotiquement consistant si le partitionnement qu'il produit, restreint à chaque lot de séquences, coïncides, à partir d'un certain temps et de manière consistante, avec le partitionnement cible. Nous montrons que les algorithmes proposés sont implémentables efficacement, et nous accompagnons nos résultats théoriques par des évaluations expérimentales. L'analyse statistique dans le cadre stationnaire ergodique est extrêmement difficile. De manière générale, il est prouvé que les vitesses de convergence sont impossibles à obtenir. Dès lors, pour deux échantillons générés indépendamment par des processus ergodiques stationnaires, il est prouvé qu'il est impossible de distinguer le cas où les échantillons sont générés par le même processus de celui où ils sont générés par des processus différents. Ceci implique que des problèmes tels le partitionnement de séries temporelles sans la connaissance du nombre de partitions ou du nombre de points de rupture ne peut admettre de solutions consistantes. En conséquence, une tâche difficile est de découvrir les formulations du problème qui en permettent une résolution dans ce cadre général. La principale contribution de cette thèse est de démontrer (par construction) que malgré ces résultats d'impossibilités théoriques, des formulations naturelles des problèmes considérés existent et admettent des solutions consistantes dans ce cadre général. Ceci inclut la démonstration du fait que le nombre de points de rupture corrects peut être trouvé, sans recourir à des hypothèses plus fortes sur les processus stochastiques. Il en résulte que, dans cette formulation, le problème des points de rupture peut être réduit à du partitionnement de séries temporelles. Les résultats présentés dans ce travail formulent les fondations théoriques pour l'analyse des données séquentielles dans un espace d'applications bien plus large.

Actor-Critic Algorithms for Risk-Sensitive MDPs [32]

In many sequential decision-making problems we may want to manage risk by minimizing some measure of variability in rewards in addition to maximizing a standard criterion. Variance-related risk measures are among the most common risk-sensitive criteria in finance and operations research. However, optimizing many such criteria is known to be a hard problem. In this paper, we consider both discounted and average reward Markov decision processes. For each formulation, we first define a measure of variability for a policy, which in turn gives us a set of risk-sensitive criteria to optimize. For each of these criteria, we derive a formula for computing its gradient. We then devise actor-critic algorithms for estimating the gradient and updating the policy parameters in the ascent direction. We establish the convergence of our algorithms to locally risk-sensitive optimal policies. Finally, we demonstrate the usefulness of our algorithms in a traffic signal control application.

Bayesian Policy Gradient and Actor-Critic Algorithms [31]

Policy gradient methods are reinforcement learning algorithms that adapt a parameterized policy by following a performance gradient estimate. Many conventional policy gradient methods use Monte-Carlo techniques to estimate this gradient. The policy is improved by adjusting the parameters in the direction of the gradient estimate. Since Monte-Carlo methods tend to have high variance, a large number of samples is required to attain accurate estimates, resulting in slow convergence. In this paper, we first propose a Bayesian framework for policy gradient, based on modeling the policy gradient as a Gaussian process. This reduces the number of samples needed to obtain accurate gradient estimates. Moreover, estimates of the natural gradient as well as a measure of the uncertainty in the gradient estimates, namely, the gradient covariance, are provided at little extra cost. Since the proposed Bayesian framework considers system trajectories as its basic observable unit, it does not require the dynamic within each trajectory to be of any special form, and thus, can be easily extended to partially observable problems. On the downside, it cannot take advantage of the Markov property when the system is Markovian. To address this issue, we then extend our Bayesian policy gradient framework to actor-critic algorithms and present a new actor-critic learning model in which a Bayesian class of non-parametric critics, based on Gaussian process temporal difference learning, is used. Such critics model the action-value function as a Gaussian process, allowing Bayes' rule to be used in computing the posterior distribution over action-value functions, conditioned on the observed data. Appropriate choices of the policy parameterization and of the prior covariance (kernel) between action-values allow us to obtain closed-form expressions for the posterior distribution of the gradient of the expected return with respect to the policy parameters. We perform detailed experimental comparisons of the proposed Bayesian policy gradient and actor-critic algorithms with classic Monte-Carlo based policy gradient methods, as well as with each other, on a number of reinforcement learning problems.

SERPICO Project-Team

6. New Results

6.1. Lifetime estimation in photon counting-based fluorescence lifetime imaging microscopy

Participants: Philippe Roudot, Charles Kervrann.

In this study, we investigated a Maximum Likelihood (ML) framework for photon counting-based fluorescence lifetime estimation in Fluorescence Lifetime Imaging Microscopy (FLIM). Data collected at a given pixel consist of photon counts exponentially decreasing along the time and are assumed to follow Poisson statistics (see Fig. 6). A careful analysis of the biophysical phenomenon and instrument models are used to derive a proper ML framework for lifetime estimation. Unlike usual pointwise approaches, a neighborhood-wise approach is proposed to take explicitly into account the spatial correlation of data [15]. The application to real biological data allowed us to prove the spatial localisation of interactions, a new result which was not achievable with conventional methods. For future work, the main challenge is to extend the framework to deal with multi-exponential decay estimate and adaptive neighborhoods, a challenge we need to address for a large class of biological studies.
Reference: [15]
Partners: A. Chessel (University of Cambridge, UK), F. Waharte and J. Boulanger (UMR 144, PICT IBiSA, CNRS-Institut Curie)

6.2. Vesicle segmentation method with automatic scale selection in TIRF microscopy

Participants: Antoine Basset, Charles Kervrann, Patrick Bouthemy.

Accurately detecting cellular structures in fluorescence microscopy is of primary interest for further quantitative analysis such as counting, tracking or classification. We aimed at segmenting vesicles in Total Internal Reflection Fluorescence (TIRF) microscopy images.

In this study, we have proposed an original and efficient method – called SLT-LoG – for vesicle segmentation with fewer parameters than the state-of-the-art methods. It exploits the Laplacian of Gaussian (LoG) of the images at several scales. Since the vesicles size is almost constant in space and time, a prominent mode is expected in the empirical distribution of the scales at which the minima of LoG values are detected. It precisely corresponds to the optimal sought scale. The vesicle segmentation map is then derived by thresholding the LoG values obtained at this optimal scale. To set the threshold, we assume that the values of the LoG locally follow a normal distribution (see Fig. 7). For each point, we estimate the local mean and variance, and the threshold is deduced from a user-selected probability of false alarm.

We have evaluated our method on classical synthetic sequences for which the performances of many detection methods are available [52], [56]. The comparative results on the dataset demonstrated that our method outperforms well-known unsupervised methods. We have also obtained very satisfactory results on real complex TIRF sequences.
Partners: Jean Salamero, J. Boulanger (UMR 144, PICT IBiSA, CNRS-Institut Curie)

6.3. Conditional random fields for vesicle traffic analysis with background estimation

Participants: Thierry Pécot, Patrick Bouthemy, Charles Kervrann.

../../../../projets/serpico/IMG/FLIM-TCSPC.png

Figure 6. Example of typical Time-Correlated Single Photon Counting (TCSPC) FLIM data. Total fluorescence intensity is shown in the center and corresponds to the sum of fluorescence intensities along the time axis at each pixel. The four side graphs correspond to time dependent photon counts in four different regions with variable sizes. By considering large regions, we observe an exponential decreasing along the time of fluorescence lifetime (see D). A: one pixel region; B and C: 3×3 *patches at different locations; D:* 15×15 *patch and liftetime estimation by least-square fitting.*

../../../../projets/serpico/IMG/segment.png

Figure 7. Segmentation method applied to a real TIRF microscopy sequence showing the Rab11-mCherry protein. The SLT-LoG method is able to provide the entire spatial support of the vesicles.

../../../../projets/serpico/IMG/CRFS/orig-a.png ../../../../projets/serpico/IMG/CRFS/orig-b.png ../../../../projets/serpico/IMG/CRFS

Figure 8. Left: Real fluorescence microscopy image depicting GFP-Rab6 proteins. Center: estimated vesicular component. Right: estimated background component.

Image analysis applied to fluorescence live cell microscopy has become a key tool in molecular biology since it enables to characterize biological processes in space and time at the subcellular level. In fluorescence microscopy imaging, the moving tagged structures of interest, such as vesicles, appear as bright spots over a static or non-static background. In this work, we consider the problem of vesicle segmentation and time-varying background estimation at the cellular scale. The main idea is to formulate the joint segmentation-estimation problem in the general Conditional Random Field (CRF) framework. Furthermore, segmentation of vesicles and background estimation are alternatively performed by energy minimization using a min cut-max flow algorithm. The proposed approach relies on a detection measure computed from intensity contrasts between neighboring patches in fluorescence microscopy images. We have demonstrated the competitiveness of the proposed method through an experimental comparison with state-of-the-art methods in fluorescence videomicroscopy, for single cell studies. We have also characterized the density of Rab6 transport carriers spatially dispersed at the cell periphery, for two different specific adhesion geometries.

Partners: Jean Salamero, J. Boulanger (UMR 144, PICT IBiSA, CNRS-Institut Curie)

6.4. Exemplar-based occlusion handling and sparse continuous aggregation for optical flow computation

Participants: Denis Fortun, Patrick Bouthemy, Charles Kervrann.

Handling large displacements, motion details and occlusions all together remains an open issue for reliable computation of optical flow in a video sequence. Our recently investigated aggregation paradigm is an attractive approach supplying motion candidates at every pixel in a first step, and combining them in a second step to determine the global optical flow field [16]. We experimentally demonstrate that simple and purely local parametric estimations combined with patch correspondences are sufficient to produce highly accurate motion candidates. Nevertheless, the performances are limited by the presence of large occlusion areas. Therefore we have proposed an exemplar-based occlusion handling scheme integrated in the two steps of the aggregation process. At the first stage, local motion candidates sets are extended at the detected occluded pixels with candidates from non-occluded pixels, and specific occlusions due to camera motion are handled by estimating the dominant motion in the image. Local occlusion cues are extracted from this first step. Then, we define a global energy function which cooperatively selects the best motion candidates for each point while recovering the occlusion areas and ensuring smoothness properties. Results on small displacement sequences are competitive with state-of-the-art methods, and great improvements are observed in the case of large displacements and occlusions (Fig. 9).

Alternatively to the discrete aggregation based on graph cut optimization, a new continuous aggregation model has been designed. In accordance with the demonstrated evidence that the set of candidates always contains at least one accurate motion vector, the aggregation is formulated in a sparse framework restricting the number of non negligible weights associated to the candidates. The continuous framework is less dependent on the quality of the candidates and thus allows us to considerably reduce the computational cost of both aggregation and candidates estimation.
Reference: [16]

6.5. Correlation and variational approaches for motion and diffusion estimation

Participants: Denis Fortun, Charles Kervrann.

Diffusion coefficient estimation in live cell fluorescence imaging is usually achieved with correlation-based methods related to Image Correlation Spectroscopy (ICS) [42]. This approach requires a high computational cost and the spatial resolution of the resulting diffusion map is limited by the inherent block-based principle of the method. To overcome these drawbacks, we propose a novel diffusion estimation method in a variational framework providing dense and discontinuity-preserving diffusion fields. The diffusion equation is integrated in a global energy via a neighborhood-wise data term, positivity constraint and temporal integration. The

Image I_1

Image I_2

Optical flow ground-truth

Our method

Weinzaepfel et al. [63]

Xu et al. [64]

Occlusion ground truth

Occlusions recovered by our method

Figure 9. Comparative evaluation of optical flow estimation for large displacements between our method, [63] and [64]. First row : two successive frames I_1 and I_2 and the ground truth motion field; second row: comparative estimation results; third row: evaluation of our occlusion map estimation.

performances of the variational and ICS approaches were compared on simulated sequences. We have demonstrated the accuracy of ICS in stationarity conditions, and we pointed out the advantages of dense variational estimation to accurately recover spatial and temporal discontinuities (Fig. 10).
Reference: [17]
Partners: Perrine Paul-Gilloteaux, Francois Waharte and Chen Chen (UMR 144, PICT IBiSA, CNRS-Institut Curie)

6.6. Classification of membrane dynamics in TIRF microscopy

Participants: Antoine Basset, Charles Kervrann, Patrick Bouthemy.

Recognizing dynamic protein behaviors in live cell fluorescence microscopy is of paramount importance to understand cell mechanisms. In the case of membrane traffic, cargo molecules are transferred from a donor to an acceptor compartments [49]. At each step, dedicated molecular platforms are acting to form, transport and address selected proteins. In microscopy imaging, this sequence of processes leads to a series of heterogeneous dynamics, which need to be untangled in order to understand the spatiotemporal coordination of the molecular actors. In this study, we aim at locating and recognizing temporal events in TIRF microscopy image sequences related to membrane dynamics. After segmenting the time-varying vesicles in the image, we exploit space-time information extracted from three successive images only to model, locate and recognize the two dynamic configurations of interest: translational motion or local fluorescence diffusion (see Fig. 11). A likelihood ratio test is defined to solve this issue. Results on synthetic sequences and real TIRF sequences demonstrated the accuracy and efficiency of the proposed method.
Partners: Jean Salamero, J. Boulanger (UMR 144, PICT IBiSA, CNRS-Institut Curie)

6.7. Crowd motion classification

Participants: Antoine Basset, Charles Kervrann, Patrick Bouthemy.

Important research efforts have been devoted to crowd analysis for several years [58], [65]. We are interested in this topic for two main reasons. First, views of crowded scenes are not that different of light microscopy intracellular images. Second, the addressed problem, i.e. motion understanding, is common, and we are investigating similar data-driven methodological approaches. This a way to cross-fertilize two domains.

We address the problem of classifying coherent crowd motions in videos recorded by a fixed camera. In contrast to most existing methods, which are based on trajectories or tracklets, our approach for crowd motion analysis provides a crowd motion classification on a frame-by-frame basis. Indeed, we only compute affine motion models from pairs of two consecutive video images. The classification itself relies on simple rules on the coefficients of the computed affine motion models, and therefore does not imply any prior learning stage. The overall method proceeds in three steps: we first compute a set of motion model candidates on a collection of windows of different sizes in the image, then we select the motion model at each point owing to a ML criterion, finally we determine the crowd motion class map with a hierarchical classification tree regularized by majority votes. The algorithm is almost parameter-free, and is extremely efficient in terms of memory and computation load. Experiments on computer-generated sequences [28] and real video sequences demonstrate that our method is accurate, and can successfully handle complex situations (see Fig. 12).
References:[14], [23]

6.8. Estimation of the flow of particles without tracking algorithm in fluorescence imaging

Participants: Thierry Pécot, Patrick Bouthemy, Charles Kervrann.

../../../../projets/serpico/IMG/diff.png

(a) Frame 1

../../../../projets/serpico/IMG/mask_diff

(b) Ground-truth diffusion field

../../../../projets/serpico/IMG/profile2.png

(c) Estimated diffusion field D

../../../../projets/serpico/IMG/histogram

(d) Histogram of D

../../../../projets/serpico/IMG/profile.png

(e) Segmented diffusion field

../../../../projets/serpico/IMG/plot9v2.p

(f) Profiles

Figure 10. Variational diffusion estimation on a simulated sequence with spatially variant diffusion. The curves of (f) are profiles of the dashed lines in (b),(c) and (e)

../../../../projets/serpico/IMG/VesicleMotionTypes.png

Figure 11. Classification results for a real TIRF sequence, whose estimated PSNR is 28.6. Results are displayed for a representative frame. The only classification error – framed in yellow – is a diffusion classified as translation. However, this vesicle has a very low intensity and changes its shape while diffusing. Two vesicles framed in green are detected as a single connected component. The vesicle framed in red corresponds to diffusing vesicles

../../../../projets/serpico/IMG/MarathonMotion.png

Figure 12. Two frames of the Marathon bend sequence. People run from upper left to upper right, describing a U. The movement is quite constant in the whole sequence and so is the classification: in the left branch, people go South (magenta), then turn counterclockwise (red) until the end of the bend. Some Eastward translation (yellow) is sometimes found here because of the large radius of curvature. Finally the North translation is recovered (blue). The points in the upper right corner of the image are classified as translations to the West (purple), but the translation direction is closer to North than to West (North-North-West): it is also due to the lateral presence of pedestrians walking to the left.

../../../../projets/serpico/IMG/EFP.png

Figure 13. Vesicle flows estimated with our method when considering a simple partition of 5 regions for an image sequence acquired in TIRF microscopy and showing the protein Clip170.

Automatic analysis of the dynamic content in fluorescence video-microscopy is crucial for understanding molecular mechanisms involved in cell functions. We have proposed an original approach for analyzing particle trafficking in these sequences. Instead of individually tracking every particle, we only need to locally count particles on regions over time and minimize a global energy function. We have specified three methods to determine the particle flow. We especially compared the NNLS algorithm [44] and the PPXA algorithm [33] known as well suited to non differentiable convex minimization problem [24]. We have conducted comparative experiments on synthetic and real fluorescence image sequences. We have shown that adding a sparsity constraint on the number of detected events allows us to reduce the number of false alarms. Compared to usual tracking methods, our approach is simpler and the results are very stable with respect to the only two parameters involved (see Fig. 13).
Reference: [24]

Partners: Jean Salamero, J. Boulanger (UMR 144, PICT IBiSA, CNRS-Institut Curie)

6.9. Probabilistic Tracking of fluorescent objects

Participants: Philippe Roudot, Charles Kervrann.

Image tracking of fluorescent objects, from labeled molecules to organelles and entire cells, is an essential task in the analysis of cellular functions. During the last decade, several algorithms have been tailored to cope with different types of cellular and subcellular motion down to Brownian single molecule behavior [8]. One of the remaining big challenges in this area of technology development has been the tracking of extremely heterogeneous movements of objects in crowded scenes. We tested several state-of-the-art algorithms [36], [40] to follow dense populations of diffusing particles, which suddenly change to directed motion. A frequent cellular scenario with this property is the jerky motion of vesicles and viruses switching between cytoplasmic diffusion and motor-mediated, fast displacements (see Fig. 14).

../../../../projets/serpico/IMG/VimentinMotion.png

Figure 14. Vimentin motility seems to present a large proportion of confined Brownian motion and rare, sudden, motor-mediated transport. Colored tracks have been computed with an advanced U-track parametrization (Unit length filament of Vimentin Y117L mutant fused to GFP and transfected into vimentin null epithelial cell (cell line SW13). Image acquired with a spinning disk confocal microscope with a 100x objective zoom 1.5 (Numerical Aperture 1.4, pixel size 0.10905μm/pixel).

../../../../projets/serpico/IMG/TrackPerf.png

Figure 15. A) Example of tracks simulation presenting a density of 3 spots/μm^2. B) Correct linking percentage wrt density and motion type switching probability. Our method outperforms U-track by 15% in the hardest case. C) True positive and false positive ratio on the same simulation with a density of 3 spots/μm^2, comparing our method with U-track, U-track with an on-line process noise estimator and an IMM algorithm with forward-backward initialization.

../../../../projets/serpico/IMG/SpeedHet.png

Figure 16. Correct linking and false positive percentage wrt speed switching probability.

These switches are particularly challenging to detect because they occur rarely. The presence of numerous detected objects in the expected range of particle displacements makes the tracking ambiguous and induces wrong associations. Lowering the ambiguity by reducing the search range, on the other hand, is not an option, as this would increase the rate of false negatives.

We first explored the existing methods in the literature to analyze their strenghts and weakness for tracking objects with heterogeneous motion and high density. Based on the conclusion we draw, we proposed a new method build on the U-track platform [40]. More specifically, we propose an interacting multiple state model that exploits recursive tracking in multiple rounds in forward and backward temporal directions. As a result, it achieves convergence of the instantaneous speed estimate time-point-by-time-point. This allows us to predict and recover abrupt transitions from freely or confined diffusive to directed motion. To address the issue of a particle that disappears as a neighboring particle appears in the same image and thus to better detect track termination, we also exploit this recursive tracking by proposing a locally adaptive on-line estimation of the search window radius for assignment (a.k.a. gating), while most of state-of-the-art algorithms propose only a global search window radius or weak per-track search radius estimations. We have shown on simulated data that our method outperforms state-of-the-art algorithms that model motion heterogeneity on different scenarios, e.g. heterogeneous motion type (see Figure 15) and speed heterogeneity (see Fig. 16), while keeping the computational cost of a deterministic method (10% overhead with respect to U-track).

Partners: Gaudenz Danuser (Harvard Medical School, Boston, USA)

6.10. Microtubules modeling for variational assimilation analysis

Participants: Pierre Allain, Charles Kervrann.

Microtubules (MT) are highly dynamic tubulin polymers that are involved in many cellular processes such as mitosis, intracellular cell organization and vesicular transport. Nevertheless, the modeling of cytoskeleton and MT dynamics based on physical properties is difficult to achieve. We proposed to model microtubules as rigid and growing cylinders alike (Nedelec and Foethk 2007) [45] but including Newtonian dynamics. Using the Euler-Bernoulli beam theory, we have proposed then to model the rigidity of microtubules on a physical basis using forces, mass and acceleration. In addition, we linked microtubules growth and shrinkage to the presence of molecules (e.g. GTP-tubulin) in the cytosol. The overall model enables linking cytosol to microtubules dynamics in a constant state space, thus allowing usage of data assimilation techniques (see Fig. 17).

../../../../projets/serpico/IMG/MT-simu2D.png ../../../../projets/serpico/IMG/MT-simu3D.png

Figure 17. Left: Simulation of a 2D radial microtubule network. The results show growing and shrinking phases yielding inhomogeneous “pseudo-tubulin” concentration in the cytosol. MTs are bended according to fluid forces. Right: 3D simulation of MT nucleation and growth that mimics MT dynamics seeded onto a two vertical bar-shaped fibronectin pattern and observed in TIRF microscopy (courtesy of iRTSV/LPCV/PCM CEA-Grenoble).

6.11. Spot localization for TMA image analysis

Participants: Nam-Hoai Nguyen, Charles Kervrann.

A very first task of TMA (Tissue MicroArray) image analysis is to accurately localize spots (separate tissue core) representing arrays of 512 x 512 pixels each, in very large images of several thousands of pixels. For this purpose, we have investigated a three-stage methodological approach. First, since tissue cores are separately assembled in array (grid structure). We started to design a graphical model to eliminate image defects due to the presence of dusts or the imperfection of TMA blocks fabrication. In the second stage, a wavelet-like transform is currently used to recognize interested features (spots) given the size of spots *a prori*. Third, we started to investigate the superpixel-based image representation (SLIC) [27], [55] to handle very large images and biological details inside each spot.

Partners: V. Paveau (Innopys company)

SHACRA Project-Team

6. New Results

6.1. Electrophysiology

Cardiac arrhythmia is a very frequent pathology that comes from an abnormal electrical activity in the myocardium. This work aims at developing a training simulator for interventional radiology and thermo-ablation of these arrhythmias. After tackling the issue of fast electrophysiology, a first version of our training simulator was proposed.

../../../../projets/shacra/IMG/Electrophysiology.png

Figure 3. Cardiac electrophysiology computed on a patient-specific geometry

The first main contribution of this work is the interactive catheter navigation inside a moving venous system and a beating heart. The virtual catheterization reproduces navigation issues that can be solved using a bending catheter. Second, our real-time GPU electrophysiology model allows interactions during the simulation such as extra-cellular potential measurement, RF ablation, and electrical stimulation. An innovative management of the computational units based on multithreading offers performances close to real-time. This framework is therefore a substantial step towards realistic and highly efficient virtual training systems in cardiology. As future work, we intend to use patient-specific data in our framework so that cardiologists could quantitatively assess the realism of our virtual training.

(a) Setup of the System (b) Simulation of Catheter Navigation (c) Electrophysiological Signals

Figure 4. The first simulation dedicated to electrocardiology training

6.2. Cryoablation

A new project started this year around cryotherapy. This technique consists in inserting needles that freezing the surrounding tissues, thus immediately leading to cellular death of the tissues. Cryoablation procedure is used in many medical fields for tumor ablation, and even starts being used in cardiology. In this scope, we build a simulator able to place the cryoprobes and run a simulation representing the evolution of iceballs in living tissues.

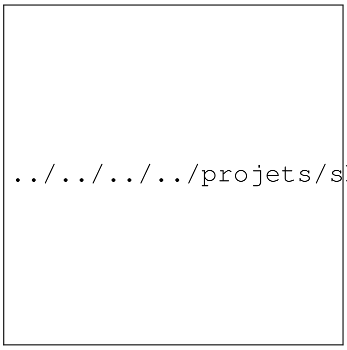


Figure 5. Simulation framework for cryoablation planning

6.3. Stapedotomy

Stapedotomy is a challenging procedure of the middle ear microsurgery, since the surgeons is in direct contact with sensitive structures such as the ossicular chain. This procedures is taught and performed in the last phase of the surgical apprenticeship. To improve surgical teaching, we propose to use a virtual surgical simulator based on a finite element model of the middle ear. The static and dynamic behavior of the developed finite element model was successfully compared to published data on human temporal bones specimens. A semi-automatic algorithm was developed to perform a quick and accurate registration of our validated mechanical atlas to match the patient dataset. This method avoids a time-consuming work of manual segmentation, parameterization, and evaluation. A registration is obtained in less than 260 seconds with an accuracy close to a manual process and within the imagery resolution. The computation algorithms, allowing carving, deformation of soft and hard tissus, and collision response, are compatible with a real-time interactive simulation of a middle ear procedure. As a future work, we propose to investigate new robotized procedures of the middle ear surgery in order to develop new applications for the RobOtol device and to provide a training tool for the surgeons.

../../../../projets/shacra/IMG/ENT.png

Figure 6. Simulation of the stapedotomy procedure.

6.4. Radiotherapy planning

The main challenge of radiotherapy treatment is to irradiate the tumor while sparing the surrounding healthy tissues. In the case of throat cancer, the complexity of the therapy treatment is due to the proximity of organs at risk such as the two parotide glands. The parotide glands are the main salivary glands. An overdose of radiation in these glands may cause xerostomia, which is a medical term for the symptom of dryness in the mouth, or in other words, a lack of saliva. This disease affect significantly the life of the patient: difficulty talking, tasting, chewing, swallowing, excessive thirst, constant pain in the throat etcetera. A radiation therapy treatment of throat cancer takes from 5 to 7 weeks. The treatment is planned several days before the therapy. The planning consist in contouring each organ of the area on CT-scan images and defining the dose of radiation to deliver to each of these organs. This stage is lengthy and takes around two hours per patient. Yet, some anatomical variations occur in the course of the treatment, mainly due to the weight loss of the patients. These variations compromise the safety of the healthy tissues, because the planned treatments is no more up to date. For now, the physicians have no solution goog enough to handle these changes. Xerostomia affects around 20 per-cent of the patients suffering from throat cancer.

The main idea of this work is to create an interface that the physicians could use to redo the planning when it is needed, when the anatomical changes are significant. The purpose is to give to them the possibility to use what they see on images, to recreate the right shape of the contours without recontouring each images, and in a reasonable time. This interface will use their knowledge to determine the new shape of the organs. The work does not aim at providing a fully automatic method because it would reduce its acceptation by the physicians. As the method is based on the input of the physician, they can control the deformation based on images but also on their knowledge.

6.5. Image-based diagnoses

In the context of the female pelvic medicine, image-based diagnoses of pelvic floor disorders like prolapse or endometriosis rely on mechanical indicators, such as mobilities of organs and shear displacements between

../../../../projets/shacra/IMG/parotide22.png

Figure 7. Screenshot of our radotherapy planning tool.

organs. This information would be useful for both precise diagnoses and planning of surgical procedure. Involving numerical tools for diagnoses and surgery planing becomes increasingly interesting for physicians in clinical uses. The advantages of numerical models are not only in visualization, but also in quantitative measurements on a group of organs, such as their shapes and their relative movements. The processing pipeline includes patient data retrieval, image analysis, patient-specific modeling and biomechanical simulation. Our work consists in proposing new methods and algorithms for modeling the 3D anatomy of specific patients based on image data. This model should be compatible with the requirements of a biomechanical simulation. Moreover, we aim at developing new image processing tools for analyzing 2D dynamic MRI (to assess the mobilities of the pelvic system by extracting certain mechanical indicators from images) and for comparison with simulations.

Registration between geometric models and images remains a major challenge in these applications. We proposed a new model-to-image registration approach which was developed and tested for segmentation of organs in 2D images and for tracking the motion of pelvic organs from 2D dynamic MRI. Thanks to this technique, evaluation of the level of shear strain that is encountered by the fascias (connective tissues between organs) during the motion became possible. This tool could help in early diagnostic of prolapse. In the next step, our objective is to extend this method for adapting it to 3D reconstruction (with 3D geometric models and 3D MR images) and for the comparison of 3D simulations with deformable images.

6.6. Dynamic Deformations Simulated at Different Frequencies

The dynamic response of deformable bodies varies significantly in dependence on mechanical properties of the objects: while the dynamics of a stiff and light object (e. g. wire or needle) involves high-frequency phenomena such as vibrations, much lower frequencies are sufficient for capturing dynamic response of an object composed of a soft tissue. Yet, when simulating mechanical interactions between soft and stiff deformable models, a single time-step is usually employed to compute the time integration of dynamics of both objects. However, this can be a serious issue when haptic rendering of complex scenes composed of various bodies is considered. In this work, we present a novel method allowing for dynamic simulation of a scene composed of colliding objects modelled at different frequencies: typically, the dynamics of soft objects are calculated at frequency about 50 Hz, while the dynamics of stiff object is modeled at 1 kHz, being directly connected to the computation of haptic force feedback. The collision response is performed at both low and high frequencies employing data structures which describe the actual constraints and are shared between the high and low frequency loops. During the simulation, the realistic behaviour of the objects according to the mechanical principles (such as non-interpenetration and action-reaction principle) is guaranteed. We have shown several scenarios involving different bodies in interaction, demonstrating the benefits of the proposed method. This research has been published at IROS 2013.

6.7. Simulation of Lipofilling Reconstructive Surgery

We have developed a method to simulate the outcome of reconstructive facial surgery based on fat-filling. Facial anatomy is complex: the fat is constrained between layers of tissues which behave as walls along the face; in addition, connective tissues that are present between these different layers also influence the fat-filling procedure. To simulate the end result, we have proposed a method which couples a 2.5D Eulerian fluid model for the fat and a finite element model for the soft tissues. The two models are coupled using the computation of the mechanical compliance matrix. We had two contributions: a solver for fluids which couples properties of solid tissues and fluid pressure, and an application of this solver to fat-filling surgery procedure simulation. This research has been published at MICCAI 2013.

6.8. Real-time simulation of contact and cutting of heterogeneous soft-tissues

We have developed a new numerical method for interactive (real-time) simulations, which considerably improves the accuracy of the response of heterogeneous soft-tissue models undergoing contact, cutting and other topological changes. It provides an integrated methodology able to deal both with the ill-conditioning issues associated with material heterogeneities, contact boundary conditions which are one of the main sources of inaccuracies, and cutting which is one of the most challenging issues in interactive simulations. Our approach is based on an implicit time integration of a non-linear finite element model. To enable real-time computations, we propose a new preconditioning technique, based on an asynchronous update at low frequency. The preconditioner is not only used to improve the computation of the deformation of the tissues, but also to simulate the contact response of homogeneous and heterogeneous bodies with the same accuracy. We also address the problem of cutting the heterogeneous structures and propose a method to update the preconditioner according to the topological modifications. Finally, we have applied our approach to three challenging demonstrators: i) a simulation of cataract surgery ii) a simulation of laparoscopic hepatectomy iii) a brain tumor surgery. This research was done in collaboration with the University of Cardiff and has been published in the journal Media this year.

6.9. Control of Elastic Soft Robots

In this work, we present a new method for the control of soft robots with elastic behavior, piloted by several actuators. The central contribution of this work is the use of the Finite Element Method (FEM), computed in real-time, in the control algorithm. The FEM based simulation computes the nonlinear deformations of the robots at interactive rates. The model is completed by Lagrange multipliers at the actuation zones and at the end-effector position. A reduced compliance matrix is built in order to deal with the necessary inversion of the model. Then, an iterative algorithm uses this compliance matrix to find the contribution of the actuators (force and/or position) that will deform the structure so that the terminal end of the robot follows a given position. Additional constraints, like rigid or deformable obstacles, or the internal characteristics of the actuators are integrated in the control algorithm. We illustrate our method using simulated examples of both serial and parallel structures and we validate it on a real 3D soft robot made of silicone.

SIERRA Project-Team

6. New Results

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

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

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

In [16] 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 a similar 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, it yields an online algorithm that has the same low iteration complexity as primal stochastic subgradient methods. However, unlike stochastic subgradient methods, the block-coordinate 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.

6.2. Minimizing Finite Sums with the Stochastic Average Gradient.

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

In [32] we propose the stochastic average gradient (SAG) method for optimizing the sum of a finite number of smooth convex functions. Like stochastic gradient (SG) methods, the SAG method's iteration cost is independent of the number of terms in the sum. However, by incorporating a memory of previous gradient values the SAG method achieves a faster convergence rate than black-box SG methods. The convergence rate is improved from $O(1/\sqrt{k})$ to $O(1/k)$ in general, and when the sum is strongly-convex the convergence rate is improved from the sub-linear $O(1/k)$ to a linear convergence rate of the form $O(\rho^k)$ for $\rho < 1$. Further, in many cases the convergence rate of the new method is also faster than black-box deterministic gradient methods, in terms of the number of gradient evaluations. Numerical experiments indicate that the new algorithm often dramatically outperforms existing SG and deterministic gradient methods.

The primary contribution of this work is the analysis of a new algorithm that we call the *stochastic average gradient* (SAG) method, a randomized variant of the incremental aggregated gradient (IAG) method of [43]. The SAG method has the low iteration cost of SG methods, but achieves the convergence rates stated above for the FG method. The SAG iterations take the form

$$x^{k+1} = x^k - \frac{\alpha_k}{n}\sum_{i=1}^{n} y_i^k, \tag{51}$$

where at each iteration a random index i_k is selected and we set $y_i^k = f_i'(x^k)$ if $i = i_k$, and y_i^{k-1} otherwise. That is, like the FG method, the step incorporates a gradient with respect to each function. But, like the SG method, each iteration only computes the gradient with respect to a single example and the cost of the iterations is independent of n. Despite the low cost of the SAG iterations, we show in this paper that with a constant step-size *the SAG iterations have an $O(1/k)$ convergence rate for convex objectives and a linear convergence rate for strongly-convex objectives*, like the FG method. That is, by having access to i_k and by keeping a *memory* of the most recent gradient value computed for each index i, this iteration achieves a faster convergence rate than is possible for standard SG methods. Further, in terms of effective passes through the data, we will also see that for many problems the convergence rate of the SAG method is also faster than is possible for standard FG methods.

6.3. Fast Convergence of Stochastic Gradient Descent under a Strong Growth Condition

Participants: Mark Schmidt [correspondent], Nicolas Le Roux [correspondent].

In [33] we consider optimizing a function smooth convex function f that is the average of a set of differentiable functions f_i, under the assumption considered by [87] and [90] that the norm of each gradient f_i' is bounded by a linear function of the norm of the average gradient f'. We show that under these assumptions the basic stochastic gradient method with a sufficiently-small constant step-size has an $O(1/k)$ convergence rate, and has a linear convergence rate if g is strongly-convex.

We write our problem

$$\min_{x \in \mathbb{R}^P} f(x) := \frac{1}{N}\sum_{i=1}^{N} f_i(x), \tag{52}$$

where we assume that f is convex and its gradient f' is Lipschitz-continuous with constant L, meaning that for all x and y we have

$$||f'(x) - f'(y)|| \leq L||x - y||.$$

If f is twice-differentiable, these assumptions are equivalent to assuming that the eigenvalues of the Hessian $f''(x)$ are bounded between 0 and L for all x.

Deterministic gradient methods for problems of this form use the iteration

$$x_{k+1} = x_k - \alpha_k f'(x_k), \tag{53}$$

for a sequence of step sizes α_k. In contrast, *stochastic gradient* methods use the iteration

$$x_{k+1} = x_k - \alpha_k f_i'(x_k), \tag{54}$$

for an individual data sample i selected uniformly at random from the set $\{1, 2, \cdots, N\}$.

The stochastic gradient method is appealing because the cost of its iterations is *independent of*N. However, in order to guarantee convergence stochastic gradient methods require a decreasing sequence of step sizes $\{\alpha_k\}$ and this leads to a slower convergence rate. In particular, for convex objective functions the stochastic gradient method with a decreasing sequence of step sizes has an expected error on iteration k of $O(1/\sqrt{k})$ [78], meaning that

$$\mathbb{E}[f(x_k)] - f(x^*) = O(1/\sqrt{k}).$$

In contrast, the deterministic gradient method with a *constant* step size has a smaller error of $O(1/k)$ [79]. The situation is more dramatic when f is *strongly* convex, meaning that

$$f(y) \geq f(x) + \langle f'(x), y - x\rangle + \frac{\mu}{2}||y - x||^2, \tag{55}$$

for all x and y and some $\mu > 0$. For twice-differentiable functions, this is equivalent to assuming that the eigenvalues of the Hessian are bounded below by μ. For strongly convex objective functions, the stochastic gradient method with a decreasing sequence of step sizes has an error of $O(1/k)$ [77] while the deterministic method with a constant step size has an *linear* convergence rate. In particular, the deterministic method satisfies

$$f(x_k) - f(x^*) \leq \rho^k [f(x_0) - f(x^*)],$$

for some $\rho < 1$ [71].

We show that if the individual gradients $f_i'(x_k)$ satisfy a certain strong growth condition relative to the full gradient $f'(x_k)$, the stochastic gradient method with a sufficiently small constant step size achieves (in expectation) the convergence rates stated above for the deterministic gradient method.

6.4. Non-strongly-convex smooth stochastic approximation with convergence rate $O(1/n)$

Participants: Eric Moulines, Francis Bach [correspondent].

Large-scale machine learning problems are becoming ubiquitous in many areas of science and engineering. Faced with large amounts of data, practitioners typically prefer algorithms that process each observation only once, or a few times. Stochastic approximation algorithms such as stochastic gradient descent (SGD) and its variants, although introduced more than sixty years ago, still remain the most widely used and studied method in this context. In [8], we consider the stochastic approximation problem where a convex function has to be minimized, given only the knowledge of unbiased estimates of its gradients at certain points, a framework which includes machine learning methods based on the minimization of the empirical risk. We focus on problems without strong convexity, for which all previously known algorithms achieve a convergence rate for function values of $O(1/\sqrt{n})$ after n iterations. We consider and analyze two algorithms that achieve a rate of $O(1/n)$ for classical supervised learning problems. For least-squares regression, we show that *averaged* stochastic gradient descent *with constant step-size* achieves the desired rate. For logistic regression, this is achieved by a simple novel stochastic gradient algorithm that (a) constructs successive local quadratic approximations of the loss functions, while (b) preserving the same running-time complexity as stochastic gradient descent. For these algorithms, we provide a non-asymptotic analysis of the generalization error (in expectation, and also in high probability for least-squares), and run extensive experiments showing that they often outperform existing approaches.

6.5. Streaming Bayesian Inference

Participant: Michael Jordan [correspondent].

Large, streaming data sets are increasingly the norm in science and technology. Simple descriptive statistics can often be readily computed with a constant number of operations for each data point in the streaming setting, without the need to revisit past data or have advance knowledge of future data. But these time and memory restrictions are not generally available for the complex, hierarchical models that practitioners often have in mind when they collect large data sets. Significant progress on scalable learning procedures has been made in recent years. But the underlying models remain simple, and the inferential framework is generally non-Bayesian. The advantages of the Bayesian paradigm (e.g., hierarchical modeling, coherent treatment of uncertainty) currently seem out of reach in the Big Data setting.

An exception to this statement is provided by Hofmann et al. (2010), who have shown that a class of approximation methods known as *variational Bayes* (VB) can be usefully deployed for large-scale data sets. They have applied their approach, referred to as *stochastic variational inference* (SVI), to the domain of topic modeling of document collections, an area with a major need for scalable inference algorithms. VB traditionally uses the variational lower bound on the marginal likelihood as an objective function, and the idea of SVI is to apply a variant of stochastic gradient descent to this objective. Notably, this objective is based on the conceptual existence of a full data set involving D data points (i.e., documents in the topic model setting), for a fixed value of D. Although the stochastic gradient is computed for a single, small subset of data points (documents) at a time, the posterior being targeted is a posterior for D data points. This value of D must be specified in advance and is used by the algorithm at each step. Posteriors for D' data points, for D' not equal to D, are not obtained as part of the analysis.

We view this lack of a link between the number of documents that have been processed thus far and the posterior that is being targeted as undesirable in many settings involving streaming data. In this project we aim at an approximate Bayesian inference algorithm that is scalable like SVI but is also truly a streaming procedure, in that it yields an approximate posterior for each processed collection of D' data points—and not just a pre-specified "final" number of data points D. To that end, we return to the classical perspective of Bayesian updating, where the recursive application of Bayes theorem provides a sequence of posteriors, not a sequence of approximations to a fixed posterior. To this classical recursive perspective we bring the VB framework; our updates need not be exact Bayesian updates but rather may be approximations such as VB.

Although the empirical success of SVI is the main motivation for our work, we are also motivated by recent developments in computer architectures, which permit distributed and asynchronous computations in addition to streaming computations. A streaming VB algorithm naturally lends itself to distributed and asynchronous implementations.

6.6. Convex Relaxations for Permutation Problems

Participants: Fajwel Fogel [correspondent], Rodolphe Jenatton, Francis Bach, Alexandre d'Aspremont.

Seriation seeks to reconstruct a linear order between variables using unsorted similarity information. It has direct applications in archeology and shotgun gene sequencing for example. In [12] we prove the equivalence between the seriation and the combinatorial 2-sum problem (a quadratic minimization problem over permutations) over a class of similarity matrices. The seriation problem can be solved exactly by a spectral algorithm in the noiseless case and we produce a convex relaxation for the 2-sum problem to improve the robustness of solutions in a noisy setting. This relaxation also allows us to impose additional structural constraints on the solution, to solve semi-supervised seriation problems. We performed numerical experiments on archeological data, Markov chains and gene sequences.

6.7. Phase retrieval for imaging problems

Participants: Fajwel Fogel [correspondent], Irène Waldspurger, Alexandre d'Aspremont.

In [29] we study convex relaxation algorithms for phase retrieval on imaging problems. We show that structural assumptions on the signal and the observations, such as sparsity, smoothness or positivity, can be exploited to both speed-up convergence and improve recovery performance. We detail experimental results in molecular imaging problems simulated from PDB data.

Phase retrieval seeks to reconstruct a complex signal, given a number of observations on the *magnitude* of linear measurements, i.e. solve

$$\begin{array}{ll} \text{find} & x \\ \text{such that} & |Ax| = b \end{array}$$

in the variable x, where A and b. This problem has direct applications in X-ray and crystallography imaging, diffraction imaging, Fourier optics or microscopy for example, in problems where physical limitations mean detectors usually capture the intensity of observations but cannot recover their phase. In this project, we focus on problems arising in diffraction imaging, where A is usually a Fourier transform, often composed with one or multiple masks (a technique sometimes called ptychography). The Fourier structure, through the FFT, often considerably speeds up basic linear operations, which allows us to solve large scale convex relaxations on realistically large imaging problems. We also observe that in most of the imaging problems we consider, the Fourier transform is very sparse, with known support (we lose the phase but observe the magnitude of Fourier coefficients), which allows us to considerably reduce the size of our convex phase retrieval relaxations.

6.8. Learning Sparse Penalties for Change-point Detection using Max Margin Interval Regression

Participants: Toby Hocking, Guillem Rigaill, Jean-Philippe Vert, Francis Bach [correspondent].

In segmentation models, the number of change-points is typically chosen using a penalized cost function. In [22] we propose to learn the penalty and its constants in databases of signals with weak change-point annotations. We propose a convex relaxation for the resulting interval regression problem, and solve it using accelerated proximal gradient methods. We show that this method achieves state-of-the-art change-point detection in a database of annotated DNA copy number profiles from neuroblastoma tumors.

6.9. Maximizing submodular functions using probabilistic graphical models

Participants: K. S. Sesh Kumar [correspondent], Francis Bach.

In [34] we consider the problem of maximizing submodular functions; while this problem is known to be NP-hard, several numerically efficient local search techniques with approximation guarantees are available. In this paper, we propose a novel convex relaxation which is based on the relationship between submodular functions, entropies and probabilistic graphical models. In a graphical model, the entropy of the joint distribution decomposes as a sum of marginal entropies of subsets of variables; moreover, for any distribution, the entropy of the closest distribution factorizing in the graphical model provides an bound on the entropy. For directed graphical models, this last property turns out to be a direct consequence of the submodularity of the entropy function, and allows the generalization of graphical-model-based upper bounds to any submodular functions. These upper bounds may then be jointly maximized with respect to a set, while minimized with respect to the graph, leading to a convex variational inference scheme for maximizing submodular functions, based on outer approximations of the marginal polytope and maximum likelihood bounded treewidth structures. By considering graphs of increasing treewidths, we may then explore the trade-off between computational complexity and tightness of the relaxation. We also present extensions to constrained problems and maximizing the difference of submodular functions, which include all possible set functions.

Optimizing submodular functions has been an active area of research with applications in graph-cut-based image segmentation [44], sensor placement [69], or document summarization [70]. A set function F is a function defined on the power set 2^V of a certain set V. It is submodular if and only if for all $A, B \subseteq V$, $F(A) + F(B) \geqslant F(A \cap B) + F(A \cup B)$. Equivalently, these functions also admit the diminishing returns property, i.e., the marginal cost of an element in the context of a smaller set is more than its cost in the context of a larger set. Classical examples of such functions are entropy, mutual information, cut functions, and covering functions—see further examples in [58], [38].

Submodular functions form an interesting class of discrete functions because minimizing a submodular function can be done in polynomial time [58], while maximization, although NP-hard, admits constant factor approximation algorithms [76]. In this paper, our ultimate goal is to provide the first (to the best of our knowledge) generic convex relaxation of submodular function maximization, with a hierarchy of complexities related to known combinatorial hierarchies such as the Sherali-Adams hierarchy [83]. Beyond the graphical model tools that we are going to develop, having convex relaxations may be interesting for several reasons:

(1) they can lead to better solutions, (2) they provide online bounds that may be used within branch-and-bound optimization and (3) they ease the use of such combinatorial optimization problems within structured prediction framework [91].

We make the following contributions:

- For any directed acyclic graph G and a submodular function F, we define a bound $F_G(A)$ and study its properties (monotonicity, tightness), which is specialized to decomposable graphs.
- We propose an algorithm to maximize submodular functions by maximizing the bound $F_G(A)$ with respect to A while minimizing with respect to the graph G, leading to a convex variational method based on outer approximation of the marginal polytope [93] and inner approximation of the hypertree polytope.
- We propose extensions to constrained problems and maximizing the difference of submodular functions, which include all possible set functions.
- We illustrate our results on small-scale experiments.

6.10. Reflection methods for user-friendly submodular optimization

Participants: Stefanie Jegelka, Suvrit Sra, Francis Bach [correspondent].

Recently, it has become evident that submodularity naturally captures widely occurring concepts in machine learning, signal processing and computer vision. Consequently, there is need for efficient optimization procedures for submodular functions, especially for minimization problems. While general submodular minimization is challenging, we propose in [15] a new method that exploits existing decomposability of submodular functions. In contrast to previous approaches, our method is neither approximate, nor impractical, nor does it need any cumbersome parameter tuning. Moreover, it is easy to implement and parallelize. A key component of our method is a formulation of the discrete submodular minimization problem as a continuous best approximation problem that is solved through a sequence of reflections, and its solution can be easily thresholded to obtain an optimal discrete solution. This method solves *both* the continuous and discrete formulations of the problem, and therefore has applications in learning, inference, and reconstruction. In our experiments, we illustrate the benefits of our method on two image segmentation tasks.

6.11. Convex Relaxations for Learning Bounded Treewidth Decomposable Graphs

Participants: K. S. Sesh Kumar [correspondent], Francis Bach.

In [24] we consider the problem of learning the structure of undirected graphical models with bounded treewidth, within the maximum likelihood framework. This is an NP-hard problem and most approaches consider local search techniques. In this paper, we pose it as a combinatorial optimization problem, which is then relaxed to a convex optimization problem that involves searching over the forest and hyperforest polytopes with special structures, independently. A supergradient method is used to solve the dual problem, with a run-time complexity of $O(k^3 n^{k+2} \log n)$ for each iteration, where n is the number of variables and k is a bound on the treewidth. We compare our approach to state-of-the-art methods on synthetic datasets and classical benchmarks, showing the gains of the novel convex approach.

Graphical models provide a versatile set of tools for probabilistic modeling of large collections of interdependent variables. They are defined by graphs that encode the conditional independences among the random variables, together with potential functions or conditional probability distributions that encode the specific local interactions leading to globally well-defined probability distributions [42], [93], [67].

In many domains such as computer vision, natural language processing or bioinformatics, the structure of the graph follows naturally from the constraints of the problem at hand. In other situations, it might be desirable to estimate this structure from a set of observations. It allows (a) a statistical fit of rich probability distributions that can be considered for further use, and (b) discovery of structural relationship between different variables. In the former case, distributions with tractable inference are often desirable, i.e., inference with run-time complexity does not scale exponentially in the number of variables in the model. The simplest constraint to ensure tractability is to impose tree-structured graphs [52]. However, these distributions are not rich enough, and following earlier work [73], [39], [75], [48], [59], [89], we consider models with *treewidth* bounded, not simply by one (i.e., trees), but by a small constant k.

Beyond the possibility of fitting tractable distributions (for which probabilistic inference has linear complexity in the number of variables), learning bounded-treewidth graphical models is key to design approximate inference algorithms for graphs with higher treewidth. Indeed, as shown by [82], [93], [68], approximating general distributions by tractable distributions is a common tool in variational inference. However, in practice, the complexity of variational distributions is often limited to trees (i.e., $k = 1$), since these are the only ones with exact polynomial-time structure learning algorithms. The convex relaxation we designed enables us to augment the applicability of variational inference, by allowing a finer trade-off between run-time complexity and approximation quality.

We make the following contributions:

- We provide a novel convex relaxation for learning bounded-treewidth decomposable graphical models from data in polynomial time. This is achieved by posing the problem as a combinatorial optimization problem, which is relaxed to a convex optimization problem that involves the graphic and hypergraphic matroids.
- We show how a supergradient ascent method may be used to solve the dual optimization problem, using greedy algorithms as inner loops on the two matroids. Each iteration has a run-time complexity of $O(k^3 n^{k+2} \log n)$, where n is the number of variables. We also show how to round the obtained fractional solution.
- We compare our approach to state-of-the-art methods on synthetic datasets and classical benchmarks showing the gains of the novel convex approach.

6.12. Large-Margin Metric Learning for Partitioning Problems

Participants: Rémi Lajugie [correspondent], Sylvain Arlot, Francis Bach.

In [31] we consider unsupervised partitioning problems, such as clustering, image segmentation, video segmentation and other change-point detection problems. We focus on partitioning problems based explicitly or implicitly on the minimization of Euclidean distortions, which include mean-based change-point detection, K-means, spectral clustering and normalized cuts. Our main goal is to learn a Mahalanobis metric for these unsupervised problems, leading to feature weighting and/or selection. This is done in a supervised way by assuming the availability of several potentially partially labelled datasets that share the same metric. We cast the metric learning problem as a large-margin structured prediction problem, with proper definition of regularizers and losses, leading to a convex optimization problem which can be solved efficiently with iterative techniques. We provide experiments where we show how learning the metric may significantly improve the partitioning performance in synthetic examples, bioinformatics, video segmentation and image segmentation problems.

Unsupervised partitioning problems are ubiquitous in machine learning and other data-oriented fields such as computer vision, bioinformatics or signal processing. They include (a) traditional *unsupervised clustering* problems, with the classical K-means algorithm, hierarchical linkage methods [61] and spectral clustering [80], (b) *unsupervised image segmentation* problems where two neighboring pixels are encouraged to be in the same cluster, with mean-shift techniques [51] or normalized cuts [84], and (c) *change-point detection* problems adapted to multivariate sequences (such as video) where segments are composed of contiguous elements, with typical window-based algorithms [54] and various methods looking for a change in the mean of the features (see, e.g., [49]).

All the algorithms mentioned above rely on a specific distance (or more generally a similarity measure) on the space of configurations. A good metric is crucial to the performance of these partitioning algorithms and its choice is heavily problem-dependent. While the choice of such a metric has been originally tackled manually (often by trial and error), recent work has considered learning such metric directly from data. Without any supervision, the problem is ill-posed and methods based on generative models may learn a metric or reduce dimensionality (see, e.g., [53]), but typically with no guarantees that they lead to better partitions. In this paper, we follow [41], [95], [40] and consider the goal of learning a metric for potentially several partitioning problems sharing the same metric, assuming that several fully or partially labelled partitioned datasets are available during the learning phase. While such labelled datasets are typically expensive to produce, there are several scenarios where these datasets have already been built, often for evaluation purposes. These occur in video segmentation tasks, image segmentation tasks as well as change-point detection tasks in bioinformatics (see [62]).

W consider partitioning problems based explicitly or implicitly on the minimization of Euclidean distortions, which include K-means, spectral clustering and normalized cuts, and mean-based change-point detection. We make the following contributions:

- We review and unify several partitioning algorithms, and cast them as the maximization of a linear function of a rescaled equivalence matrix, which can be solved by algorithms based on spectral relaxations or dynamic programming.
- Given fully labelled datasets, we cast the metric learning problem as a large-margin structured prediction problem, with proper definition of regularizers, losses and efficient loss-augmented inference.
- Given partially labelled datasets, we propose an algorithm, iterating between labeling the full datasets given a metric and learning a metric given the fully labelled datasets. We also consider extensions that allow changes in the full distribution of univariate time series (rather than changes only in the mean), with application to bioinformatics.
- We provide experiments where we show how learning the metric may significantly improve the partitioning performance in synthetic examples, video segmentation and image segmentation problems.

6.13. Comparison between multi-task and single-task oracle risks in kernel ridge regression

Participant: Matthieu Solnon [correspondent].

In [35] we study multi-task kernel ridge regression and try to understand when the multi-task procedure performs better than the single-task one, in terms of averaged quadratic risk. In order to do so, we compare the risks of the estimators with perfect calibration, the oracle risk. We are able to give explicit settings, favorable to the multi-task procedure, where the multi-task oracle performs better than the single-task one. In situations where the multi-task procedure is conjectured to perform badly, we also show the oracle does so. We then complete our study with simulated examples, where we can compare both oracle risks in more natural situations. A consequence of our result is that the multi-task ridge estimator has a lower risk than any single-task estimator, in favorable situations.

Increasing the sample size is the most common way to improve the performance of statistical estimators. In some cases (see, for instance, the experiments of [56] on customer data analysis or those of [63] on molecule binding problems), having access to some new data may be impossible, often due to experimental limitations. One way to circumvent those constraints is to use datasets from several related (and, hopefully, “similar”) problems, as if it gave additional (in some sense) observations on the initial problem. The statistical methods using this heuristic are called “multi-task” techniques, as opposed to “single-task” techniques, where every problem is treated one at a time. In this paper, we study kernel ridge regression in a multi-task framework and try to understand when multi-task can improve over single-task.

The first trace of a multi-task estimator can be found in the work of [88]. In this article, Charles Stein showed that the usual maximum-likelihood estimator of the mean of a Gaussian vector (of dimension larger than 3, every dimension representing here a task) is not admissible—that is, there exists another estimator that has a lower risk for every parameter. He showed the existence of an estimator that uniformly attains a lower quadratic risk by shrinking the estimators along the different dimensions towards an arbitrary point. An explicit form of such an estimator was given by [64], yielding the famous James-Stein estimator. This phenomenon, now known as the "Stein's paradox", was widely studied in the following years and the behaviour of this estimator was confirmed by empirical studies, in particular the one from [55]. This first example clearly shows the goals of the multi-task procedure: an advantage is gained by borrowing information from different tasks (here, by shrinking the estimators along the different dimensions towards a common point), the improvement being scored by the global (averaged) squared risk. Therefore, this procedure does not guarantee individual gains on every task, but a global improvement on the sum of those task-wise risks.

We consider $p \geq 2$ different regression tasks, a framework we refer to as "multi-task" regression, and where the performance of the estimators is measured by the fixed-design quadratic risk. Kernel ridge regression is a classical framework to work with and comes with a natural norm, which often has desirable properties (such as, for instance, links with regularity). This norm is also a natural "similarity measure" between the regression functions. [56] showed how to extend kernel ridge regression to a multi-task setting, by adding a regularization term that binds the regression functions along the different tasks together. One of the main questions that is asked is to assert whether the multi-task estimator has a lower risk than any single-task estimator. It was recently proved by [86] that a fully data-driven calibration of this procedure is possible, given some assumptions on the set of matrices used to regularize—which correspond to prior knowledge on the tasks. Under those assumptions, the estimator is showed to verify an *oracle inequality*, that is, its risk matches (up to constants) the best possible one, the *oracle risk*. Thus, it suffices to compare the oracle risks for the multi-task procedure and the single-task one to provide an answer to this question.

We study the oracle multi-task risk and compare it to the oracle single-task risk. We then find situations where the multi-task oracle is proved to have a lower risk than the single-task oracle. This allows us to better understand which situation favors the multi-task procedure and which does not. After having defined our model, we write down the risk of a general multi-task ridge estimator and see that it admits a convenient decomposition using two key elements: the mean of the tasks and the resulting variance. This decomposition allows us to optimize this risk and get a precise estimation of the oracle risk, in settings where the ridge estimator is known to be minimax optimal. We then explore several repartitions of the tasks that give the latter multi-task rates, study their single-task oracle risk and compare it to their respective multi-task rates. This allows us to discriminate several situations, depending whether the multi-task oracle either outperforms its single-task counterpart, underperforms it or whether both behave similarly. We also show that, in the cases favorable to the multi-task oracle detailed in the previous sections, the estimator proposed by [86] behaves accordingly and achieves a lower risk than the single-task oracle. We finally study settings where we can no longer explicitly study the oracle risk, by running simulations, and we show that the multi-task oracle continues to retain the same virtues and disadvantages as before.

6.14. Sharp analysis of low-rank kernel matrix approximations

Participant: Francis Bach [correspondent].

Kernel methods, such as the support vector machine or kernel ridge regression, are now widely used in many areas of science and engineering, such as computer vision or bioinformatics. However, kernel methods typically suffer from at least quadratic running-time complexity in the number of observations n, as this is the complexity of computing the kernel matrix. In large-scale settings where n may be large, this is usually not acceptable. In [7], we consider supervised learning problems within the positive-definite kernel framework, such as kernel ridge regression, kernel logistic regression or the support vector machine. Low-rank approximations of the kernel matrix are often considered as they allow the reduction of running time complexities to $O(p^2 n)$, where p is the rank of the approximation. The practicality of such methods thus depends on the required rank p. In this paper, we show that in the context of kernel ridge regression, for approximations based on a random subset of columns of the original kernel matrix, the rank p may be chosen to be linear in the *degrees of freedom* associated with the problem, a quantity which is classically used in the statistical analysis of such methods, and is often seen as the implicit number of parameters of non-parametric estimators. This result enables simple algorithms that have sub-quadratic running time complexity, but provably exhibit the same *predictive performance* than existing algorithms, for any given problem instance, and not only for worst-case situations.

6.15. fMRI encoding and decoding models

Participant: Fabian Pedregosa [correspondent].

In [20] we show that HRF estimation improves sensitivity of fMRI encoding and decoding models and propose a new approach for the estimation of Hemodynamic Response Functions from fMRI data. The model we propose is based on the linearity assumption behind the General Linear Model and can be computed using standard gradient-based solvers. We use the activation patterns computed by our model as input data for encoding and decoding studies and report performance improvement in both settings.

This work proves that significant improvements in recovery of brain activation patterns can be made by estimating the form of the Hemodynamic Response Function instead of using a canonical form for this response.

6.16. Structured Penalties for Log-linear Language Models

Participants: Anil Nelakanti [correspondent], Cédric Archambeau, Francis Bach, Guillaume Bouchard.

Language models can be formalized as log-linear regression models where the input features represent previously observed contexts up to a certain length m. The complexity of existing algorithms to learn the parameters by maximum likelihood scale linearly in nd, where n is the length of the training corpus and d is the number of observed features. In [19] we present a model that grows logarithmically in d, making it possible to efficiently leverage longer contexts. We account for the sequential structure of natural language using tree-structured penalized objectives to avoid overfitting and achieve better generalization.

Language models are crucial parts of advanced natural language processing pipelines, such as speech recognition [45], machine translation [47], or information retrieval [92]. When a sequence of symbols is observed, a language model predicts the probability of occurrence of the next symbol in the sequence. Models based on so-called back-off smoothing have shown good predictive power [60]. In particular, Kneser-Ney (KN) and its variants [66] are still achieving state-of-the-art results for more than a decade after they were originally proposed. Smoothing methods are in fact clever heuristics that require tuning parameters in an ad-hoc fashion. Hence, more principled ways of learning language models have been proposed based on maximum entropy [50] or conditional random fields [81], or by adopting a Bayesian approach [94].

We focus on penalized maximum likelihood estimation in log-linear models. In contrast to language models based on *unstructured* norms such as ℓ_2 (quadratic penalties) or ℓ_1 (absolute discounting), we use *tree-structured* norms [96], [65]. Structured penalties have been successfully applied to various NLP tasks, including chunking and named entity recognition [74], but not language modeling. Such penalties are particularly well-suited to this problem as they mimic the nested nature of word contexts. However, existing optimizing techniques are not scalable for large contexts m.

We show that structured tree norms provide an efficient framework for language modeling. Furthermore, we give the first algorithm for structured ℓ_∞ tree norms with a complexity nearly linear in the number of nodes. This leads to a memory-efficient *and* time-efficient learning algorithm for generalized linear language models.

6.17. Distributed Large-scale Natural Graph Factorization

Participants: Amr Ahmed, Nino Shervashidze [correspondent], Shravan Narayanamurthy, Vanja Josifovski, Alexander Smola.

Natural graphs, such as social networks, email graphs, or instant messaging patterns, have become pervasive through the internet. These graphs are massive, often containing hundreds of millions of nodes and billions of edges. While some theoretical models have been proposed to study such graphs, their analysis is still difficult due to the scale and nature of the data. We propose a framework for large-scale graph decomposition and inference. To resolve the scale, our framework in [6] is distributed so that the data are partitioned over a shared-nothing set of machines. We propose a novel factorization technique that relies on partitioning a graph so as to minimize the number of neighboring vertices rather than edges across partitions. Our decomposition is based on a streaming algorithm. It is network-aware as it adapts to the network topology of the underlying computational hardware. We use local copies of the variables and an efficient asynchronous communication protocol to synchronize the replicated values in order to perform most of the computation without having to incur the cost of network communication. On a graph of 200 million vertices and 10 billion edges, derived from an email communication network, our algorithm retains convergence properties while allowing for almost linear scalability in the number of computers.

6.18. Evaluating Speech Features with the Minimal-Pair ABX task

Participants: Thomas Schatz [correspondent], Vijayaditya Peddinti, Francis Bach, Aren Jansen, Hynek Hermansky, Emmanuel Dupoux.

In [23] we introduce a new framework for the evaluation of speech representations in zero-resource settings, that extends and complements previous work by Carlin, Jansen and Hermansky [46]. In particular, we replace their Same/Different discrimination task by several Minimal-Pair ABX (MP-ABX) tasks. We explain the analytical advantages of this new framework and apply it to decompose the standard signal processing pipelines for computing PLP and MFC coefficients. This method enables us to confirm and quantify a variety of well-known and not-so-well-known results in a single framework.

Speech recognition technology crucially rests on adequate speech features for encoding input data. Several such features have been proposed and studied (MFCCs, PLPs, etc), but they are often evaluated indirectly using complex tasks like phone classification or word identification. Such an evaluation technique suffers from several limitations. First, it requires a large enough annotated corpus in order to train the classifier/recognizer. Such a resource may not be available in all languages or dialects (the so-called “zero or limited resource” setting). Second, supervised classifiers may be too powerful and may compensate for potential defects of speech features (for instance noisy/unreliable channels). However, such defects are problematic in unsupervised learning techniques. Finally, the particular statistical assumptions of the classifier (linear, Gaussian, etc.) may not be suited for specific speech features (for instance sparse neural codes as in Hermansky [85]). It is therefore important to replace these complex evaluation schemes by simpler ones which tap more directly the properties of the speech features.

We extend and complement the framework proposed by Carlin, Jansen and Hermansky [46] for the evaluation of speech features in zero resource settings. This framework uses a Same-Different word discrimination task that does not depend on phonetically labelled data, nor on training a classifier. It assumes a speech corpus segmented into words, and derives a word-by-word acoustic distance matrix computed by comparing every word with every other one using Dynamic Time Warping (DTW). Carlin et al. then compute an average precision score which is used to evaluate speech features (the higher average precision, the better the features).

We explore an extension of this technique through Minimal-Pair ABX tasks (MP-ABX tasks) tested on a phonetically balanced corpus [57]. This improves the interpretability of the Carlin et al evaluation results in three different ways. First, the Same/Different task requires the computation of a ROC curve in order to derive average precision. In contrast, the ABX task is a discrimination task used in psychophysics (see [72], chapter 9) which allows for the direct computation of an error rate or a d’ measure that are easier to interpret than average precision [46] and involve no assumption about ROC curves. Second, the Same/Different task compares *sets of words*, and as a result is influenced by the mix of similar versus distinct words or short versus long words in the corpus. The ABX task, in contrast, is computed on *word pairs*, and therefore enables to make linguistically precise comparisons, as in word *minimal pairs*, i.e. words differing by only one phoneme. Variants of the task enable to study phoneme discrimination across talkers and/or phonetic contexts, as well as talker discrimination across phonemes. Because it is more controlled and provides a parameter and model-free metric, the MP-ABX error rate also enables to compare performance across databases or across languages. Third, we compute bootstrap-based estimates of the variability of our performance measures, which allows us to derive confidence intervals for the error rates and tests of the significance of the difference between the error rates obtained with different representations.

6.19. Hidden Markov Tree Models for Semantic Class Induction

Participants: Edouard Grave [correspondent], Guillaume Obozinski, Francis Bach.

In [13] we propose a new method for semantic class induction. First, we introduce a generative model of sentences, based on dependency trees and which takes into account homonymy. Our model can thus be seen as a generalization of Brown clustering. Second, we describe an efficient algorithm to perform inference and learning in this model. Third, we apply our proposed method on two large datasets (10^8 tokens, 10^5 words types), and demonstrate that classes induced by our algorithm improve performance over Brown clustering on the task of semi-supervised supersense tagging and named entity recognition.

Most competitive learning methods for computational linguistics are supervised, and thus require labeled examples, which are expensive to obtain. Moreover, those techniques suffer from data scarcity: many words only appear a small number of time, or even not at all, in the training data. It thus helps a lot to first learn word clusters on a large amount of unlabeled data, which are cheap to obtain, and then to use this clusters as features for the supervised task. This scheme has proven to be effective for various tasks such as named entity recognition, syntactic chunking or syntactic dependency parsing. It was also successfully applied for transfer learning of multilingual structure.

The most commonly used clustering method for semi-supervised learning is known as Brown clustering. While still being one of the most efficient word representation method, Brown clustering has two limitations we want to address in this work. First, since it is a hard clustering method, homonymy is ignored. Second, it does not take into account syntactic relations between words, which seems crucial to induce semantic classes. Our goal is thus to propose a method for semantic class induction which takes into account both syntax and homonymy, and then to study their effects on semantic class learning.

We start by introducing a new unsupervised method for semantic classes induction. This is achieved by defining a generative model of sentences with latent variables, which aims at capturing semantic roles of words. We require our method to be scalable, in order to learn models on large datasets containing tens of millions of sentences. More precisely, we make the following contributions:

- We introduce a generative model of sentences, based on dependency trees, which can be seen as a generalization of Brown clustering,
- We describe a fast approximate inference algorithm, based on message passing and online EM for scaling to large datasets. It allowed us to learn models with 512 latent states on a dataset with hundreds of millions of tokens in less than two days on a single core,
- We learn models on two datasets, Wikipedia articles about musicians and the NYT corpus, and evaluate them on two semi-supervised tasks, namely supersense tagging and named entity recognition.

6.20. Domain Adaptation for Sequence Labeling using Hidden Markov Models

Participants: Edouard Grave [correspondent], Guillaume Obozinski, Francis Bach.

Most natural language processing systems based on machine learning are not robust to domain shift. For example, a state-of-the-art syntactic dependency parser trained on Wall Street Journal sentences has an absolute drop in performance of more than ten points when tested on textual data from the Web. An efficient solution to make these methods more robust to domain shift is to first learn a word representation using large amounts of unlabeled data from both domains, and then use this representation as features in a supervised learning algorithm. In this paper, we propose to use hidden Markov models to learn word representations for part-of-speech tagging. In particular, we study the influence of using data from the source, the target or both domains to learn the representation and the different ways to represent words using an HMM.

Nowadays, most natural language processing systems are based on supervised machine learning. Despite the great successes obtained by those techniques, they unfortunately still suffer from important limitations. One of them is their sensitivity to domain shift: for example, a state-of-the-art part-of-speech tagger trained on the Wall Street Journal section of the Penn treebank achieves an accuracy of 97% when tested on sentences from the Wall Street Journal, but only 90% when tested on textual data from the Web. This drop in performance can also be observed for other tasks such as syntactic parsing or named entity recognition.

One of the explanations for this drop in performance is the big lexical difference that exists accross domains. This results in a lot of out-of-vocabulary words (OOV) in the test data, *i.e.*, words of the test data that were not observed in the training set. For example, more than 25% of the tokens of the test data from the Web corpus are unobserved in the training data from the WSJ. By comparison, only 11.5% of the tokens of the test data from the WSJ are unobserved in the training data from the WSJ. Part-of-speech taggers make most of their errors on those out-of-vocabulary words.

Labeling enough data to obtain a high accuracy for each new domain is not a viable solution. Indeed, it is expensive to label data for natural language processing, because it requires expert knowledge in linguistics. Thus, there is an important need for transfer learning, and more precisely for domain adaptation, in computational linguistics. A common solution consists in using large quantities of unlabeled data, from both source and target domains, in order to learn a good word representation. This representation is then used as features to train a supervised classifier that is more robust to domain shift. Depending on how much data from the source and the target domains are used, this method can be viewed as performing semi-supervised learning or domain adaptation. The goal is to reduce the impact of out-of-vocabulary words on performance. This scheme was first proposed to reduce data sparsity for named entity recognition, before being applied to domain adaptation for part-of-speech tagging or syntactic parsing.

Hidden Markov models have already been considered in previous work to learn word representations for domain adaptation or semi-supervised learning. Our contributions in [25] are mostly experimental: we compare different word representations that can be obtained from an HMM and study the effect of training the unsupervised HMM on source, target or both domains. While previous work mostly use Viterbi decoding to obtain word representations from an HMM, we empirically show that posterior distributions over latent classes give better results.

6.21. Simple Greedy Matching for Aligning Large Knowledge Bases

Participant: Simon Lacoste-Julien [correspondent].

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

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. Here, 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 that leverages both the structural information from the relationship graph and flexible similarity measures between entity properties in a greedy local search, which makes it scalable. Despite its greedy nature, our experiments in [17] indicate that SiGMa can efficiently match some of the world's largest knowledge bases with high accuracy. We provide additional experiments on benchmark datasets which demonstrate that SiGMa can outperform state-of-the-art approaches both in accuracy and efficiency.

SIMPAF Project-Team

6. New Results

6.1. Quantitative homogenization theory

In collaboration with S. Neukamm and F. Otto, A. Gloria developed in [46] and [45] a quantitative approach of the stochastic homogenization of discrete elliptic equations. There are two main achievements. In [46] we developed a general theory which quantifies optimally in time the decay of the non-constant coefficients semi-group associated with discrete random diffusion coefficients satisfying a spectral gap assumption (namely, the environment seen from the particle). Combined with spectral theory this allowed us to make a sharp numerical analysis of the popular periodization method to approximate homogenized coefficients. In [45], we obtained a quantitative two-scale expansion result, and essentially proved that the difference between the solution of a (discrete) elliptic equation with random coefficients on the torus and the first two terms of the two-scale expansion scales as in the periodic case (except in dimension 2, for which there is a logarithmic correction).

6.2. Corrosion

The Diffusion Poisson Coupled Model [32] is a model of iron based alloy in a nuclear waste repository. It describes the growth of an oxide layer in this framework. The system is made of a Poisson equation on the electrostatic potential and convection-diffusion equations on the densities f charge carriers (electrons, ferric cations and oxygen vacancies), supplemented with coupled Robin boundary conditions. The DPCM model also takes into account the growth of the oxide host lattice and its dissolution, leading to moving boundary equations. In [44], C. Chainais-Hillairet and I. Lacroix-Violet consider a simplified version of this model, where only two charge carriers are taken into account and where there is no evolution of the layer thickness. They prove the existence of a steady-state solution to this model. More recently, C. Chainais-Hillairet and I. Lacroix-Violet have also obtained an existence result for the time-dependent simplified model. This result is submitted for publication [47].

P.-L. Colin, C.Chainais-Hillairet and I. Lacroix-Violet have recently performed the numerical analysis of the numerical scheme presented in [31]. The scheme is a Euler implicit in time scheme with Scharfetter-Gummel approximation of the convection-diffusion fluxes. They prove existence of a solution to the scheme, a priori estimates satisfied by the solution and convergence of the numerical scheme to a weak solution of the corrosion model.

Numerical experiments done for the simulation of the full DPCM model with moving boundaries shows the convergence in time towards a pseudo-steady-state. T. Gallouët has proposed a new scheme in order to compute directly this pseudo-steady-state. This scheme has been implemented in the code CALIPSO (ANDRA). Validation is in progress, as the numerical analysis of the scheme.

6.3. New results on finite volume schemes

In [5], C. Chainais-Hillairet, S. Krell and A. Mouton develop Discrete Duality Finite Volume methods for the finite volume approximation of a system describing miscible displacement in porous media (Peaceman model). They establish relevant a priori estimates satisfied by the numerical solution and prove existence and uniqueness of the solution to the scheme. They show the efficiency of the schemes through numerical experiments. Recently, they also proved the convergence of the DDFV scheme for the Peaceman model. This work will be soon submitted for publication.

In [35], M. Bessemoulin-Chatard, C. Chainais-Hillairet and F. Filbet prove several discrete Gagliardo-Nirenberg-Sobolev and Poincaré-Sobolev inequalities for some approximations with arbitrary boundary values on finite volume meshes. The keypoint of their approach is to use the continuous embedding of the space $BV(\Omega)$ into $L^{N/(N-1)}(\Omega)$ for a Lipschitz domain $\Omega \subset \mathbb{R}^N$, with $N \geq 2$. Finally, they give several applications to discrete duality finite volume (DDFV) schemes which are used for the approximation of nonlinear and on isotropic elliptic and parabolic problems.

In [22], M. Bessemoulin-Chatard, C. Chainais-Hillairet and M.-H. Vignal consider the numerical approximation of the classical time-dependent drift-diffusion system near quasi-neutrality by a fully implicit in time and finite volume in space scheme, where the convection-diffusion fluxes are approximated by Scharfetter-Gummel fluxes. They establish that all the a priori estimates needed to prove the convergence of the scheme does not depend on the Debye length λ. This proves that the scheme is asymptotic preserving in the quasi-neutral limit $\lambda \to 0$.

In [24], C. Chainais-Hillairet, A. Jüngel and S. Schuchnigg prove the time decay of fully discrete finite-volume approximations of porous-medium and fast-diffusion equations with Neumann or periodic boundary conditions in the entropy sense. The algebraic or exponential decay rates are computed explicitly. In particular, the numerical scheme dissipates all zeroth-order entropies which are dissipated by the continuous equation. The proofs are based on novel continuous and discrete generalized Beckner inequalities.

6.4. New results in numerical fluid dynamics

In the case of compressible models, as the Euler equations, a careful analysis of sharp and practical stability conditions to ensure the positivity of both density and pressure variables was performed[4]. We are also concerned with the numerical simulation of certain multi-fluids flows, which in particular arises in the modeling of powder/snow avalanches. The hybrid scheme works on unstructured meshes and can be advantageously coupled to mesh refinements strategies in order to follow fronts of high density variation [42]. In particular, we investigate the influence of the characteristics Froude number, Schmidt number and Reynolds number on the front progression. In the context of the PhD thesis of Meriem Ezzoug (University of Monastir, Tunisia), co-advised by C. Calgaro and E. Zahrouni (University of Monastir, Tunisia), we investigate theoretically and numerically the influence of a specific stress tensor, introduced for the first time by Korteweg, in some diffuse interface models which allow to describe some phase transition phenomena, such as surface tension force formulation for multiphase fluid flows. In order to answer these questions, we have developed respectively a Fortran code, a C++ code (NS2DDV-C++, see the softwares section) and a MATLAB code (NS2DDV-M, see the softwares section).

6.5. New results on a posteriori estimates

Some residual-type a posteriori error estimators were developed in the context of magnetostatic and magnetodynamic Maxwell equations, given in their potential and harmonic formulations. Here, the task was to find a relevant decomposition of the error in order to obtain the reliability of the estimator, with the use of ad-hoc interpolations. This work was realized in collaboration with the L2EP Laboratory (Laboratoire d'Electrotechnique et d'Electronique de Puissance de Lille, Lille 1 University), and gave rise to several contributions [7], [18], [19], [20], [21], obtained in the context of the Ph-D thesis and of the Post-doc position of Zuqi Tang. Then, other results about a posteriori error estimators were obtained in other contexts [6], [8]. Recently, we started working on space/time error estimators for finite element methods, arising in the context of low-frequency Maxwell equations (PhD of R. Tittarelli, CIFRE EDF R&D, see [25].

6.6. New results in control in fluid mechanics

Recently, we studied more particularly passive control techniques using porous media for incompressible aerodynamics on several bodies, with the use of the penalisation method [3].

SIROCCO Project-Team

6. New Results

6.1. Analysis and modeling for compact representation and navigation

3D modelling, multi-view plus depth videos, Layered depth images (LDI), 2D and 3D meshes, epitomes, image-based rendering, inpainting, view synthesis

6.1.1. Salient object detection

Participants: Olivier Le Meur, Zhi Liu.

Salient object detection consists in extracting in an automatic manner the most interesting object in an image or video sequence. From an input image, an object, with well-defined boundaries, is detected based on its saliency. This subject knows an renewed interest these last years. A number of datasets serving as ground truth has been released and can be used to benchmark methods.

In 2013, a new method to detect salient objects has been proposed [32], [18]. The principle relies upon low-level visual features and super-pixel segmentation. First, the original image is simplified by performing super-pixel segmentation and adaptive color quantization. On the basis of super-pixel representation, inter-super-pixel similarity measures are then calculated based on difference of histograms and spatial distance between each pair of super-pixels. For each super-pixel, its global contrast measure and spatial sparsity measure are evaluated, and refined with the integration of inter super-pixel similarity measures to finally generate the super-pixel-level saliency map. Experimental results on a dataset containing 1,000 test images with ground truths demonstrate that the proposed saliency model outperforms state-of-the-art saliency models. Figure 1 illustrates some results.

6.1.2. Image Memorability

Participant: Olivier Le Meur.

This work has been carried out in collaboration with Mattei Mancas (researcher of the University of Mons) during his visit of the team. The image memorability consists in the faculty of an image to be recalled after a period of time. Recently, the memorability of an image database was measured and some factors responsible for this memorability were highlighted. In [34] we proposed to improve an existing method by using attention-based visual features. To determine whether the visual attention plays a role in the memorability mechanism, eye tracking experiment has been performed by using a set of images of different memorability scores. Two important results have been observed. First the fixation duration is longer for the most memorable images (especially for the very first fixations) which shows a higher cognitive activity for memorable images. Second the observers congruency (agreement between observers) is significantly higher for the most memorable images. This shows that when there are areas with high attraction on all viewers, this induces higher memorability.

Following these first two observations, attention-based visual features were used to predict image memorability scores. A new set of features was then defined and used to train a model. Compared to an existing approach, we improve on the quality of the prediction of 2% while reducing the number of parameters by 14%. More specifically we replace the 512 features related to the GIST by 17 features which are directly related to visual attention.

6.1.3. Models for 3D video quality assessment

Participants: Darya Khaustova, Olivier Le Meur.

This work is carried out in collaboration with Orange labs. The goal is to design objective metrics for quality assessment of 3D video content, by establishing links between human visual perception (visual comfort) and video parameters such as quality and depth quantity, and between visual comfort and visual attention. The goal is also to study the differences in 2D visual attention in comparison with 3D visual attention.

../../../../projets/sirocco/IMG/Fig_Salient_Object_Segmentation.png

Figure 1. Illustration of the proposed approach: first row: original image; second row: saliency map; third row: extraction of the salient object.

Several subjective experiments have been carried out in order to study visual attention in different viewing conditions. The goal of the first experiment, involving 135 observers, was to study visual attention in three different conditions (2D, 3D comfortable and 3D uncomfortable), to eventually establish whether depth influences visual attention and whether there is a link between comfort and visual attention. The use of an eye-tracker allowed to record and to track observer's gaze. By analyzing the results, we found out that visual strategy to observe 2D images and 3D images with uncrossed disparity is very similar; there was no significant influence of discomfort on visual attention.

The second question which has then been addressed is how visual attention is influenced by objects with crossed disparity. A second test has been designed to answer this question, involving 51 observers. Considering scenes with crossed disparity it was revealed that objects located in front of the display plane are the most salient, even if observers experience discomfort. In the third experiment, we extended the study using scenes with crossed and uncrossed disparities. We verified the hypothesis that texture and contrast are more influential in guiding our gaze than the amount of depth. The features influencing the saliency of the objects in stereoscopic conditions were also evaluated with low-level visual stimuli. It was discovered that texture is the most salient feature in comparison to depth. Crossed disparity significantly influences the process of selecting the objects, while uncrossed disparity is less important, the process of selection being in this latter case similar to 2D conditions.

6.1.4. Epitome-based video representation

Participants: Martin Alain, Christine Guillemot.

This work is carried out in collaboration with Technicolor (D. Thoreau, Ph. Guillotel) and aims at studying novel spatio-temporal representations for videos based on epitomes. An epitome is a condensed representation of an image (or a video) signal containing the essence of the textural properties of this image. Different forms of epitomes have been proposed in the literature, such as a patch-based probability model learned either from still image patches or from space-time texture cubes taken from the input video. These probability models together with appropriate inference algorithms, are useful for content analysis inpainting or super-resolution. Another family of approaches makes use of computer vision techniques, like the KLT tracking algorithm, in order to recover self similarities within and across images. In parallel, another type of approach consists in extracting epitome-like signatures from images using sparse coding and dictionary learning.

We have in the past (in the context of the PhD thesis of S. Cherigui) developed a method for constructing epitomes for representing still images. The algorithm tracks self-similarities within the image using a block matching (BM) algorithm. The epitome is constructed from disjoint pieces of texture ("epitome charts") taken from the original image and a transform map which contains translational parameters (see Fig.2 . Those parameters keep track of the correspondences between each block of the input image and a block of the epitome. An Intra image compression scheme based on the epitome has been developed showing significant rate savings on some images, including the rate cost of the epitome texture and of the transform map. The entire image can be reconstructed from the epitome texture with the help of the transform map. The method is currently being extended to construct epitome representations of video segments rather than simple images. Such spatio-temporal epitome should pave the way for novel video coding architectures and open perspectives for other video processing problems which we have started to address such as denoising and super-resolution.

6.2. Rendering, inpainting and super-resolution

image-based rendering, inpainting, view synthesis, super-resolution

6.2.1. Image and video inpainting

Participants: Mounira Ebdelli, Christine Guillemot, Olivier Le Meur.

../../../../projets/sirocco/IMG/foreman/cif_352x288_0.rgb.png

../../../../projets/sirocco/IMG/EpitomeForeman_

Figure 2. Original image and corresponding epitome.

Image (and video) inpainting refers to the process of restoring missing or damaged areas in an image (or a video). This field of research has been very active over the past years, boosted by numerous applications: restoring images from scratches or text overlays, loss concealment in a context of impaired image transmission, object removal in a context of editing, disocclusion in image-based rendering of viewpoints different from those captured by the cameras. Inpainting is an ill-posed inverse problem: given observations, or known samples in a spatial (or spatio-tempoal) neighborhood, the goal is to estimate unknown samples of the region to be filled in. Many methods already exist for image inpainting, either based on PDE (Partial Derivative Equation)-based diffusion schemes, either using sparse or low rank priors or following texture synthesis principles exploiting statistical or self-similarity priors.

Novel methods have been developed investigating two complementary directions first for image inpainting. The first direction which has been explored is the estimation of the unknown pixel with different neighbor embedding methods, i.e. Locally Linear embedding (LLE), LLE with a low-dimensional neigborhood representation (LLE-LDNR), Non-Negative Matrix Factorization (NMF) with various solvers [16]. The second method developed uses a two-steps hierarchical (or coarse to fine) approach to reduce the execution time [17]. In this hierarchical approach, a low resolution version of the input image is first inpainted, this first step being followed by a second one which recovers the high frequency details of the inpainted regions, using a single-image super-resolution method. To be less sensitive to the parameters setting of the inpainting, the low-resolution input picture is inpainted several times with different settings. Results are then efficiently combined with a loopy belief propagation. Experimental results in a context of image editing, texture synthesis and 3D view synthesis demonstrate the effectiveness of the proposed method.

The problem of video inpainting has also been considered. A first video inpainting algorithm has been developed in 2012, using a spatio-temporal examplar-based method. The algorithm proceeds in three steps. The first one inpaints missing pixels in moving objects using motion information. Then the static background is inpainted exploiting similarity between neighboring frames. The last step fills in the remaining holes in the current frame using spatial inpainting. This approach works well with static cameras but not so well when the video has been captured by free-moving cameras.

In 2013, we have therefore addressed the problem of video inpainting with free-moving cameras. The algorithm developed first compensates the camera motion between the current frame and its neighboring frames in a sliding window, using a new region-based homography computation which better respects the geometry of the scene compared to state-of-the-art methods. The source frame is first segmented into regions in order to find homogeneous regions. Then, the homography for mapping each region into the target frame is estimated. The overlapping of all aligned regions forms the registration of the source frame into the target one. Once the neighboring frames have been aligned, they form a stack of images from which the best candidate pixels are searched in order to replace the missing ones. The best candidate pixel is found by minimizing a cost function which combines two energy terms. One energy term, called the data term, captures how stationary is the background information after registration, hence enforcing temporal coherency. The second term aims at favoring spatial consistency and preventing incoherent seams, by computing the energy of the difference between each candidate pixel and its 4-neighboring pixels in the missing region. The minimization of the energy term is performed globally using Markov Random Fields and graph cuts. The proposed approach, although less complex than state-of-the-art methods, provides more natural results (see Fig.3).

Figure 3. Mask of the image to be inpainted; Results with the proposed video inpainting algorithm.

6.2.2. Image priors for inpainting

Participants: Raul Martinez Noriega, Aline Roumy.

Image inpainting is an ill-posed inverse problem which has no well-defined unique solution. To make this problem more "well-defined" it is necessay to introduce image priors. We consider here the problem of extracting such priors to help restoring the connection of long edges across the missing region. The prior is defined as a binary image that contains the locations of salient edge points located at the boundary of the missing region as well as the linear edges that join these points across the missing region. A method has been developed to extract such priors. It first detect edges which are then successively pruned in order to keep only informative edges, i.e., which have coherent gradients and are either part of a salient structure, or at the border between two different textures. Edges which are quasi-perpendicular to the boundary of the missing region are finally retained. Directions of the retained edges are computed and pairs of edges with similar directions are then connected with straight lines. These lines are used to segment the image into different regions and to define the processing order of the patches to be inpainted. Only patches from the known part and belonging to the same region as the input patch are used. This avoids bringing details of one texture into another one, as well as the unconnected edge problem [35].

6.2.3. Image and video super-resolution

Participants: Marco Bevilacqua, Christine Guillemot, Aline Roumy.

Super-resolution (SR) refers to the problem of creating a high-resolution (HR) image, given one or multiple low-resolution (LR) images as input. The SR process aims at adding to the LR input(s) new plausible high-frequency details, to a greater extent than traditional interpolation methods (see, for example, Fig. 4 for a comparison between bicubic interpolation and SR). We mostly focused on the single-image problem, where only a single LR image is available.

We have adopted the example-based framework, where the relation between the LR and HR image spaces is modeled with the help of pairs of small "examples", i.e. texture patches. Each example pair consists of a LR patch and its HR version that also includes high-frequency details; the pairs of patches form a dictionary of patches. For each patch of the LR input image, one or several similar patches are found in the dictionary, by performing a nearest neighbor search. The corresponding HR patches in the dictionary are then combined to form a HR output patch; and finally all the reconstructed HR patches are re-assembled to build the super-resolved image.

In this procedure, one important aspect is how the dictionary of patches is built. At this regard, two choices are possible: an external dictionary, formed by sampling HR and LR patches from external training images; and an internal dictionary, where the LR/HR patch correspondences are learnt by putting in relation directly the input image and scaled versions of it. The advantage of having an external dictionary is that it is built in advance: this leads to a reduction of the computational time, whereas in the internal case the dictionary is generated online at each run of the algorithm. However, external dictionaries have a considerable drawback: they are fixed and so non-adapted to the input image. To be able to satisfactorily process any input image, we need then to include in the dictionary a large variety of patch correspondences, leading to a high computational time.

To overcome this problem, in [23] we proposed a novel method to build a compact external dictionary. The method consists in first jointly clustering LR and HR patches. The aim of this procedure, which we called JKC (Jointly K-means Clustering), is to prune the dictionary of the "bad" pairs of patches, i.e. those ones for which the cluster assignments of the related LR and HR patches do not correspond. Once the dictionary is clustered, it is summarized, by sampling some prototype patches, and applying on them simple geometrical transformations, in order to enrich the dictionary. The so constructed compact dictionary is shown to give equivalent or even better performance than the initial large dictionary with any input image.

The dictionary construction method described in [23] has been used as a basis for designing a full single-image SR algorithm. The new algorithm, presented in [25], follows the traditional scheme of example-based SR with an external dictionary, where a new way to generate the training patches is introduced. Given a HR training image H, the corresponding LR image L is generated; but instead of directly sampling patches from H and L, as usually done, the training images are further processed. An enhanced interpolation of L, using an iterated back projection, is used as a source of LR patches, and a high-frequency residual image, given by the difference between H and the interpolated LR image, is used for extracting HR patches. The JKC procedure is then applied to get the final compact dictionary. A special example-based SR algorithm has been designed, where the final HR output patches are constructed by combining selected HR residual patches from the dictionary with nonnegative weights. In the context of this study, we have also introduced a novel non-negative dictionary learning method [24]. The proposed method consists of two steps which are alternatively iterated: a sparse coding and a dictionary update stage. As for the dictionary update, an original method has been proposed, which we called K-WEB, as it involves the computation of k WEighted Barycenters.

Besides SR for still images, a preliminary work on video sequences has been also conducted [26]. In particular, we have considered the case of a LR video sequence with periodic high-resolution (HR) key frames. Given this scenario, a specific SR procedure has been designed to upscale each intermediate frame, by using the internal dictionary constructed from the two neighbor key frames.

../../../../projets/sirocco/IMG/comp_bic_sr.png

Figure 4. Comparison between bicubic interpolation and SR when upscaling the same LR image (by a factor of 3).

6.3. Representation and compression of large volumes of visual data

Sparse representations, data dimensionality reduction, compression, scalability, perceptual coding, rate-distortion theory

6.3.1. Multi-view plus depth video compression

Participants: Christine Guillemot, Laurent Guillo.

Multi-view plus depth video content represent very large volumes of input data wich need to be compressed for storage and tranmission to the rendering device. The huge amount of data contained in multi-view sequences indeed motivates the design of efficient representation and compression algorithms. The team has worked on motion vector prediction in the context of HEVC-compatible Multi-view plus depth (MVD) video compression. The HEVC compatible MVD compression solution implements a 6 candidate vector list for merge and skip modes. When a merge or a skip mode is selected, a merge index is written in the bitstream. This index is first binarized using a unary code, then encoded with the CABAC. A CABAC context is dedicated to the first bin of the unary coded index while the remaining bins are considered as equiprobable. This strategy is efficient as long as the candidate list is ordered by decreasing index occurrence probability. We have improved the construction of the candidate list by proposing two new candidates derived from disparity motion vectors in order to exploit inter-view correlation. This work has led to a joint proposal with Qualcomm and Mediatek which has been adopted in the HEVC-3DV standard in July 2013.

6.3.2. Spatio-temporal video prediction with neighbor embedding

Participants: Martin Alain, Christine Guillemot.

The problem of texture prediction can be regarded as a problem of texture synthesis. Given observations, or known samples in a spatial neighborhood, the goal is to estimate unknown samples of the block to be predicted. We have in 2012 developed texture prediction methods as well as inpainting algorithms using sparse representation as with learned dictionaries [19], or using neighbor embedding techniques [11], [30]. The methods which we have more particularly considered are Locally Linear Embedding (LLE), LLE with Low-dimensional neighborhood representation (LDNR), and Non-negative Matrix Factorization (NMF) using various solvers. In 2013, we have addressed the problem of temporal prediction for inter frame coding of video sequences using locally linear embedding (LLE). LLE-based prediction computes the predictor as a linear combination of K nearest neighbors (K-NN) searched within one or several reference frames. We have explored different K-NN search strategies in the context of temporal prediction, leading to several temporal predictor variants using or not motion information [22]. A parallel was also drawn between such multi-patch based prediction and the adaptive interpolation filtering (AIF) method. The LLE-based inter prediction techniques, when used as extra modes for inter prediction in an H.264 codec, are shown to bring significant Rate-Distortion (RD) performance gains compared to H.264 (up to 21.76 $\%$ bit-rate saving) and with respect to the use of AIF.

6.3.3. Dictionary learning for sparse coding of satellite images

Participants: Jeremy Aghaei Mazaheri, Christine Guillemot, Claude Labit.

In the context of the national partnership Inria-Astrium, we explore novel methods to encode images captured by a geostationary satellite. These pictures have to be compressed on-board before being sent to earth. Each picture has a high resolution, therefore the rate without compression is very high (about 70 Gbits/sec). The goal is to achieve a rate after compression of 600 Mbits/sec, i.e., a compression ratio higher than 100. On earth, the pictures are decompressed with a high reconstruction quality and visualized by photo-interpreters. The goal of the study is to design novel transforms based on sparse representations and learned dictionnaries for satellite images.

Sparse representation of a signal consists in representing a signal $y \in \Re^n$ as a linear combination of columns, known as atoms, from a dictionary matrix. The dictionary $D \in \Re^{n \times K}$ is generally overcomplete and contains K atoms. The approximation of the signal can thus be written $y \approx Dx$ and is sparse because a small number of atoms of D are used in the representation, meaning that the vector x has only a few non-zero coefficients. Sparsity of the representation depends on how the dictionary is representative of the data at hand, hence the need to learn appropriate dictionaries.

We have developed methods for learning adaptive tree-structured dictionaries, called Tree K-SVD [20]. Each dictionary in the structure is learned on a subset of residuals from the previous level, with the K-SVD algorithm. The tree structure offers better rate-distortion performance than a "flat" dictionary learned with K-SVD, especially when only a few atoms are selected among the first levels of the tree. The tree-structured dictionary allows efficient coding of the indices of the selected atoms. We recently developped a new sparse coding method adapted to this tree-structure to improve the results [20]. The tree-structured dictionary has been further improved by studying different branch pruning strategies. The use of these dictionaries in an HEVC-based intra coder is under study. The dictionaries are also considered for scene classification and for detecting the MTF (Modulation Transfer Function) of the optical capturing system.

6.3.4. HDR video compression

Participants: Christine Guillemot, Mikael Le Pendu.

High Dynamic Range (HDR) images contain more intensity levels than traditional image formats. Instead of 8 or 10 bit integers, floating point values are generally used to represent the pixel data. Floating point video formats are widely used in the visual effects industry. Moreover, the development of a new standardized workflow ACES intends to generalize the use of such formats to the whole cinema production pipeline. The increasing use of floating point representations, however, comes with a technical issue concerning the storage space required for those videos with higher precision than the current 8 or 10 bit standards.

In collaboration with Technicolor (D. Thoreau), we worked on floating point video compression. Different approaches exist in the literature. Several methods consists in compressing directly the floating point data using its internal representation (i.e. sign, exponent and mantissa bits). These methods are generally limited to lossless compression schemes. Another type of approach makes use of the existing compression standards such as H264/AVC or HEVC to encode a floating point sequence of images previously converted to lower bit depth integers. In this approach, the conversion is designed to be reversible with minimal loss. However the converted integer images are not intended for being displayed directly. Finally a last family of approach aims at keeping backward compatibility with an existing compression standard. The original image sequence is first tone mapped and encoded to obtain a low dynamic range (LDR) version that can be visualized on a standard LDR display. In parallel, a residual information needed to reconstruct the HDR image from the LDR version is also encoded.

In our study, a floating point to integer conversion method was developed to be applied before HEVC compression. The original floating point RGB values are converted to high bit depth integers with an approximate logarithmic encoding that is reversible without loss. The RGB values are then converted to a YUV color space. The bit depth must also be reduced to be supported by the compression standard. This bit depth reduction is performed adaptively depending on the minimum and maximum values (i.e. darkest and brightest points respectively) which characterize the real dynamic of the data. In the best case, the difference between the extreme values is sufficiently low to perform this operation without loss.

Three variants of the method have been compared. The conversion can be performed either by Groups of Pictures (GOP), or independently on each frame of the sequence, or even more locally, by blocks of pixels. The GOP-wise approach combined with spatial and temporal predictions in the encoder gives the best results for low bit rate compression. The block-wise approach can reduce the bit depth with less data loss but breaks the continuity between the blocks, which degrades the Rate Distortion (RD) performance especially at low bit rates. However, we have shown that this approach gives the best results in the context of near lossless compression. The frame-wise version is intermediate between the global (GOP-wise) and local (block-wise) versions. It is adapted to high quality compression. This method was also compared to another frame-wise conversion method in the recent literature called adaptive LogLuv transform, and a 50% rate saving was obtained at high bitrates.

6.3.5. HEVC coding optimization

Participants: Nicolas Dhollande, Christine Guillemot, Bihong Huang, Olivier Le Meur.

The team has two collaborations in the area of HEVC-based video coding optimization. The first research activity is carried out in collaboration with Orange labs (Felix Henry) and UPC (Philippe Salembier) in Barcelona. The objective is to design novel methods for predicting the residues resulting from spatio-temporal prediction. We have indeed observed that the redundancy in residual signals (hence the potential rate saving) is high. In 2013, different methods have been investigated to remove this redundancy, such as generalized lifting and different types of predictors. The generalized lifting is an extension of the lifting scheme of classical wavelet transforms which permits the creation of nonlinear and signal probability density function (pdf) dependent and adaptive transforms.

The second collaboration is with Thomson Video Networks and aims at designing an innovative architecture for effective real-time broadcast encoders of Ultra High Definition (UHD) contents. Currently, the only way to transmit acceptable UHD contents around $10-20$ Mbits/sec is the new compression standard HEVC (finalized in January 2013). Yet, UHD requires at minimum 8 times more computation than the actual HDTV formats, and HEVC has a computing complexity which is already from 2 to 10 times that of MPEG4-AVC. To reduce the encoding complexity on UHD content, a pre-analysis with a lower resolution version (HD) of the input content has been considered to infer some decisions and coding parameters on the UHD video. A speed-up of a factor 3 has already been achieved for a small rate loss of $4-5\%$.

6.4. Distributed processing and robust communication

Information theory, stochastic modelling, robust detection, maximum likelihood estimation, generalized likelihood ratio test, error and erasure resilient coding and decoding, multiple description coding, Slepian-Wolf coding, Wyner-Ziv coding, information theory, MAC channels

6.4.1. Loss concealment based on video inpainting

Participants: Mounira Ebdelli, Christine Guillemot, Ronan Le Boulch, Olivier Le Meur.

We have developed a loss concealment scheme based on a new hierarchical video examplar-based inpainting algorithm. The problem of loss concealment is to estimate unknown pixels after decoding when the corresponding transport packets have been lost on the transmission network. Before proceeding to the video texture inpainting, the motion vectors of the lost blocks must first be estimated from the motion vectors of the received blocks in the spatial neighborhood. The Motion vectors (MV) of damaged blocks are estimated using a Bilinear Motion Field Interpolation (BMFI) technique.

The algorithm follows a coarse to fine approach and first inpaints a low resolution version of the damaged video. Moving objects, detected thanks to the estimated motion vectors, are processed first. The most similar patches (similar to the known pixels of the patch to be completed) is searched within a motion-compensated window in adjacent frames, and used as an estimate of the pixels to be filled in. Then the static background is inpainted using known co-located pixels of neighboring frames. The remaining holes are filled-in using spatial inpainting.

In a second step, the high frequency details of the inpainted areas are recovered using a super-resolution technique, in the same vein as described in Section 6.2.1 for still images. The inpainted low resolution video is first interpolated using a simple lanczos interpolation. The idea is then to search for the nearest neighbor (the best match) of the interpolated version of each inpainted block, within the known part of the current image of the impaired video at the native resolution. The found correspondences form a so-called nearest neighbor field (NNF) which connects inpainted and interpolated patches of the low resolution video to high resolution patches of known parts of the high resolution (HR) video. The found NN patch is then copied to replace the low resolution inpainted patch.The two-step approach allows significantly reducing the execution time of the video inpainting process, while preserving a satisfactory quality.

6.4.2. Universal distributed coding

Participant: Aline Roumy.

In 2012, we started a new collaboration with Michel Kieffer and Elsa Dupraz (Supelec, L2S) on universal distributed source coding. Distributed source coding refers to the problem where several correlated sources need to be compressed without any cooperation at the encoders. Decoding is however performed jointly. This problem arises in sensor networks but also in video compression techniques, where the correlation between the successive frames is not directly used at the encoder, and are therefore seen as distributed. Traditional approaches (from an information theoretical but also practical point of view) assume that the correlation channel between the sources is perfectly known. Since this assumption is not satisfied in practice, a way to get around this is to use a feedback channel (from the decoder to the encoder), that can trigger the encoder.

Instead, we consider universal distributed source coding, where the correlation channel is unknown and belongs to a class parametrized by some unknown parameter vector. We proposed four uncertainty models that depend on the partial knowledge we have on the correlation channel and derived the information theoretical bounds [28]. A complete coding scheme has also been proposed that works well for any distribution in the class [27]. At the encoder, the proposed scheme encompasses the determination of the coding rate and the design of the encoding process. Both contributions result from the information-theoretical compression bounds of universal lossless source coding with side information. Then a novel decoder is proposed that takes into account the available information regarding the class. The proposed scheme avoids the use of a feedback channel or the transmission of a learning sequence, which both would result in a rate increase at finite length.

SISYPHE Project-Team

6. New Results

6.1. Neuroscience & Neuroendocrinology: Regulation of the Gonadotrope axis

6.1.1. A numerical method for transport equations with discontinuous flux functions: application to mathematical modeling of cell dynamics

Participants: Benjamin Aymard, Frédérique Clément, Frédéric Coquel, Marie Postel.

We have proposed a numerical method to handle discontinuous fluxes arising in transport-like equations [35]. More precisely, we have studied 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 has been adapted to treat the discontinuities arising at interfaces. The validation of the method has been done on one- and two-dimensional toy problems for which exact solutions are available, allowing us to do a thorough convergence study. We have then applied the method to a biological model focusing on complex cell dynamics [40] that initially motivated this study and illustrates the full potentialities of the scheme.

6.1.2. Adaptive mesh refinement strategy for a nonconservative transport problem

Participants: Benjamin Aymard, Frédérique Clément, Marie Postel.

In the framework of transport equations it is usual to need long time simulations, and therefore large physical domains to cover a phenomenon. On the other hand it can happen that only a small time varying portion of the domain is interesting. This motivates the use of adaptivity for the spatial discretization. Biological models involving cell development are often nonconservative to account for cell division. In that case the threshold controlling the spatial adaptivity may have to be time-dependent in order to keep up with the progression of the solution. We have tackled the difficulties arising when applying a multiresolution method to a transport equation with discontinuous fluxes modeling localized mitosis [76]. The analysis of the numerical method is performed on a simplified model and numerical scheme. An original threshold strategy is proposed and validated thanks to extensive numerical tests. It is then applied to a biological model in both cases of distributed and localized mitosis.

6.1.3. Coupled Somatic Cell Kinetics and Germ Cell Growth: Multiscale Model-Based Insight on Ovarian Follicular Development

Participants: Frédérique Clément, Philippe Michel, Danielle Monniaux, Thomas Stiehl.

We have designed a stochastic individual-based model describing the first stages of follicular development, where the cell population is structured with respect to age (progression within the cell cycle) and space (radial distance from the oocyte) [39]. The model accounts for the molecular dialogue existing between the oocyte and granulosa cells. Three dynamically interacting scales are considered in the model: (i) a microscopic, local scale corresponding to an individual cell embedded in its immediate environment, (ii) a mesoscopic, semi-local scale corresponding to anatomical or functional areas of follicles and (iii) a macroscopic, global scale corresponding to the morphology of the follicle. Numerical simulations were performed to reproduce the 3D morphogenesis of follicles and follow simultaneously the detailed spatial distribution of individual granulosa cells, their organization as concentric layers or functional cell clones and the increase in the follicle size. Detailed quantitative simulation results have been provided in the ovine species, in which well characterized genetic mutations lead to a variety of phenotypic follicle morphogenesis. The model can help to explain pathological situations of imbalance between oocyte growth and follicular cell proliferation.

6.1.4. Innovative computational and theoretical tools for slow-fast dynamics

Participants: Mathieu Desroches, Maciej Krupa.

Mixed-Mode Bursting Oscillations: Dynamics created by a slow passage through spike-adding canard explosion in a square-wave burster [44]. This work concerns the phenomenon of Mixed-Mode Bursting Oscillations (MMBOs). These are solutions of fast-slow systems of ordinary differential equations that exhibit both small-amplitude oscillations (SAOs) and bursts consisting of one or multiple large-amplitude oscillations (LAOs). The name MMBO is given in analogy to Mixed-Mode Oscillations, which consist of alternating SAOs and LAOs, without the LAOs being organized into burst events. In this article, we show how MMBOs are created naturally in systems that have a spike-adding bifurcation or spike-adding mechanism, and in which the dynamics of one (or more) of the slow variables causes the system to pass slowly through that bifurcation. Canards are central to the dynamics of MMBOs, and their role in shaping the MMBOs is two-fold: saddle-type canards are involved in the spike-adding mechanism of the underlying burster and permit one to understand the number of LAOs in each burst event, and folded-node canards arise due to the slow passage effect and control the number of SAOs. The analysis is carried out for a prototypical fourth-order system of this type, which consists of the third-order Hindmarsh-Rose system, known to have the spike-adding mechanism, and in which one of the key bifurcation parameters also varies slowly. We also include a discussion of the MMBO phenomenon for the Morris-Lecar-Terman system. Finally, we discuss the role of the MMBOs to a biological modeling of secreting neurons.

Canards in piecewise-linear systems: explosions and super-explosions [43]. We show that a planar slow-fast piecewise-linear (PWL) system with three zones admits limit cycles that share a lot of similarity with van der Pol canards, in particular an explosive growth. Using phase-space compactification, we show that these quasi-canard cycles are strongly related to a bifurcation at infinity. Furthermore, we investigate a limiting case in which we show the existence of a continuum of canard homoclinic connections that coexist for a single-parameter value and with amplitude ranging from an order of ε to an order of 1, a phenomenon truly associated with the non-smooth character of this system and which we call super-explosion.

Some results have been obtained concerning numerical continuation techniques for planar slow-fast systems [42] and short-term synaptic plasticity in the deterministic Tsodyks-Markram model that leads to unpredictable network dynamics [41].

6.2. Quantum engineering: controlled quantum systems

Participants: Joachim Cohen, Loïc Herviou, Mazyar Mirrahimi, Pierre Rouchon, Pierre Six.

6.2.1. Schrödinger cat states and hardware efficient quantum error correction

We introduce a new 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 [52]. Such a highly non-classical state is often called a Schrödinger cat state. 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. 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.

In a second contribution [53], we propose to use an encoding of a quantum bit of information in a four-component Schrödinger cat state to ensure its protection against the photon loss, being the major source of decoherence for such a quantum harmonic oscillator. This protection is ensured by an efficient quantum error correction scheme employing the nonlinearity provided by a single physical qubit coupled to the cavity. We describe in detail how to implement these operations in a circuit quantum electrodynamics system. This directly addresses the task of building a hardware-efficient quantum memory and can lead to important shortcuts in quantum computing architectures.

As an important step towards the realization of such a protected quantum memory, in a collaboration with the team of Robert J. Schoelkopf at Yale university, we have successfully realized the encoding protocol of [52] using a 3D transom qubit coupled to a waveguide cavity resonator with a highly ideal off-resonant coupling [60]. This dispersive interaction is much greater than decoherence rates and higher-order nonlinearities to allow simultaneous manipulation of hundreds of photons. We created cat states as large as 111 photons and created superpositions of up to four coherent states. This control creates a powerful interface between discrete and continuous variable quantum computation and could enable applications in metrology and quantum information processing. This important achievement was published in Science and was also highlighted in Science Perspectives [103].

6.2.2. Quantum reservoir (dissipation) engineering

We have studied the application of dissipation engineering techniques to perform a high-performance and fast qubit reset [64]. Qubit reset is crucial at the start of and during quantum information algorithms. In a collaboration with the team of Michel H. Devoret at Yale university, our protocol, nicknamed DDROP (Double Drive Reset of Population) was experimentally tested on a superconducting transmon qubit and achieves a ground state preparation of at least 99.5% in times less than $3\mu s$; faster and higher fidelity are predicted upon parameter optimization [46].

Next, we proposed a dissipation engineering scheme that prepares and protects a maximally entangled state of a pair of superconducting qubits [54]. This is done by off-resonantly coupling the two qubits to a low-Q cavity mode playing the role of a dissipative reservoir. We engineer this coupling by applying six continuous-wave microwave drives with appropriate frequencies. The two qubits need not be identical. We show that our approach does not require any fine-tuning of the parameters and requires only that certain ratios between them be large. This protocol was experimentally realized in a collaboration with the team of M. H. Devoret at Yale university [57]. Unlike conventional, measurement-based schemes, this autonomous approach uses engineered dissipation to counteract decoherence, obviating the need for a complicated external feedback loop to correct errors. Instead, the feedback loop is built into the Hamiltonian such that the steady state of the system in the presence of drives and dissipation is a Bell state, an essential building block for quantum information processing. Such autonomous schemes, which are broadly applicable to a variety of physical systems, will be an essential tool for the implementation of quantum error correction. This important result appeared in Nature back-to-back to another paper by the group of D.J. Wineland (2012 Nobel prize winner) at NIST implementing similar ideas on another physical system consisting of trapped ion qubits [105].

6.2.3. Quantum measurement and measurement-based feedback

Measuring a quantum system can randomly perturb its state. The strength and nature of this back-action depend on the quantity that is measured. In a partial measurement performed by an ideal apparatus, quantum physics predicts that the system remains in a pure state whose evolution can be tracked perfectly from the measurement record. This property was proved in a collaboration with the group of Michel H. Devoret (Yale university) using a superconducting qubit dispersively coupled to a cavity traversed by a microwave signal [47]. The back-action on the qubit state of a single measurement of both signal quadratures was observed and shown to produce a stochastic operation whose action is determined by the measurement result. This accurate monitoring of a qubit state is an essential prerequisite for measurement-based feedback control of quantum systems. Indeed, in another experiment performed by our collaborators at ENS (team of Benjamin Huard and François Mallet), we demonstrated stabilization of an arbitrary trajectory of a superconducting qubit by such a measurement-based feedback [37]. The protocol benefits from the long coherence time ($T_2 > 10\mu s$) of the 3D transmon qubit, the high efficiency (82%) of the phase preserving Josephson amplifier, and fast electronics ensuring less than 500 ns delay. At discrete time intervals, the state of the qubit is measured and corrected in case an error is detected. For Rabi oscillations, where the discrete measurements occur when the qubit is supposed to be in the measurement pointer states, we demonstrate an average fidelity of 85% to the targeted trajectory. Incidentally, we demonstrate a fast reset protocol allowing to cool a 3D transmon qubit down to 0.6% in the excited state.

6.3. Classical engineering: Monitoring and control of complex systems

6.3.1. Modeling, signal analysis and control with medical applications

Participants: Alexandre Guerrini, Lisa Guigue, Claire Médigue, Michel Sorine, Serge Steer.

Reduced order cardiac modeling and applications. See Section 4.3 for complements. We consider two topics:
- Personalized medecine: a first validation on clinical data of our model of controlled contraction of cardiac muscle has been obtained [55].
- Heart Failure with preserved Ejection Fraction (HFpEF): this work is done in collaboration with Bijan Gahleh (INSERM U955). Our objective is to define markers of HFpEF identifiable from noninvasive measurements. After having assembled a high precision ECG acquisition and post-processing system, we have measured multi-lead ECG on pigs treated to induce HFpEF, cf. B. Gahleh et al [109]. The analysis of the diastolic electric interval (e.g. P-wave, PR interval etc.) is ongoing.

Semiclassical analysis of cardiovascular signals. A summary of the theory is now published [51].

CGAO-REA: Computerized Glucose Control in Critically Ill Patients. The version CGAO_v1 of our controller (see Sections 4.3), has been used in a large multi-center study, CGAO-REA (35 active ICUs, more than 3500 included patients). Mortality has not been changed [70], [48] but the protocol is now formalized and tunable. CGAO-REA has proved that our controller is robust in the real life context and comparable to human control with its present tuning. Improving the tuning (in particular the glycemic target) seems possible.

6.3.2. Diagnosis of inhomogeneous insulation degradation in electric cables by distributed shunt conductance estimation

Participant: Qinghua Zhang.

For the diagnosis of inhomogeneous insulation degradation in electric cables, the estimation of distributed shunt conductance is studied in this work. Gradual growth of the shunt conductance is a consequence of degradation of the dielectric properties of the insulator. The proposed estimation method is based on voltage and current measurements at a single end of the cable. After the linearization of the bilinear term of the telegrapher's equations through a perturbation approach, the Kalman filter is applied to transform the problem of dynamic system parameter estimation to a simple linear regression problem. Numerical simulations are made to demonstrate the feasibility of the proposed method. In particular, it is shown that the weak sensitivity of the available measurements to the shunt conductance can be compensated by long time data samples. See [61] for more details.

6.3.3. Feasibility of reflectometry techniques for non destructive evaluation of external post-tensioned cables

Participants: Michel Sorine, Qinghua Zhang.

Nowadays a considerable number of bridges is reaching an age when renovating operations become necessary. For some bridges, external post-tension is realized with cables protected in ducts, with the residual internal space imperfectly filled with a fluid cement grout. Detecting the problems of injection in the ducts is visually impossible from the outside. In collaboration with IFSTTAR (Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux) through the I4S team common to Inria and IFSTTAR, the feasibility of reflectometry techniques for cable health monitoring is investigated via numerical simulations and laboratory experiments. The main idea consists in adding electrically conductive tapes along a duct so that the duct and the added tapes can be treated as an electrical transmission line. It is then possible to apply advanced reflectometry methods developed by the SISYPHE project-team, initially for true electric cables.

6.3.4. Nonlinear system identification

Participants: Boyi Ni, Michel Sorine, Qinghua Zhang.

In the framework of the joint Franco-Chinese ANR-NSFC EBONSI project (see Section 8.1.1), the topics studied this year on nonlinear system identification are mainly on the detection of asymmetric control valve stiction from oscillatory data based on a method for extended Hammerstein system identification, and on the identification of Wiener systems.

The study on control valve stiction is motivated by the detection of control valves with asymmetric stiction resulting in oscillations in feedback control loops. The joint characterization of the control valve and the controlled process is formulated as the identification of a class of extended Hammerstein systems. The input nonlinearity is described by a point-slope-based hysteretic model with two possibly asymmetric ascent and descent paths. An iterative identification method is proposed, based on the idea of separating the ascent and descent paths subject to the oscillatory input and output. The structure of the formulated extended Hammerstein system is shown to be identifiable, and the oscillatory signals in feedback control loops are proved to be informative by exploiting the cyclo-stationarity of these oscillatory signals. Numerical, experimental and industrial examples confirm the effectiveness of the proposed identification method.

Wiener system identification has been investigated this year by focusing on the estimation of the finite impulse response (FIR) of the linear subsystem. Under the assumption of Gaussian input distribution, this work mainly aims at addressing a deficiency of the well-known correlation-based method for Wiener system identification: it fails when the nonlinearity of the Wiener system is an even function. This method is, in the considered Gaussian input case, equivalent to the best linear approximation (BLA), which exhibits the same deficiency. A new method is developed this year, based on a weighted principal component analysis (wPCA). Its consistency is proved for Wiener systems with either even or non even nonlinearities. Its computational cost is almost the same as that of a standard PCA. Numerical simulations are made to compare the new wPCA-based method to the correlation-based method for different Wiener systems with nonlinearities more or less close to an even function.

6.3.5. Model-based fault diagnosis for descriptor systems

Participants: Abdouramane Moussa Ali, Qinghua Zhang.

This work is about fault diagnosis for linear time varying descriptor systems, the discrete time counterpart of dynamic systems described by differential-algebraic equations. The Kalman filter for descriptor systems is first revisited by completing existing results about its properties that are essential for the purpose of fault diagnosis. Based on the analysis of the effects of the considered actuator and sensor faults on the innovation of the Kalman filter, it is shown that the considered fault diagnosis problem in linear time varying descriptor systems can be transformed to a classical linear regression problem formulated by appropriately filtering the input-output data. Following this result, algorithms for fault diagnosis through maximum likelihood estimation are then developed.

In the framework of the ITEA2 MODRIO project (see Section 8.2.1), this work is in preparation for studying hybrid system monitoring, aiming at extending existing results from state-space systems to descriptor systems in the modes of a hybrid system.

6.3.6. Analysis of the Behavior of Networks of Dynamical Systems

Participant: Pierre-Alexandre Bliman.

We have established convergence results for some continuous-time dynamics which are analogs to ant colony optimization algorithms that solve shortest path problems. Global asymptotic stability has been shown, and the speed of convergence has been calculated explicitly and shown to be proportional to the difference between the reciprocals of the second shortest and the shortest paths. Such precise results are missing in the context of ant colony optimization algorithms (which are discrete-time dynamical systems). The systems studied are special instances of networks of dynamical systems which represent the evolution of some state variable on each path, coupled in a competitive way through global macroscopic quantity. Such models are related to simple forms of models studied in mathematical epidemiology, which will be the subject of further work. This work is done in cooperation with Amit Bhaya from COPPE, Universidade Federal de Rio de Janeiro. Papers have been submitted [77].

SMIS Project-Team

6. New Results

6.1. Minimum Exposure

Participants: Nicolas Anciaux, Marouane Fazouane, Benjamin Nguyen [correspondent], 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.

Following our 2012 seminal works on the general Minimum Exposure framework, we have pursued its general study in 2013 [15], [29]. We have also developed a prototype system, using a low powered and highly secure smartcard [21], which is used to support hidden decision rules.

6.2. Flash-Based Data Management

Participants: Nicolas Anciaux, Matias Bjørling, Philippe Bonnet, Luc Bouganim [correspondent], Niv Dayan, Philippe Pucheral.

Mass-storage secure portable tokens are emerging and provide a real breakthrough in the management of sensitive data. They can embed personal data and/or metadata referencing documents stored encrypted in the Cloud and can manage them under holder's control. Mass on-board storage requires efficient embedded database techniques. These techniques are however very challenging to design due to a combination of conflicting NAND Flash constraints and scarce RAM constraint, disqualifying known state of the art solutions. To tackle this challenge, we proposed a log-only based storage organization and an appropriate indexing scheme, which (1) produce only sequential writes compatible with the Flash constraints and (2) consume a tiny amount of RAM, independent of the database size [13].

Solid State Drives (SSDs) are a moving target for system designers: they are black boxes, their internals are undocumented, and their performance characteristics vary across models. There is no appropriate analytical model and experimenting with commercial SSDs is cumbersome, as it requires a careful experimental methodology to ensure repeatability. Worse, performance results obtained on a given SSD cannot be generalized. Overall, it is impossible to explore how a given algorithm, say a hash join or LSM-tree insertions, leverages the intrinsic parallelism of a modern SSD, or how a slight change in the internals of an SSD would impact its overall performance. In 2013, we worked on a new SSD simulation framework, named EagleTree, which addresses these problems, and enables a principled study of SSD-Based algorithms. We published a demonstration on EagleTree at VLDB'13 [20]. The demonstration scenario illustrates the design space for algorithms based on an SSD-based IO stack, and shows how researchers and practitioners can use EagleTree to perform tractable explorations of this complex design space.

6.3. Secure Global Computing on Asymmetric Architecture

Participants: Benjamin Nguyen [correspondent], Philippe Pucheral, Cuong Quoc To.

Current applications, from complex sensor systems (e.g. quantified self) to online e-markets acquire vast quantities of personal information which usually ends-up on central servers. Decentralized architectures, devised to help individuals keep full control of their data, hinder global treatments and queries, impeding the development of services of great interest. In this study, we promote the idea of pushing the security to the edges of applications, through the use of secure hardware devices controlling the data at the place of their acquisition. To solve this problem, we propose secure distributed querying protocols based on the use of a tangible physical element of trust, reestablishing the capacity to perform global computations without revealing any sensitive information to central servers. This leads to execute global treatments on an asymmetric architecture, composed of a powerful, available and untrusted computing infrastructure (server or cloud), and a large set of low powered, highly disconnected trusted devices. Given our large scale data centric applications (e.g. nationwide surveys), we discard solutions based on secure multi-party computation, which do not scale. We have primarily 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 [12]. This work is an extension of [26]. A vulgarization paper on the scientific and societal challenges related to PPDP techniques has been published in a newspaper [24]. We are now trying to support general SQL queries in this same execution context. We concentrate first on the subset of SQL queries without joins, but including Group By and aggregates, and show how to secure their execution in the presence of honest-but-curious attackers. Cost models and experiments demonstrate that this approach can scale to nationwide infrastructures [23], [42]. This work is part of Cuong Quoc To's Ph.D. thesis started in sept. 2012.

6.4. Trusted Cells

Participants: Nicolas Anciaux, Philippe Bonnet, Luc Bouganim, Benjamin Nguyen, Pilippe Pucheral [correspondent], 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). This work was based on a thorough analysis of existing and potential threats on personal data, which led to a tutorial on data privacy [18], [30].

In parallel, we revisited the Trusted Cells vision to the context of Least Developed Countries (LDCs). The main barrier to the development of IT services in these regions is not only the lack of communication facilities, but also the lack of consistent information systems, security procedures, economic and legal support. We propose, Folk-enabled Information System (Folk-IS), a new paradigm based on a fully decentralized and participatory approach, where each individual implements a small subset of a complete information system without the need for infrastructure. As trusted cells, Folk-IS builds upon the emergence of highly secure, portable, low-cost storage and computing devices, called hereafter Smart Tokens. Here, however, the focus is on the low cost of ownership, deployment and maintenance, and on the absence of a networked infrastructure. With Folk-IS, and thanks to smart tokens, people will transparently and opportunistically perform data management and networking tasks as they physically move, so that IT services are truly delivered by the crowd [17].

SOCRATE Project-Team

6. New Results

6.1. Flexible Radio Front-End

The contributions on hardware design are twofold. First, the development of a Full-Duplex architecture for OFDM systems. Second, a proposal of a Wake-Up scheme for home networking with reduced power consumption.

6.1.1. Full-Duplex systems

Zhan et al. [23] focused on the study of active analog self-interference cancellation (AASIC) techniques in full-duplex OFDM systems. This original approach aims at proposing a cancellation technique at RF level for wideband systems. A theoretical study confronted to simulations was proposed with a particular emphasis on the channel estimation of the interfering signal. This study was completed with an analysis on the phase noise and the thermal noise impact.

6.1.2. Wake-Up Architectures

Khoumeri et al. [28] proposed radio architectures for allowing energy savings by letting devices to switch off part of the transmission components when they are not in use. Based on classical WiFi systems, the proposed architecture offers the ability to use a conventional emitter, using only a particular subcarrier fingerprint to identify the node to wake-up, hence avoiding a high level of false wake-up.

6.2. Agile Radio Resource Sharing

The contributions of the axis in *agile radio resource sharing* can be gathered in three groups: (a) green communications; (b) performance analysis; and (c) scheduling and power allocation techniques.

6.2.1. Green Communications

The main contributions in the subject of green communications focus on the problem of increasing the energy efficiency of Orthogonal Frequency-Division Multiple Access (OFDMA) wireless networks. In particular, Tsilimantos et al. in [21] and Hasan et al. in [15] studied different techniques to strategically switch off some of the base stations in cellular systems while guaranteeing a given quality of service (QoS). In [21], the authors use methods from stochastic geometry to determine the number of active cells that can be switched off while the outage probability, or equivalently the signal to interference plus noise ratio (SINR), remains the same. In [15], this problem is studied from a decentralized point of view using methods from coalitional game theory.

6.2.2. Performance Analysis

The contributions in performance analysis are mainly oriented to the field of body area networks (BANs), [17] indoor adaptive OFDM wireless networks and relay channels.

In [17], Lauzier et al. presented the results of a measurement campaign whose primary objective was to characterize the complete mesh of a BAN and analyze the quality of every radio link between the different nodes. In [18], [19], the Multi-Resolution Frequency Domain ParFlow (MRFDPF) model is used to calculate the bit error rate (BER) and study the feasibility of adaptive modulation in OFDM systems.

In the context of relay channels, Ferrand et al. [32] studied the asymptotic *coding gain* of the packet error rate of relay channels in which the radio links are subject to both fading and log-normal shadowing effects simultaneously.

6.2.3. Scheduling and Power Allocation Techniques

Power allocation techniques and scheduling were studied by Ferrand et al. [33] and Wang et al. [9]. More specifically, advances in the study of the achievable rate region of relay channels in the case of global power constraints were reported in [33]. Cooperative scheduling techniques in the context of BAN were proposed in [9] to reduce inter-BAN interference using tools from game theory.

6.3. Software Radio Programming Model

Software defined radio (SDR) technology has evolved rapidly and is now reaching market maturity. Still, a lot of issues have yet to be studied. Mickaël Dardaillon, Kevin Marquet, Tanguy Risset and others highlighted the constraints imposed by recent radio protocols and presented current architectures, solutions, and challenges for programming SDR [31].

6.3.1. Dataflow programming

To enable dynamic adaptation of computation intensive multimedia dataflow applications, Lionel Morel, Kevin Marquet and others have studied language extensions, together with the corresponding run-time support. They show that this approach can be used to monitor and control throughput [20] and offer quality of service [29], with a low impact on the overall performance.

6.3.2. Energy-efficient Localization

Guillaume Salagnac and others address the tradeoff between energy consumption and localization performance in a mobile sensor network application [7]. The focus is on augmenting GPS location with more energy-efficient location sensors to bound position estimate uncertainty while GPS is off. Such combined strategies can cut node energy consumption by one third while still meeting application-specific positioning criteria.

6.3.3. Swap Fairness for Thrashing Mitigation

In the context of shared hosting or virtualization, where multiple users run uncoordinated and selfish workloads, François Goichon, Guillaume Salagnac and Stéphane Frénot introduced an accounting layer that forces swap fairness among processes competing for main memory [13]. It ensures that a process cannot monopolize the swap subsystem by delaying the swap operations of abusive processes, reducing the number of system-wide page faults while maximizing memory utilization.

SPADES Team

6. New Results

6.1. Components and Contracts

Participants: Gregor Goessler, Quentin Sabah, Jean-Bernard Stefani.

6.1.1. Analysis of logical causality

The failure of one component may entail a cascade of failures in other components; several components may also fail independently. In such cases, elucidating the exact scenario that led to the failure is a complex and tedious task that requires significant expertise.

The notion of causality *(did an event e cause an event e'?)* has been studied in many disciplines, including philosophy, logic, statistics, and law. The definitions of causality studied in these disciplines usually amount to variants of the counterfactual test "e is a cause of e' if both e and e' have occurred, and in a world that is as close as possible to the actual world but where e does not occur, e' does not occur either". Surprisingly, the study of logical causality has so far received little attention in computer science, with the notable exception of [69] and its instantiations. However, this approach relies on a causal model that may not be known, for instance in presence of black-box components.

Improving on previous results, we have proposed in [21] an approach to enhance the fault diagnosis in black-box component-based systems, in which only events on component interfaces are observable. For such systems, we have described a causality analysis framework that helps us establish the causal relationship between component failures and system failures, given an observed system execution trace. The analysis is based on a formalization of counterfactual reasoning, and applicable to real-time systems. We have illustrated the analysis with a case study from the medical device domain.

In [5] we have proposed a formal framework for reasoning about causality, and blaming system-level failures on the component(s) that caused them. The framework is general in the sense that it applies to many different models of computation and communication (MoC), such as synchronous and asynchronous computation, and communication by messages or shared variables. We are currently instantiating the framework to specific MoC, in particular, to timed automata, and developing a refinement of our original approach that reduces the number of false positives.

6.1.2. Supporting isolation for actors in shared memory

The actor model of concurrency, as supported *e.g.*, by the Erlang programming language, is an appealing programming model for the construction of concurrent and distributed systems, and multicore programming in particular. Although much work has taken place in particular during the past ten years on efficient implementations of the actor model, the design space is far from being completely understood.

As part of Quentin Sabah's thesis [10], we have developed a variant of the actor model that, in contrast to previous works, ensures a strict isolation between actors while imposing no restriction on the form of data exchanged in messages. We have formally specified an abstract machine, called SIAAM (see Sec.5.4.5), for an extension of the Java language with our actor model, and implemented it as a modified Jikes virtual machine, a state of the art Java virtual machine. A combination of points-to and live variable analyses has been implemented using the Soot framework, that can be used to remove unnecessary read and write checks for isolation. A diagnosis tool built on top of the analyses helps programmers to pinpoint potential problems (exceptions raised indicating a potential violation of isolation). We have shown with artificial and small applicative benchmarks that, using our analyses to improve performance, our implementation is reasonably efficient and imposes low overhead for the benefit of strict isolation.

In addition, we have developed a Coq proof of the isolation property enforced by SIAAM, namely that no information between actors can take place outside of message exchanges, despite the presence of a shared heap between actors.

6.2. Real-Time multicore programming

Participants: Vagelis Bebelis, Gwenaël Delaval, Pascal Fradet, Alain Girault, Gregor Goessler, Bertrand Jeannet, Gideon Smeding, Jean-Bernard Stefani.

6.2.1. A time predictable programming language for multicores

Time predictability (PRET) is a topic that emerged in 2007 as a solution to the ever increasing unpredictability of today's embedded processors, which results from features such as multi-level caches or deep pipelines [57]. For many real-time systems, it is mandatory to compute a strict bound on the program's execution time. Yet, in general, computing a tight bound is extremely difficult [90]. The rationale of PRET is to simplify both the programming language and the execution platform to allow more precise execution times to be easily computed [39].

Following our past results on the PRET-C programming language [35], we have proposed a time predictable synchronous programming language for multicores, called FOREC. It extends C with a small set of ESTEREL-like synchronous primitives to express concurrency, interaction with the environment, looping, and a synchronization barrier [22] (like the `pause` statement in ESTEREL). FOREC threads communicate with each other via shared variables, the values of which are combined at the end of each tick to maintain deterministic execution. FOREC is compiled into threads that are then statically scheduled for a target multicore chip. Our WCET analysis takes into account the access to the shared TDMA bus and the necessary administration for the shared variables. We achieve a very precise WCET (the over-approximation being less than 2%) thanks to a reachable space exploration of the threads' states.

This work has been conducted within the RIPPES associated team.

6.2.2. WCET analysis

Our past work on the WCET analysis of PRET-C programs has led us to design static analyses, for instance to prune unfeasible paths in the control flow graph [36]. In 2013, we have worked on how to take into account direct mapped instruction caches in WCET analysis. Instruction caches are essential to address if one wants to analyze large embedded programs. Our cache analysis technique offers the same precision as the most precise techniques [80], while improving analysis time by up to 240 times. This improvement is achieved by analyzing individual blocks of the control flow graph separately, and by proposing a tailored abstract domain to represent efficiently the cache state [14], [25]. In contrast with previous abstract analysis methods [88], [85], our analysis is able to offer the same precision as the concrete approaches [80].

6.2.3. Tradeoff exploration between reliability, power consumption, and execution time

For autonomous critical real-time embedded systems (*e.g.*, satellites), guaranteeing a very high level of reliability is as important as keeping the power consumption as low as possible. We have designed an off-line ready list scheduling heuristics which, from a given software application graph and a given multiprocessor architecture (homogeneous and fully connected), produces a static multiprocessor schedule that optimizes three criteria: its *length* (crucial for real-time systems), its *reliability* (crucial for dependable systems), and its *power consumption* (crucial for autonomous systems). Our tri-criteria scheduling heuristics, *TSH*, uses the *active replication* of the operations and the data-dependencies to increase the reliability, and uses *dynamic voltage and frequency scaling* to lower the power consumption [37], [38]. TSH implements a ready list scheduling heuristics, and we have formulated a new multi-criteria cost function such that we are able to prove rigorously that the static schedules we generate meet both the reliability constraint and the power consumption constraint [12].

By running TSH on a single problem instance, we are able to provide the Pareto front for this instance in 3D, therefore exposing the user to several tradeoffs between the power consumption, the reliability and the execution time. Thanks to extensive simulation results, we have shown how TSH behaves in practice. Firstly, we have compared TSH versus an optimal Mixed Linear Integer Program on small instances; the experimental results show that TSH behaves very well compared to the ILP. Secondly, we have compared TSH with the ECS heuristic (Energy-Conscious Scheduling [77]); the experimental results show that TSH performs systematically better than ECS.

This is a joint work with Ismail Assayad (U. Casablanca, Morocco) and Hamoudi Kalla (U. Batna, Algeria), who both visit the team regularly.

6.2.4. Modular distribution

Synchronous programming languages describe functionally centralized systems, where every value, input, output, or function is always directly available for every operation. However, most embedded systems are nowadays composed of several computing resources. The aim of this work is to provide a language-oriented solution to describe *functionally distributed reactive systems*. This research started within the Inria large scale action SYNCHRONICS and is a joint work with Marc Pouzet (ENS, PARKAS team from Rocquencourt) and Xavier Nicollin (Grenoble INP, VERIMAG lab).

We are working on type systems to formalize, in a uniform way, both the clock calculus and the location calculus of a synchronous data-flow programming language (the HEPTAGON language, inspired from LUCID SYNCHRONE [49]). On one hand, the clock calculus infers the clock of each variable in the program and checks the clock consistency: *e.g.*, a time-homogeneous function, like +, should be applied to variables with identical clocks. On the other hand, the location calculus infers the spatial distribution of computations and checks the spatial consistency: *e.g.*, a centralized operator, like +, should be applied to variables located at the same location. Compared to the PhD of Gwenaël Delaval [55], [56], the goal is to achieve *modular* distribution. By modular, we mean that we want to compile each function of the program into a single function capable of running on any computing location. We make use of our uniform type system to express the computing locations as first-class abstract types, exactly like clocks. It allows us to compile a typed variable (typed by both the clock and the location calculi) into `if ... then ... else ...` structures, whose conditions will be valuations of the clock and location variables.

We currently work on an example of software-defined radio. We have shown on this example how to use a modified clock calculus to describe the localisation of values as clocks, and the architecture as clocks (for the computing resources) and their relations (for communication links).

6.2.5. Distribution of synchronous programs under real-time constraints

The goal of Gideon Smeding's PhD thesis [11] was to propose a quasi-synchronous framework encompassing constraints on the relative speed of clocks, together with a formalism for reasoning about clock-dependent properties within the model. This framework should provide a seamless link between synchronous models and their asynchronous implementation.

The quasi-synchronous approach developed in [11] considers independently clocked, synchronous components that interact via communication-by-sampling or FIFO channels. We have defined relative drift bounds on pairs of recurring events such as clock ticks or the arrival of a message. Drift bounds express constraints on the stability of clocks, *e.g.*, at least two ticks of one per three consecutive ticks of the other. We can thus move from total synchrony, where all clocks tick simultaneously, to global asynchrony by relaxing the drift bounds. As constraints are more relaxed, behavior diverges more and more from synchronous system behavior. In many systems, such as distributed control systems, occasional deviations of input and output signals of the controller from their behavior in the synchronous model may be acceptable as long as the frequency of such deviations is bounded. The approach of [11] takes as inputs a program written in a Lustre-like language extended with asynchronous communication by sampling, application requirements on the distribution in the form of weakly-hard constraints [45] bounding *e.g.*, the tolerated loss of data tokens, and platform assertions (*e.g.*, relative clock speeds, available communication resources), and verifies whether the program meets the requirements under the platform assertions.

6.2.6. Analysis and scheduling of parametric dataflow models

Recent data-flow programming environments support applications whose behavior is characterized by dynamic variations in resource requirements. The high expressive power of the underlying models (*e.g.*, Kahn Process Networks or the CAL actor language) makes it challenging to ensure predictable behavior. In particular, checking *liveness* (*i.e.*, no part of the system will deadlock) and *boundedness* (*i.e.*, the system can be executed in finite memory) is known to be hard or even undecidable for such models. This situation is troublesome for the design of high-quality embedded systems.

Last year, we have introduced the *schedulable parametric data-flow (SPDF)* MoC for dynamic streaming applications [60]. SPDF extends the standard dataflow model by allowing rates to be parametric. SPDF was designed to be statically analyzable while retaining sufficient expressive power.

Following the same lines, we have recently proposed the *Boolean Parametric Data Flow (BPDF)* MoC which combines integer parameters (to express dynamic rates) and boolean parameters (to express the activation and deactivation of communication channels) [15], [26], [24]. High dynamism is provided by integer parameters which can change at each basic iteration and boolean parameters which can change even within the iteration. We have presented static analyses which ensure statically the liveness and the boundedness of BDPF graphs. Our case studies are video decoders for high definition video streaming such as VC-1.

We have proposed a generic and flexible framework to generate parallel ASAP schedules targeted to the new STHORM many-core platform designed by STMicroelectronics [29], [23]. The parametric dataflow graph is associated with generic or user-defined specific constraints aimed at minimizing, timing, buffer sizes, power consumption, or other criteria. The scheduling algorithm executes with minimal overhead and can be adapted to different scheduling policies just by changing some constraints. The safety of both the dataflow graph and constraints can be checked statically and all schedules are guaranteed to be bounded and deadlock free. This parallel scheduling framework has been developed for a parametric MoC without booleans. We are now focusing on extending it to BPDF applications.

This research is the central topic of Vagelis Bebelis' PhD thesis. It is conducted in collaboration with STMicroelectronics.

6.2.7. Abstract Acceleration of general linear loops

We have investigated abstract acceleration techniques for computing loop invariants for numerical programs with linear assignments and conditionals. Whereas abstract interpretation techniques typically over-approximate the set of reachable states iteratively, abstract acceleration captures the effect of the loop with a single, non-iterative transfer function applied to the initial states at the loop head.

In contrast to previous acceleration techniques, our approach applies to any linear loop without restrictions. Its novelty lies in the use of the Jordan normal form decomposition of the loop body to derive symbolic expressions for the entries of the matrix modeling the effect of $n >= 0$ iterations of the loop. The entries of such a matrix depend on n through complex polynomial, exponential and trigonometric functions. Therefore, we introduced an abstract domain for matrices that captures the linear inequality relations between these complex expressions. This results in an abstract matrix for describing the fixpoint semantics of the loop. We also developed a technique to take into account the guard of the loop by bounding the number of loop iterations, which relies again on the Jordan normal form decomposition.

Our approach integrates smoothly into standard abstract interpreters and can handle programs with nested loops and loops containing conditional branches. We evaluate it over small but complex loops that are commonly found in control software, comparing it with other tools for computing linear loop invariants. The loops in our benchmarks typically exhibit polynomial, exponential and oscillatory behaviors that present challenges to existing approaches, that are either too unprecise (classical abstract interpretation) or limited to a restricted class of loops (*e.g.*, translation with resets in the case of abstract acceleration, or stable loops, in the sense of control theory, for ellipsoid methods). Our approach finds non-trivial invariants to prove useful bounds on the values of variables for such loops, clearly outperforming the existing approaches in terms of precision while exhibiting good performance.

A paper presenting this technique has been accepted to POPL'2014. An extended version has been published in arXiv [30].

6.2.8. Synthesis of switching controllers using approximately bisimilar multiscale abstractions

The use of discrete abstractions for continuous dynamics has become standard in hybrid systems design (see *e.g.*, [87] and the references therein). The main advantage of this approach is that it offers the possibility to leverage controller synthesis techniques developed in the areas of supervisory control of discrete-event systems [82]. The first attempts to compute discrete abstractions for hybrid systems were based on traditional systems behavioral relationships such as simulation or bisimulation, initially proposed for discrete systems most notably in the area of formal methods. These notions require inclusion or equivalence of observed behaviors which is often too restrictive when dealing with systems observed over metric spaces. For such systems, a more natural abstraction requirement is to ask for closeness of observed behaviors. This leads to the notions of approximate simulation and bisimulation introduced in [61].

These approaches are based on sampling of time and space where the sampling parameters must satisfy some relation in order to obtain abstractions of a prescribed precision. In particular, the smaller the time sampling parameter, the finer the lattice used for approximating the state-space; this may result in abstractions with a very large number of states when the sampling period is small. However, there are a number of applications where sampling has to be fast; though this is generally necessary only on a small part of the state-space. We have been exploring two approaches to overcome this state-space explosion.

In [52], we have proposed a technique for the synthesis of safety controllers for switched systems using multi-scale abstractions that allow us to deal with fast switching while keeping the number of states in the abstraction at a reasonable level. The finest scales of the abstraction are effectively explored only when fast switching is needed, that is when the system approaches the unsafe set. We have implemented these results in the tool COSYMA (COntroller SYnthesis using Multi-scale Abstractions, see Sec. 5.4.2) [20]. The tool accepts a description of a switched system represented by a set of differential equations and the sampling parameters used to define an approximation of the state-space on which discrete abstractions are computed. The tool generates a controller — if it exists — for the system that enforces a given safety or time-bounded reachability specification.

In [19], we have presented an approach using mode sequences of given length as symbolic states for our abstractions. We have shown that the resulting symbolic models are approximately bisimilar to the original switched system and that an arbitrary precision can be achieved by considering sufficiently long mode sequences. The advantage of this approach over existing ones is double: first, the transition relation of the symbolic model admits a very compact representation under the form of a shift operator; second, our approach does not use lattices over the state-space and can potentially be used for higher dimensional systems. We have provided a theoretical comparison with the lattice-based approach and presented a simple criterion enabling to choose the most appropriate approach for a given switched system. We have applied the approach to a model of road traffic for which we have synthesized a schedule for the coordination of traffic lights under constraints of safety and fairness.

6.3. Language Based Fault-Tolerance

Participants: Dmitry Burlyaev, Pascal Fradet, Alain Girault, Jean-Bernard Stefani.

6.3.1. Automatic Transformations for Fault tolerant Circuits

In the recent years, we have studied the implementation of specific fault tolerance techniques in real-time embedded systems using program transformation [1]. We are now investigating the use of automatic transformations to ensure fault-tolerance properties in digital circuits. To this aim, we consider program transformations for hardware description languages (HDL). We have designed a simple hardware description language inspired from LUSTRE and Lucid Synchrone. It is a core functional language manipulating synchronous boolean streams. We consider both single-event upsets (SEU) and single-event transients (SET) and all fault models of the form "at most 1 SEU or SET within n clock signals". The language's semantics as well as fault modes have been formalized in Coq and many basic (library) properties have been shown on that language.

We have expressed several variants of triple modular redundancy (TMR) as program transformations. We have proposed a verification-based approach to minimize the number of voters in TMR [16]. Our technique guarantees that the resulting circuit *(i)* is fault tolerant to the soft-errors defined by the fault model and *(ii)* is functionally equivalent to the initial one. Our approach operates at the logic level and takes into account the input and output interface specifications of the circuit. Its implementation makes use of graph traversal algorithms, fixed-point iterations, and BDDs. Experimental results on the ITC'99 benchmark suite indicate that our method significantly decreases the number of inserted voters which entails a hardware reduction of up to 55% and a clock frequency increase of up to 35% compared to full TMR. We address scalability issues arising from formal verification with approximations and assess their efficiency and precision.

We are currently studying the definition of other fault-tolerant techniques (*e.g.*, time redundancy, mixed time/space redundancy) as program transformations. We are also considering the use of the Coq proof assistant to certify that the transformations make the programs fault tolerant *w.r.t.* specific fault models. Our long term goal is to design an aspect-like language allowing users to specify and tune a wide range of fault tolerance techniques, while ensuring that the corresponding transformations ensure well-defined fault-tolerance properties. The advantage would be to produce fault-tolerant circuits by specifying fault-tolerant properties/strategies separately from their functional specifications.

6.3.2. Concurrent flexible reversibility

In the recent years, we have been investigating reversible concurrent computation, and investigated various reversible concurrent programming models, with the hope that reversibility can shed some light on the common semantic features underlying various forms of fault recovery techniques (including, exceptions, transactions, and checkpoint/rollback schemes).

As part of this research program, we have devised a reversible variant of the higher-order π-calculus, equipped with an imperative rollback operation that allows a concurrent program to be rolled back to a past execution state, and a primitive form of compensation to control (forward execution) after a rollback operation [18]. We have shown that these two extensions provide very powerful primitives for programming different forms of rollback/compensation schemes. We have shown in particular that they are powerful enough to provide a faithful encoding of a notion of communicating transaction proposed in the literature. We have started the development of a behavioral theory for this $\mathsf{croll}\pi$ calculus, and proved in particular a context lemma, similar to that of the π-calculus, although the reversible machinery makes its proof more involved.

This work was done in collaboration with Inria teams FOCUS in Bologna and CELTIQUE in Rennes, and as part of the ANR REVER project.

Specfun Team

6. New Results

6.1. Creative telescoping for bivariate hyperexponential functions

In [8], we gave a new algorithm for the symbolic integration of bivariate hyperexponential functions, which outperforms state-of-the-art implementations like Maple's function *DEtools[Zeilberger]*. The approach was to extend Hermite's reduction for rational functions and the Hermite-like reduction for hyperexponential functions in a suitable way. A key feature of the algorithm is that it can avoid the costly computation of certificates.

6.2. Creative telescoping for rational functions

In [10] we described a precise and elementary algorithmic version of the Griffiths–Dwork method for the creative telescoping of rational functions. This leads to bounds on the order and degree of the coefficients of the differential equation, and to the first complexity result which is single exponential in the number of variables. One of the important features of the algorithm is that it does not need to compute certificates. The approach is vindicated by a prototype implementation.

6.3. Complexity of the uncoupling of linear functional systems

Uncoupling algorithms transform a linear differential system of first order into one or several scalar differential equations. We examined in [9] two approaches to uncoupling: the cyclic-vector method (*CVM*) and the Danilevski-Barkatou-Zürcher algorithm (*DBZ*). We gave tight size bounds on the scalar equations produced by *CVM*, and designed a fast variant of *CVM* whose complexity is quasi-optimal with respect to the output size. We exhibited a strong structural link between *CVM* and *DBZ* enabling to show that, in the generic case, *DBZ* has polynomial complexity and that it produces a single equation, strongly related to the output of *CVM*. We proved that algorithm *CVM* is faster than *DBZ* by almost two orders of magnitude, and provided experimental results that validate the theoretical complexity analyses.

6.4. Computation of integrals related to the Ising model

We showed in [2] that the n-fold integrals of the magnetic susceptibility of the Ising model, as well as various other n-fold integrals of the "Ising class", or n-fold integrals from enumerative combinatorics, like lattice Green functions, correspond to a distinguished class of functions generalising algebraic functions: they are actually diagonals of rational functions. This algebraic structure explains many remarkable properties of the integrals of the Ising class.

6.5. Non-D-finite excursions in the quarter plane

The number of excursions (finite paths starting and ending at the origin) having a given number of steps and obeying various geometric constraints is a classical topic of combinatorics and probability theory. We proved in [3] that the sequence of numbers of excursions in the quarter plane corresponding to a nonsingular step set $S \subseteq \{0, \pm 1\}^2$ with infinite group does not satisfy any nontrivial linear recurrence with polynomial coefficients. Accordingly, in those cases, the trivariate generating function of the numbers of walks with given length and prescribed ending point is not D-finite. This solves an open problem in the field of lattice path combinatorics.

6.6. A human proof of Gessel's lattice path conjecture

Gessel walks are planar walks confined to the positive quarter plane, that move by unit steps in any of the following directions: West, North-East, East, and South-West. In 2001, Ira Gessel conjectured a closed-form expression for the number of Gessel walks of a given length starting and ending at the origin. In 2008, Kauers, Koutschan, and Zeilberger gave a computer-aided proof of this conjecture. The same year, Bostan and Kauers showed, using again computer algebra tools, that the trivariate generating function of Gessel walks is algebraic. We proposed in [15] the first "human proofs" of these results. They are derived from a new expression for the generating function of Gessel walks.

6.7. Efficient algorithms for rational first integrals

We presented in [14] fast algorithms for computing rational first integrals with bounded degree of a planar polynomial vector field. Our approach is inspired by an idea of Ferragut and Giacomini. We improve upon their work by proving that rational first integrals can be computed via systems of linear equations instead of systems of quadratic equations. This leads to a probabilistic algorithm with arithmetic complexity $\tilde{\mathrm{O}}(N^{2\omega})$ and to a deterministic algorithm solving the problem in $\tilde{\mathrm{O}}(d^2 N^{2\omega+1})$ arithmetic operations, where N denotes the given bound for the degree of the rational first integral, and where $d \leq N$ is the degree of the vector field, and ω the exponent of linear algebra. By comparison, the best previous algorithm uses at least $d^{\omega+1} N^{4\omega+4}$ arithmetic operations. The new algorithms are very efficient in practice.

6.8. Reactive document checking in Coq

In an effort to improve the reactivity of Coq, the way it processes and checks a single document has been completely redesigned [7]. The current development version is able to reschedule the tasks to be performed in order to minimize the time required to give interactive feedback to the user. On typical documents taken from the formal proof of the Odd Order Theorem, the worst reaction time of the tool dropped from 5 minutes to 9 seconds. This improvement will be part of the next stable release of the Coq system.

6.9. Efficient normalization of ring/field expressions in Coq

The implementation of Coq's proof commands for manipulation of ring/field expressions has been improved in response to the demand for better efficiency that emerged in the formalization of Apéry's irrationality proof of $\zeta(3)$. The data structure used for the abstract syntax tree of ring/field expressions has been refined to enable a more efficient and more precise interpretation into concrete ring/field expressions. Moreover the collection of non-nullity conditions for denominators in a field expressions has been speeded up, making the type-checking time of a field normalization proof not be dominated by this collecting phase.

6.10. Documentation of Coq's canonical structures

The device employed to model a hierarchy of algebraic structures with overloaded notations in Coq has been documented in [6] and in the user manual of the tool.

6.11. Maintenance and development of the SSReflect extension for Coq and its user manual

The Small Scale Reflection extension of Coq has been maintained together with its user manual. Some new linguistic constructs to model non-structural reasoning and to enable the user to better factor out repeated arguments have been developed and documented. Some language constructs have been made compatible with the type-classes mechanism offered by Coq. The release of version 1.5 has been prepared.

6.12. Efficient proof-search techniques in sequent calculus

We have proposed in [11] a sequent calculus which is focussed, polarized, and parameterized by an abstract notion of theory. This new combination of features aims at proposing a framework which is adapted to the simulation in sequent calculus of efficient, general-purpose decision procedures (tableaux methods, satisfiability, ...) that can interact with theory-specific decision procedures (for linear arithmetics, arrays, ...). In particular we propose a tight simulation of the Davis–Putnam–Logemann–Loveland algorithm modulo theory, and show how to simulate some advanced optimizations that are crucial to realistic implementations of SMT solvers.

6.13. A formal proof of the irrationality of $\zeta(3)$

We have obtained a formal proof, machine-checked by the Coq proof assistant, of the irrationality of the constant $\zeta(3)$, under the single assumption of the asymptotic behavior of the least common multiple of the first n natural numbers. The core of this formal proof is based on (untrusted) computer-algebra calculations performed outside the proof assistant with the Algolib Maple library. Then, we verify formally and a posteriori the desired properties of the objects computed by Maple and complete the proof of irrationality.

6.14. Documentation of the Mathematical Components libraries

The approach to finite-group theory adopted in the libraries formalizing in Coq the proof of the Odd Order Theorem has been documented in [5].

STARS Project-Team

6. New Results

6.1. Introduction

This year Stars has proposed new algorithms related to its three main research axes : perception for activity recognition, semantic activity recognition and software engineering for activity recognition.

6.1.1. Perception for Activity Recognition

Participants: Julien Badie, Slawomir Bak, Vasanth Bathrinarayanan, Piotr Bilinski, François Brémond, Guillaume Charpiat, Duc Phu Chau, Etienne Corvée, Carolina Garate, Vaibhav Katiyar, Ratnesh Kumar, Srinidhi Mukanahallipatna, Marco San Biago, Silviu Serban, Malik Souded, Kartick Subramanian, Anh Tuan Nghiem, Monique Thonnat, Sofia Zaidenberg.

This year Stars has extended an algorithm for tuning automatically the parameters of the people tracking algorithm. We have evaluated the algorithm for re-identification of people through a camera network while taking into account a large variety of potential features together with practical constraints. We have designed several original algorithms for the recognition of short actions and validated its performance on several benchmarking databases (e.g. ADL). We have also worked on video segmentation and representation, with different approaches and applications.

More precisely, the new results for perception for activity recognition concern:

- Background Subtraction and People Detection in Videos (6.2)
- Tracking and Video Representation (6.3)
- Video segmentation with shape constraint (6.4)
- Articulating motion (6.5)
- Lossless image compression (6.6)
- People detection using RGB-D cameras (6.7)
- Online Tracking Parameter Adaptation based on Evaluation (6.8)
- People Detection, Tracking and Re-identification Through a Video Camera Network (6.9)
- People Retrieval in a Network of Cameras (6.10)
- Global Tracker : an Online Evaluation Framework to Improve Tracking Quality (6.11)
- Human Action Recognition in Videos (6.12)
- 3D Trajectories for Action Recognition Using Depth Sensors (6.13)
- Unsupervised Sudden Group Movement Discovery for Video Surveillance (6.14)
- Group Behavior Understanding (6.15)

6.1.2. Semantic Activity Recognition

Participants: Guillaume Charpiat, Serhan Cosar, Carlos -Fernando Crispim Junior, Hervé Falciani, Baptiste Fosty, Qiao Ma, Rim Romdhane.

During this period, we have thoroughly evaluated the generic event recognition algorithm using both sensors (RGB and RGBD video cameras). This algorithm has been tested on more than 70 videos of older adults performing 15 min of physical exercises and cognitive tasks. In Paris subway, we have been able to demonstrate the recognition in live of group behaviours. We have also been able to store the meta-data (e.g. people trajectories) generated from the processing of 8 video cameras, each of them lasting 2 or 3 days. From these meta-data, we have automatically discovered few hundreds of rare events, such as loitering, collapsing, ... to display on the screen of subway security operators.

Concerning semantic activity recognition, the contributions are :

- Evaluation of an Activity Monitoring System for Older People Using Fixed Cameras (6.16)
- A Framework for Activity Detection of Older People Using Multiple Sensors (6.17)
- Walking Speed Detection on a Treadmill using an RGB-D Camera (6.18)
- Serious Game for older adults with dementia (6.19)
- Unsupervised Activity Learning and Recognition (6.20)
- Extracting Statistical Information from Videos with Data Mining (6.21)

6.1.3. Software Engineering for Activity Recognition

Participants: François Brémond, Daniel Gaffé, Julien Gueytat, Sabine Moisan, Anh Tuan Nghiem, Annie Ressouche, Jean-Paul Rigault, Luis-Emiliano Sanchez.

This year Stars has continued the development of the SUP platform. This latter is the backbone of the team experiments to implement the new algorithms. We continue to improve our meta-modelling approach to support the development of video surveillance applications based on SUP. This year we have focused on metrics to drive dynamic architecture changes and on component management. We continue the development of a scenario analysis module (SAM) relying on formal methods to support activity recognition in SUP platform. We improve the CLEM toolkit and we rely on it to build SAM. Finally, we are improving the way we perform adaptation in the definition of a multiple services for device adaptive platform for scenario recognition.

The contributions for this research axis are:

- SUP (6.22)
- Model-Driven Engineering for Activity Recognition (6.23)
- Scenario Analysis Module (6.24)
- The Clem Workflow (6.25)
- Multiple Services for Device Adaptive Platform for Scenario Recognition (6.26)

6.2. Background Subtraction and People Detection in Videos

Participants: Vasanth Bathrinarayanan, Srinidhi Mukanahallipatna, Silviu Serban, François Brémond.

Keywords: Background Subtraction, People detection, Automatic parameter selection for algorithm

Background Subtraction Background subtraction is a vital real time low-level algorithm, which differentiates foreground and background objects in a video. We have thoroughly evaluated our Extended Gaussian Mixture model containing a shadows-removal algorithm, which performs better than other state of the art methods. Figure 10 shows the comparison of 13 background subtraction algorithms results on a challenging railway station monitoring video dataset from Project CENTAUR, which includes illumination change, shadows, occlusion and moving trains. Our algorithms performs the best in terms of result and with good processing speed too. Figure 11 is an example of our background subtraction algorithm's output on an indoor sequence of a surveillance footage from the Project SUPPORT.

Ongoing research include automatic parameter selection for this algorithm based on some learnt context. Since tuning the parameters is a daunting task for a non-experienced person, we try to learn some context information in a video like occlusion, contrast variation, density of foreground, texture etc. and map them to appropriate parameters of segmentation algorithm. Thus designing a controller to automatically adapt parameters of a algorithm as the scene context changes over time.

People Detection

A new robust real-time person detection system was proposed [45] , which aims to serve as solid foundation for developing solutions at an elevated level of reliability. Our belief is that clever handling of input data correlated with efficacious training algorithms are key for obtaining top performance. A comprehensive training method on very large training database and based on random sampling that compiles optimal classifiers with minimal bias and overfit rate is used. Building upon recent advances in multi-scale feature computations, our approach attains state-of-the-art accuracy while running at high frame rate.

../../../../projets/stars/IMG/bgres1.png ../../../../projets/stars/IMG/bgres2.png

Figure 10. Background Subtraction result on confidential dataset from CENTAUR project

../../../../projets/stars/IMG/bg-support.png

Figure 11. Background Subtraction result on a video to count number of people walking through the door after using their badge inside the terminal area (Project SUPPORT) - Autonomous Monitoring for Securing European Ports

Our method combines detection techniques that greatly reduce computational time without compromising accuracy. We use efficient LBP and MCT features which we compute on integral images for optimal retrieval of rectangular region intensity and nominal scaling error. AdaBoost is used to create cascading classifiers with significantly reduced detection time. We further refine detection speed by using the soft cascades approach and by transferring all important computation from the detection stage to the training stage. Figure 12 shows some output samples from various datasets which it was tested on.

../../../../projets/stars/IMG/detection1.png ../../../../projets/stars/IMG/detection2.png

../../../../projets/stars/IMG/detection3.png ../../../../projets/stars/IMG/detection4.png

Figure 12. Detection output from PETS (left-top), VANAHEIM (right-top), Hospital (left-bottom) and ETISEO (right-bottom)

6.3. Tracking and Video Representation

Participants: Ratnesh Kumar, Guillaume Charpiat, Monique Thonnat.

keywords: Fibers, Graph Partitioning, Message Passing, Iterative Conditional Modes, Video Segmentation, Video Inpainting

Multiple Object Tracking The objective is to find trajectories of objects (belonging to a particular category) in a video. To find possible occupancy locations, an object detector is applied to all frames of a video, yielding bounding boxes. Detectors are not perfect and may provide false detections; they may also miss objects sometimes. We build a graph of all detections, and aim at partitioning the graph into object trajectories. Edges in the graph encode factors between detections, based on the following :

- Number of common point tracks between bounding boxes (the tracks are obtained from an optical-flow-based point tracker)
- Global appearance similarity (based on the pixel colors inside the bounding boxes)
- Trajectory straightness : for three bounding boxes at different frames, we compute the Laplacian (centered at the middle frame) of the centroids of the boxes.
- Repulsive constraint : Two detections in a same frame cannot belong to the same trajectory.

We compute the partitions by using sequential tree re-weighted message passing (TRW-S). To avoid local minima, we use a label flipper motivated from the Iterative Conditional Modes algorithm.
We apply our approach to typical surveillance videos where object of interest are humans. Comparative quantitative results can be seen in Tables 1 and 2 for two videos. The evaluation metrics considered are : Recall, Precision, Average False Alarms Per Frame (FAF), Number of Groundtruth Trajectories (GT), Number of Mostly Tracked Trajectories, Number of Fragments (Frag), Number of Identity Switches (IDS), Multiple Object Tracking Accuracy (MOTA) and Multiple Object Tracking Precision (MOTP).

This work has been submitted to CVPR' 14.

Table 1. Towncenter Video Output

Method	MOTA	MOTP	Detector
[59] (450-750)	56.8	79.6	HOG
Ours (450-750)	53.5	69.1	HOG

Table 2. Comparison with recent proposed approaches on PETS S2L1 Video

Method	Recall	Precision	FAF	GT	MT	Frag	IDS
[77]	96.9	94.1	0.36	19	18	15	22
Ours	95.4	93.4	**0.28**	19	18	42	**13**

Video Representation We continued our work from the previous year on Fiber-Based Video Representation. During this year we focused on obtaining competitive results with the state-of-the-art (Figure 13).

The usefulness of our novel representation is demonstrated by a simple video inpainting task. Here a user input of only 7 clicks is required to remove the dancing girl disturbing the news reporter (Figure 14).

This work has been accepted for publication next year [41].

6.4. Video segmentation with shape constraint

Participant: Guillaume Charpiat.

keywords: video segmentation, graph-cut, shape growth, shape statistics, shape prior, dynamic time warping

6.4.1. Video segmentation with growth constraint

This is joint work with Yuliya Tarabalka (Ayin Inria team) and Björn Menze (ETH Zurich, also MIT and collaborator of Asclepios Inria team).

../../../../projets/stars/IMG/marple8_originals.png ../../../../projets/stars/IMG/marple8_tm7_update

../../../../projets/stars/IMG/marple8_highest_hierarchy7.png ../../../../projets/stars/IMG/grundmann_vol.png

*Figure 13. **Top Row**: Left image displays a sequence as a volumetric display. Right image displays all fibers found, clustered at a particular hierarchy. **Bottom Row** : Left Image displays the highest level of the hierarchical clustering, with fiber extension. Right Image shows the result obtained from [71]. **Our result demonstrates better long term temporal coherency.***

Figure 14. Inpainting task. ***Left*** *: Original video (top) and xt slice (bottom) showing trajectories.* ***Right*** *: Our result. Clusters of fibers were computed and selected with only 7 mouse clicks to distinguish the disturbing girl from the reporter and background. The girl was removed and the hole was filled by extending the background fibers in time.*

Context : One of the important challenges in computer vision is the automatic segmentation of objects in videos. This task becomes more difficult when image sequences are subject to low signal-to-noise ratio or low contrast between intensities of neighboring structures in the image scene. Such challenging data is acquired routinely, for example in medical imaging or in satellite remote sensing. While individual frames could be analyzed independently, temporal coherence in image sequences provides crucial information to make the problem easier. In this work, we focus on segmenting shapes in image sequences which only grow or shrink in time, and on making use of this knowledge as a constraint to help the segmentation process.

Approach and applications : We had proposed last year an approach based on graph-cut (see Figure 15), able to obtain efficiently (linear time in the number of pixels in practice), for any given video, its globally-optimal segmentation satisfying the growth constraint. This year we applied this method to three different applications :

- forest fires in satellite images,
- organ development in medical imaging (brain tumor, in multimodal MRI 3D volumes),
- sea ice melting in satellite observation, with a shrinking constraint instead of growth (see Figure 16).

The results on the first application were published in IGARSS (International Geoscience and Remote Sensing Symposium) [48], while the last two applications and the theory were published in BMCV [47]. A journal paper is also currently under review. A science popularization article was also published [53]. Not related but also with the Ayin Inria team was published the last of a series of articles about optimizers for point process models [40], introducing graph-cuts in the multiple birth and death approach in order to detect numerous objects that should not overlap.

6.4.2. Video segmentation with statistical shape prior

This is joint work with Maximiliano Suster (leader of the Neural Circuits and Behaviour Group at Bergen University, Norway).

../../../../projets/stars/IMG/graphmulti3.png

Figure 15. The approach : segmenting all frames together with a single graph-cut, with growth or shrinkage constraint, instead of segmenting independently each frame.

../../../../projets/stars/IMG/icy.png

Figure 16. Example of a noisy, challenging sequence of melting sea ice, from satellite observations (first column). The frames were aligned beforehand, using inside/outside histograms of pixel intensities. The following columns show the results of varied approaches, ranging from frame-by-frame segmentation to techniques ensuring shape smoothness in time. Our approach (last column) has the right suitable prior (shape shrinkage) and thus performs the best.

Context : The zebrafish larva is a model organism widely used in biology to study genetics. Therefore, analyzing its behavior in video sequences is particularly important for this research field. For this, there is a need to segment the animal in the video, in order to estimate its speed, and also more precisely to extract its shape, in order to express for instance how much it is bent, how fast it bends, etc. However, as the animal is stimulated by the experimenter with a probe, the full zebrafish larva is not always visible because of occlusion.

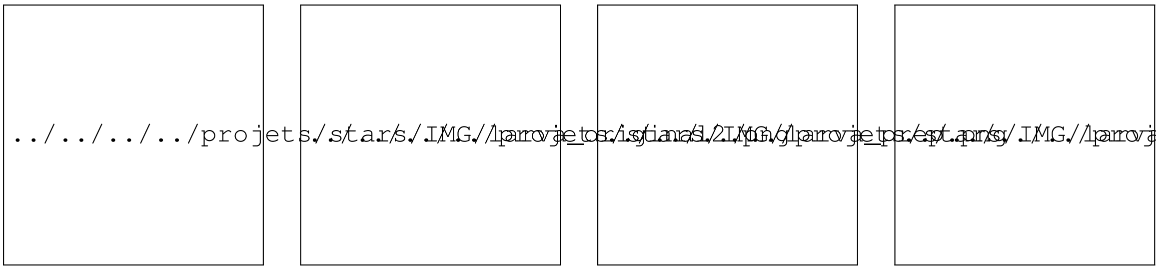


Figure 17. Example of a segmentation : initial image, processed image (based on video coherency), initialization of the active contour evolution, result. The total time spent per frame on average is reasonable for practical applications (magnitude order of 1 second).

Approach : We build a shape prior based on a training set of examples of non-occluded shapes, and use it to segment new images where the animal is occluded. This is however not straightforward.

- **Building a training set of shape deformations :** Given a set of training images containing non-occluded animals, we extract their contours via multiple robust thresholdings and morphomathematical operations. For each contour, we then estimate automatically the location of the tip of the tail. We then compute point-to-point correspondences between all contours, using a modified version of Dynamic Time Warping, as well as the approximate tip location information. This is done in a translation- and rotation-invariant way.
- **Building the shape prior :** Based on these matchings, the mean shape is computed, as well as modes of deformation with PCA.
- **Segmenting occluded images :** Images with occluded shapes are pre-processed in a similar way to non-occluded ones; however, the resulted segmentation does not contain only the parts of the larva but also the probe, which has potentially similar colors and location, and is moving. To identify the probe, whose shape depends on the video sequence, we make use of its rigidity and of temporal coherency. Then a segmentation criterion is designed to push an active contour towards the zones of interest (in a way that is robust to initialization), while keeping a shape which is feasible according to the shape prior.

Examples of data and results for a preliminary algorithm are shown in Figure 17 , with the associated shape prior shown in Figure 18 .

6.5. Articulating motion

Participant: Guillaume Charpiat.

keywords: shape evolution, metrics, gradient descent, Finsler gradient, Banach space, piecewise-rigidity, piecewise-similarity

This is joint work with Giacomo Nardi, Gabriel Peyré and François-Xavier Vialard (Ceremade, Paris-Dauphine University).

../../../../projets/stars/IMG/larva_modes.png

Figure 18. First deformation modes of the shape prior used in the segmentation above.

Context in optimization : A fact which is often ignored when optimizing a criterion with a gradient descent is that the gradient of a quantity depends on the metric chosen. In many domains, people choose by default the underlying L^2 metric, while it is not always relevant. Here we extend the set of metrics that can be considered, by building gradients for metrics that do not derive from inner products, with examples of metrics involving the L^1 norm, possibly of a derivative.

Mathematical foundations : This work introduces a novel steepest descent flow in Banach spaces. This extends previous works on generalized gradient descent, notably the work of Charpiat et al. [6], to the setting of Finsler metrics. Such a generalized gradient allows one to take into account a prior on deformations (e.g., piecewise rigid) in order to favor some specific evolutions. We define a Finsler gradient descent method to minimize a functional defined on a Banach space and we prove a convergence theorem for such a method. In particular, we show that the use of non-Hilbertian norms on Banach spaces is useful to study non-convex optimization problems where the geometry of the space might play a crucial role to avoid poor local minima.

Application to shape evolution : We performed some applications to the curve matching problem. In particular, we characterized piecewise-rigid deformations on the space of curves and we studied several models to perform piecewise-rigid evolutions (see Figure 19). We also studied piecewise-similar evolutions. Piecewise-rigidity intuitively corresponds to articulated motions, while piecewise-similarity further allows the elastic stretching of each articulated part independently. One practical consequence of our work is that any deformation to be applied to a shape can be easily and optimally transformed into an articulated deformation with few articulations, the number and location of the articulations being not known in advance. Surprisingly, this problem is actually convex.

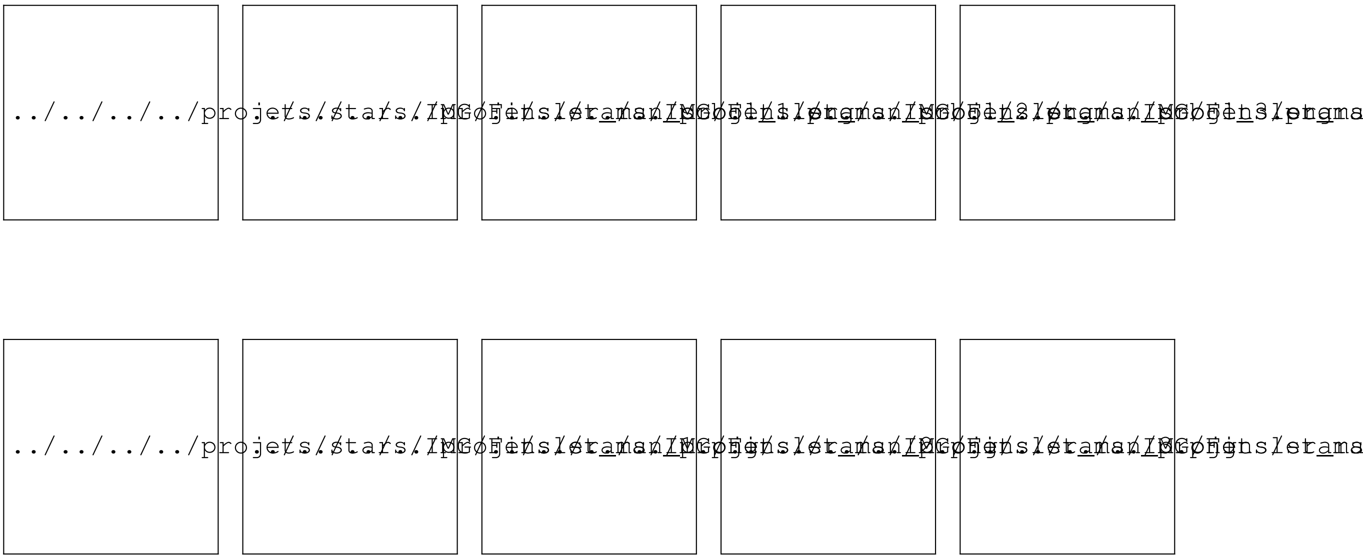


Figure 19. Example of use of the Finsler gradient for the piecewise-rigid evolution of curves. Given an initial shape S and a target shape T, as well as a shape dissimilarity measure $E(S) = Dissim(S,T)$, any classical gradient descent on $E(S)$ would draw the evolving shape S towards the target T. However the metric considered to compute the gradient changes the path followed. The top row is the evolution obtained with a Sobolev gradient H^1, which has the property of smoothing spatially the flow along the curve, to avoid irregular deformations. This is however not sufficient. The bottom row makes use of the Finsler gradient instead, with a metric favoring piecewise-rigid deformations.

An article was submitted to the journal Interfaces and Free Boundaries [52].

6.6. Lossless image compression

Participant: Guillaume Charpiat.

keywords: image compression, entropy coding, graph-cut

This is joint work with Yann Ollivier and Jamal Atif from the TAO Inria team.

Context : Understanding, modelling, predicting and compressing images are tightly linked, in that any good predictor can be turned into a good compressor via entropy coding (such as Huffman coding or arithmetic coding). Indeed, with such techniques, the more predictable an event E is, i.e. the higher its probability $p(E)$, the easier to compress it will be, with coding cost $-\log(p(E))$. Therefore we are interested in image compression, in order to build better models of images.

MDL approach : The state-of-the-art sequential prediction of time series based on the advice of various experts combines the different expert predictions, with weights depending on their individual past performance (cf. Gilles Stoltz and Peter Grünwald's work). This approach originates from the Minimum Description Length principle (MDL). This work was however designed for 1D data such as time series, and is not directly applicable to 2D data such as images. Consequently, our aim has been to adapt such an approach to the case of image compression, where time series are replaced with 2D series of pixel colors, and where experts are predictors of the color of a pixel given the colors of neighbors.

New method and results : This year, we have focused on lossless greyscale image compression, and proposed to encode any image with two maps, one storing the choice of the expert made for each pixel, and one storing the encoding of the intensity of each pixel according to its expert. In order to compress efficiently the first map, we ask the choices of experts to be coherent in space, and then encode the boundaries of the experts' areas. To find a suitable expert map, we optimize the total encoding cost explicitely, set as an energy minimization problem, solved with graph-cuts. An example of expert map obtained is shown in Figure 20 . Preliminary results with a hierarchical ordering scheme already compete with standard techniques in lossless compression (PNG, lossless JPEG2000, JPEG-LS).

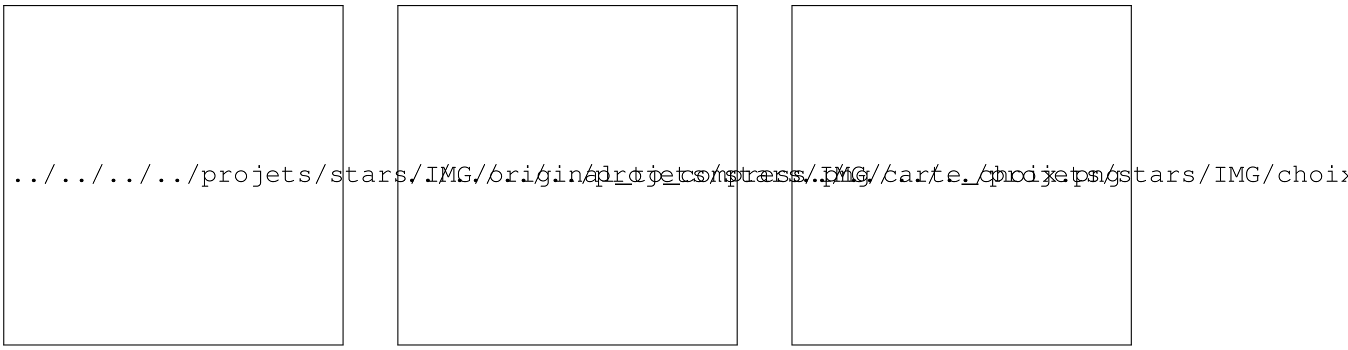


Figure 20. An image to encode; the map of the best experts, chosen independently for each pixel [each expert is represented by one color]; the expert map obtained with our approach enforcing spatial coherency in the expert choice.

6.7. People detection using RGB-D cameras

Participants: Anh-Tuan Nghiem, François Brémond.

keywords: people detection, HOG, RGB-D cameras

With the introduction of low cost RGB-D cameras like Kinect of Microsoft, video monitoring systems have another option for indoor monitoring beside conventional RGB cameras. Comparing with conventional RGB camera, reliable depth information from RGB-D cameras makes people detection easier. Besides that, constructors of RGB-D cameras also provide various libraries for people detection, skeleton detection or hand detection etc. However, perhaps due to high variance of depth measurement when objects are too far from the camera, these libraries only work when people are in the range of 0.5 to around 4.5 m from the cameras. Therefore, for our own video monitoring system, we construct our own people detection framework consisting of a background subtraction, a people classifier, a tracker and a noise removal component as illustrated in figure 21 .

../../../../projets/stars/IMG/overallHorizontal.jpg

Figure 21. The people detection framework

In this system, the background subtraction algorithm is designed specifically for depth data. Particularly, the algorithm employs temporal filters to detect noise related to imperfect depth measurement on some special surface.

The people classification part is the extension of the work in [79]. From the foreground region provided by the background subtraction algorithm, the classification first searches for people head and then extracts HOG like features (Histogram of Oriented Gradient on binary image) above the head and the shoulder. Finally, these features are classified by a SVM classifier to recognise people.

The tracker links detected foreground regions in the current frame with the ones from previous frames. By linking objects in different frames, the tracker provides useful history information to remove noise as well as to improve the sensitivity of the people classifier.

Finally, the noise removal algorithm uses the object history constructed by the tracker to remove two types of noise: noise detected by temporal filter at the background subtraction algorithm and noise from high variance of depth measurement on objects far from the camera. Figure 22 illustrates the performance of noise removal on the detection results.

../../../../projets/stars/IMG/noiseRemovalPerformance.jpg

Figure 22. The people detection framework

The overall performance of our people detection framework is comparable to the one provided by Primesense, the constructor of RGB-D camera Microsoft Kinect.

Currently, we are doing extensive evaluation of the framework and the results will be submitted to a conference in the near future.

6.8. Online Tracking Parameter Adaptation based on Evaluation

Participants: Duc Phu Chau, Julien Badie, Kartick Subramanian, François Brémond, Monique Thonnat.

Keywords: Object tracking, parameter tuning, online evaluation, machine learning

Several studies have been proposed for tracking mobile objects in videos [50]. For example we have proposed recently a new tracker which is based on co-inertia analysis (COIA) of object features [44]. However the parameter tuning is still a common issue for many trackers. In order to solve this problem, we propose an online parameter tuning process to adapt a tracking algorithm to various scene contexts. The proposed approach brings two contributions: (1) an online tracking evaluation, and (2) a method to adapt online tracking parameters to scene contexts.

In an offline training phase, this approach learns how to tune the tracker parameters to cope with different contexts. Different learning schemes (e.g. neural network-based) are proposed. A context database is created at the end of this phase to support the control process of the considered tracking algorithm. This database contains satisfactory parameter values of this tracker for various contexts.

In the online control phase, once the tracking quality is evaluated as not good enough, the proposed approach computes the current context and tunes the tracking parameters using the learned values.

The experimental results show that the proposed approach improves the performance of the tracking algorithm and outperforms recent state of the art trackers. Figure 23 shows the correct tracking results of four people while occlusions happen. Table 3 presents the tracking results of the proposed approach and of some recent trackers from the state of the art. The proposed controller increases significantly the performance of an appearance-based tracker [63]. We obtain the best MT value (i.e. mostly tracked trajectories) compared to state of the art trackers.

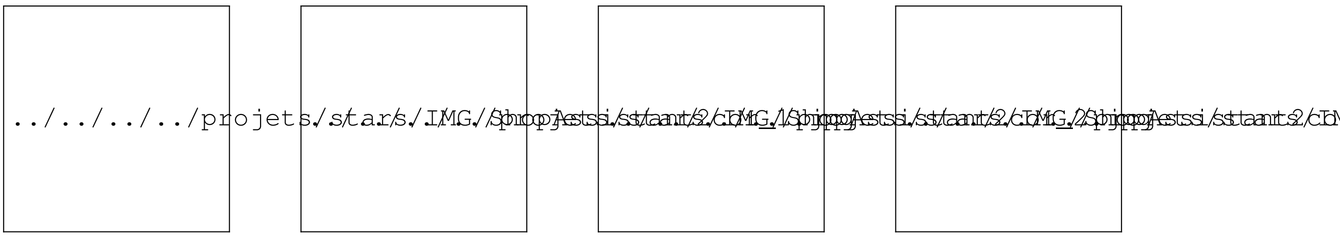


Figure 23. Tracking results of four people in the sequence ShopAssistant2cor (Caviar dataset) are correct, even when occlusions happen.

This work has been published in [33], [34].

6.9. People Detection, Tracking and Re-identification Through a Video Camera Network

Participants: Malik Souded, François Brémond.

keywords: People detection, Object tracking, People re-identification, Region covariance descriptors, SIFT descriptor, LogitBoost, Particle filters.

Table 3. Tracking results for the Caviar dataset. The proposed controller improves significantly the tracking performance. MT: Mostly tracked trajectories, higher is better. PT: Partially tracked trajectories. ML: Most lost trajectories, lower is better. The best values are printed in bold.

Approaches	MT (%)	PT (%)	ML (%)
Xing et al. [92]	84.3	12.1	3.6
Li et al. [76]	84.6	14.0	1.4
Kuo et al. [74]	84.6	14.7	**0.7**
D.P Chau et al. [63] without the proposed approach	78.3	16.0	5.7
D.P Chau et al. [63] with the proposed approach	**85.5**	9.2	5.3

This works aims at proposing a whole framework for people detection, tracking and re-identification through camera networks. Three main constraints have guided this work: high performances, real-time processing and genericity of the proposed methods (minimal human interaction/parametrization). This work is divided into three separate but dependent tasks:

6.9.1. People detection:

The proposed approach optimizes state-of-the-art methods [89], [93] which are based on training cascades of classifiers using the LogitBoost algorithm on region covariance descriptors. The optimization consists in clustering negative data before the training step, and speeds up both the training and detection processes while improving the detection performance. This approach has been published this year in [46]. The evaluation results and examples of detection are shown in Figures 24 and 25 .

6.9.2. Object tracking:

The proposed object tracker uses a state-of-the-art background subtraction algorithm to initialize objects to track, with a collaboration of the proposed people detector in the case of people tracking. The object modelling is performed using SIFT features, detected and selected in a particular manner. The tracking process is performed at two levels: SIFT features are tracked using a specific particle filter, then object tracking is deduced from the tracked SIFT features using the proposed data association framework. A fast occlusion management is also proposed to achieve the object tracking process. The evaluation results are shown in Figure 26 .

6.9.3. People re-identification:

A state-of-the-art method for people re-identification [67] is used as a baseline and its performance has been improved. A fast method for image alignment for multiple-shot case is proposed first. Then, texture information is added to the computed visual signatures. A method for people visible side classification is also proposed. Camera calibration information is used to filter candidate people who do not match spatio-temporal constraints. Finally, an adaptive feature weighting method according to visible side classification concludes the improvement contributions. The evaluation results are shown in Figure 27 .

This work has been published in [28].

6.10. People Retrieval in a Network of Cameras

Participants: Sławomir Bąk, Marco San Biago, Ratnesh Kumar, Vasanth Bathrinarayanan, François Brémond.

keywords: Brownian statistics, re-identification, retrieval

../../../../projets/stars/IMG/People_Detector_Results.png

Figure 24. People detector evaluation and comparison on Inria, DaimlerChrysler, Caltech and CAVIAR datasets.

../../../../projets/stars/IMG/People_Detection_Examples.png

Figure 25. Some examples of detection using the proposed people detector.

../../../../projets/stars/IMG/Object_Tracking_Results.png

Figure 26. Object tracking evaluation on: (a) CAVIAR dataset using MT, PT and ML metrics. (b) ETI-VS1-BE-18-C4 sequence from ETISEO dataset, using ETISEO metrics.

../../../../projets/stars/IMG/People_ReIdentification_Results_VIPeR_iLids119_CAVIAR4R

Figure 27. People re-identification evaluation on VIPeR (left), iLids-119 (middle) and CAVIAR4REID (right) datasets.

Task. Person re-identification (also known as multi-camera tracking) is defined as the process of determining whether a given individual has already appeared over a network of cameras. In most video surveillance scenarios, features such as face or iris are not available due to video low-resolution. Therefore a robust modeling of the global appearance of an individual (clothing) is necessary for re-identification. This problem is particularly hard due to significant appearance changes caused by variations in view angle, lighting conditions and different person pose. In this year, we focused on the two following aspects: new image descriptors and a design of a retrieval tool.

New image region descriptors. We have evaluated different image descriptors *w.r.t.* their recognition accuracy. As the covariance descriptor achieved the best results, we have employed this descriptor using different learning strategies to achieve the most accurate model for representing a human appearance [51]. We have also proposed a new descriptor based on recent advances in mathematical statistics related to Brownian motion [31]. This new descriptor outperforms the classical covariance in terms of matching accuracy and efficiency. We show that the proposed descriptor can capture richer characteristics than covariance, especially when fusing nonlinearly dependent features, which is often the case for images. The effectiveness of the approach is validated on three challenging vision tasks: object tracking & person re-identification [31] and pedestrian classification (the paper submitted to conference CVPR 2014). In all our experiments, we demonstrate competitive results while in person re-identification and tracking we significantly outperform the state-of-the-art.

New design of retrieval tool for a large network of cameras. Owing to the complexity of the re-identification problem, current state of the art approaches have relatively low retrieval accuracy, thus a fully automated system is still unattainable. However, we propose a retrieval tool [30], [29] that helps a human operator to solve the re-identification task (see Figure 28). This tool allows a human operator to browse images of people extracted from a network of cameras: to detect a person on one camera and to re-detect the same person few minutes later on another camera. The main stream is displayed on the left of the screen, while retrieval results are shown on the right. The results show lists of the most similar signatures extracted from each camera (green boxes indicate the correctly retrieved person). Below the main stream window a topology of the camera network is displayed. Detection and single camera tracking (see the main stream) are fully automatic. The human operator only needs to select a person of interest, thus producing retrieval results (right screen). The operator can easily see a preview of the retrieval results and can go directly to the original video content.

Perspectives. Currently, we are working not only on invariant image descriptors, which provide high recognition accuracy, but also on improving the alignment of the person pose, while matching appearance from cameras with significant difference in viewpoint. In addition to re-identification technology, we also work on designing an intuitive graphical interface, an important tool for the human operator analyzing retrieval results. Displaying retrieval results from a large camera network is still an issue, even after applying time-space constraints (the usage of topology of cameras).

../../../../projets/stars/IMG/Slawomir_fig.jpg

(a)

Figure 28. Re-identification tool

Acknowledgements
This work has been supported by PANORAMA and CENTAUR European projects.

6.11. Global Tracker : an Online Evaluation Framework to Improve Tracking Quality

Participants: Julien Badie, Slawomir Bak, Duc Phu Chau, François Brémond, Monique Thonnat.

keywords: online quality estimation, improving tracking results

This work addresses the problem of estimating the quality of a tracking algorithm during runtime and correcting the anomalies found. Evaluating and tuning a tracking algorithm generally requires multiple runs and a ground truth. The proposed online evaluation framework, called global tracker, overcome these limitations by proposing a three-steps algorithm to improve tracking results in any kind of situations (mono-camera, camera network, 3D camera, ...).

The first step aims at correcting small tracking errors (when detections in consecutive frames are missing from an object trajectory) by interpolating the detected object data.

The second steps aims at detecting and correcting any anomaly found in the output of the tracking algorithm. For each detected object on each frame, we compute three different sets of features : the features that are computed using only data from the object (eg. appearance, size, ...), the features that measure the level of interaction between two objects (eg. occlusion level, density) and the features that measure the level of interaction between the object and the environment (eg. occlusion with background element, entering or leaving zones, ...). By studying the evolution of the coefficients of variation of each feature, some anomalies can be detected. Then, by minimizing an energy function based on the object-only features, we correct the output of the tracking algorithm.

The final step uses re-acquisition and re-identification methods to merge detected objects representing the same real object. This algorithm aims at correcting ID change errors when an object leaves the scene and reappears in another camera or when the object reappears after a long-term occlusion. The method used here is a constrained clustering algorithm that create groups of detections representing the same real object.

Figure 29. An example of the global tracker goal : offline learning of occlusion zones to manage online occlusion risks and optimize object trajectory.

This approach has been tested on several datasets (PETS 2009, Caviar, TUD-Stadtmitte). The results show that the global tracker, even associated with a tracking algorithm that does not have good results, can perform nearly as well as the state of the art and even exactly as well when a good tracker is used. On the Caviar dataset, the global tracker is even able to slightly outperform the result of the state of the art.

A part of this approach is described in an article published in AVSS 2013 [33]. This article shows how a tracking algorithm can use the global tracker results to tune its parameters and improve its results. This work was also used to improve the tracking results in 2 papers [38] [54].

Table 4. Comparison of the tracking results using CLEAR metrics on the sequence S2.L1 of the PETS 2009 dataset with and without the global tracker

Method	MOTA	MOTP	$\overline{M}$
Berclaz et al. [60]	0.80	0.58	0.69
Shitrit et al. [58]	0.81	0.58	0.70
Henriques et al. [72]	*0.85*	0.69	0.77
Chau et al. [33] without global tracker	0.62	0.63	0.63
Chau et al. [33] with global tracker	*0.85*	*0.71*	*0.78*

6.12. Human Action Recognition in Videos

Participants: Piotr Bilinski, Etienne Corvée, Slawomir Bak, François Brémond.

keywords: action recognition, tracklets, head detection, relative tracklets, bag-of-words.

In this work we address the problem of recognizing human actions in video sequences for home care applications.

Recent studies have shown that approaches which use a bag-of-words representation reach high action recognition accuracy. Unfortunately, these approaches have problems to discriminate similar actions, ignoring spatial information of features.

We propose a feature representation for action recognition based on dense point tracklets, head position estimation, and a dynamic coordinate system. Our main idea is that action recognition ought to be performed using a dynamic coordinate system corresponding to an object of interest. Therefore, we introduce a relative tracklet descriptor based on relative positions of a tracklet according to the central point of our dynamic coordinate system. As a center of our dynamic coordinate system, we choose the head position, providing description invariant to camera viewpoint changes. We use the bag-of-words approach to represent a video sequence and we capture global distribution of tracklets and relative tracklet descriptors over a video sequence. The proposed descriptors introduce spatial information to the bag-of-words model and help to distinguish similar features detected at different positions (*e.g.* to distinguish similar features appearing on hands and feet). Then we apply the Support Vector Machines with exponential chi-squared kernel to classify videos and recognize actions.

We report experimental results on three action recognition datasets (publicly available KTH and ADL datasets, and our locally collected dataset). Our locally collected dataset has been created in cooperation with the CHU Nice Hospital. It contains people performing daily living activities such as: standing up, sitting down, walking, reading a magazine, *etc*. Consistently, experiments show that our representation enhances the discriminative power of the tracklet descriptors and the bag-of-words model, and improves action recognition performance.

Sample video frames with extracted tracklets and estimated head positions are presented in Figure 30 .

This work has been published in [32].

6.12.1. Acknowledgments

This work was supported by the Région Provence-Alpes-Côte d'Azur. However, the views and opinions expressed herein do not necessarily reflect those of the financing institution

6.13. 3D Trajectories for Action Recognition Using Depth Sensors

Participants: Michal Koperski, Piotr Bilinski, François Brémond.

keywords: action recognition, computer vision, machine learning, 3D sensors

The goal of our work is to extend recently published approaches ([61], [62], [32], [90]) for Human Action Recognition to take advantage of the depth information from 3D sensors.

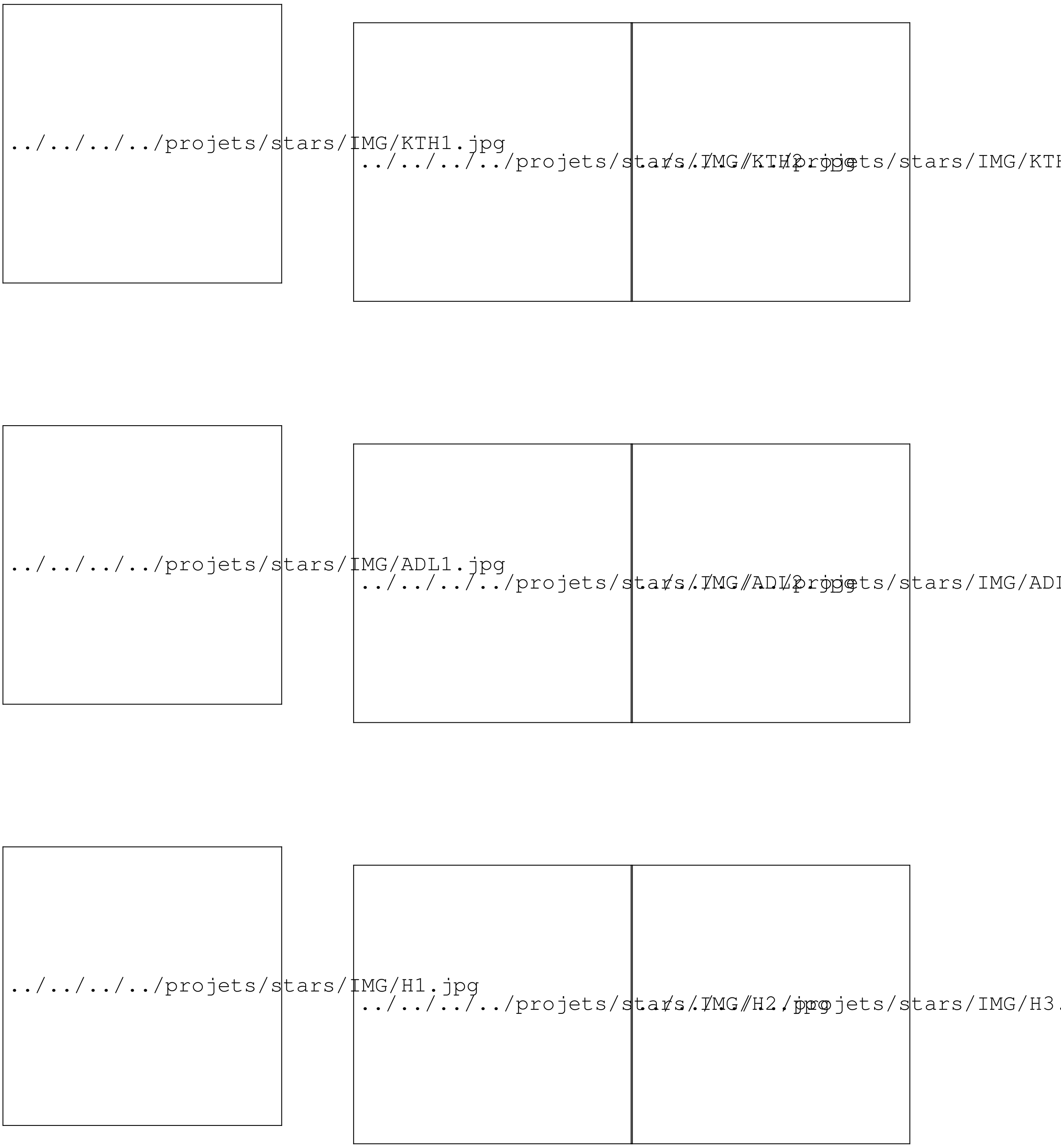


Figure 30. Sample video frames with extracted tracklets and estimated head positions for the KTH (first row), ADL (second row) and our locally collected dataset (third row).

We propose to add depth information to trajectory based algorithms ([32], [90]). Currently mentioned algorithms compute trajectories by sampling video frames and then tracking points of interest - creating the trajectory. Our contribution is to create even more discriminative features by adding depth information to previously detected trajectories. In our work we propose methods to deal with noise and missing measurements in depth information map. Such computed 3D trajectories, combined with other appearance features (HOG, HOF), are subject to a Bag of Words model and SVM classifier.

../../../../projets/stars/IMG/MSR_DA3D_example.jpg

Figure 31. Visualization of MSR Dailiy Activty 3D data set. Left : video input frame; Middle : frame with detected trajectories (red = static points, green = detected trajectories); Right : corresponding depth map.

The evaluation of our method was conducted on the ”Microsoft Daily Activity3D” data set [91] which consist of 16 actions (drink, eat, read book, call cellphone, write on a paper, use laptop etc.) performed by 10 subjects.

The experiments showed that adding depth information to Dense Trajectories descriptor [90] gave gain in efficiency 57.72% to 64.12%. The mentioned work is going to be submitted in December 2013.

6.14. Unsupervised Sudden Group Movement Discovery for Video Surveillance

Participants: Sofia Zaidenberg, Piotr Bilinski, François Brémond.

keywords: Sudden Group Movement Discovery, Video Surveillance.

In this work we address the problem of discovering "sudden" movements in video surveillance videos. We propose an unsupervised approach which automatically detects quick motions in a video, corresponding to any action. A set of possible actions is not required and the proposed method successfully detects potentially alarm-raising actions without training or camera calibration. Moreover the system uses a group detection and event recognition framework to relate detected sudden movements and groups of people, and provides a semantic interpretation of the scene. We have tested our approach on a dataset of nearly eight hours of videos recorded from two cameras in the Parisian subway for a European Project. For evaluation we annotated one hour of sequences containing 50 sudden movements. Our system, if parametrized to a high sensitivity, detects 100% of what the annotator considered as sudden potentially dangerous events, with a false positive rate of 21.2%. Setting the sensitivity to lower values we decrease the false positive rate to only 5.3% but we also decrease the success rate to 76%. An example of an unusual sudden movement annotated by a human and detected by our approach is presented in Figure 32 . This work has been published in [49].

6.14.1. Acknowledgments

This work was supported by the Région Provence-Alpes-Côte d'Azur and by the European Community's Seventh Framework Programme FP7/2007-2013 - Challenge 2 - Cognitive Systems, Interaction, Robotics - under grant agreement number: 248907-VANAHEIM. However, the views and opinions expressed herein do not necessarily reflect those of the financing institution.

Figure 32. Example of an unusual sudden movement detected by our approach.

6.15. Group Behavior Understanding

Participants: Carolina Gárate, Sofia Zaidenberg, Julien Badie, François Brémond.

The goal is to recognize group behavior from videos. Dangerous and criminal behaviors are mostly observed within groups of people. The idea is to detect potentially dangerous situations while they are happening in the context of underground railway station security.
keywords: group tracking, scene understanding, group behavior recognition, video surveillance, event detection.

This research work considers a process consisting of 5 consecutive steps for video processing. The steps are: 1) segmentation, 2) blob detection, 3) physical objects tracking, 4) group tracking and 5) behavior recognition. Here, we are focussing on the last two phases: group tracking and behavior recognition.

The group tracking approach characterizes a group through three features: the average of the intra-object distance, the average standard deviations of speed and direction. The input for this algorithm is a set of trajectories for the physical objects (output of the stage 3: physical objects tracking) tracked by the algorithm described in [64]. The trajectories are processed using Mean-Shift clustering to create more reliable groups, see Figure 33 .

The behavior recognition approach identifies 2 steps: knowledge modeling and the event recognition algorithm. The ontology is implemented with the ScReK declarative language [94]. The grammar describes the objects and events using the extended BNF (Backus Naur Form) representation.

We process large amounts of long video surveillance data from Paris and Turin underground railway station to perform statistical analysis. This analysis automatically brings forward data about the usage of the station and the various behaviors of groups for different hours of the day. We present the results and interpretation of one month of processed data from a video surveillance camera in Turin subway.

One of the measures obtained in the experimentation is the agitation level which is represented by the variation of the size of the bounding box of a group. We consider 3 categories from *no agitation* ("Calm_Group", having a bounding box with stable size) to *little agitation* ("Active_Group") to *high agitation* ("Lively_Group", the bounding box's size varies a lot, meaning that group members move around more often). Figure 34 shows that most of the time, this middle category predominates. Groups are neither too calm, nor too agitated. Moreover, it is more common for a group to be lively rather than calm. The rest of the results obtained were presented in [54].

6.16. Evaluation of an Activity Monitoring System for Older People Using Fixed Cameras

Participants: Carlos F. Crispim-Junior, Baptiste Fosty, Vasanth Bathrinarayanan, Salma Zouaoui-Elloumi, Monique Thonnat, François Brémond.

keywords: 2D-RGB cameras, RGB-D cameras, model-based activity recognition, older people

We have continued the evaluation of our model-based algorithm for complex activity recognition, now extending it to a larger dataset containing 38 older people participants undertaking instrumental activities of daily living (IADL) during 15 minutes (570 min. in total). The recordings have taken place in the observation room of the Memory Center of Nice hospital. Figure 35 presents the algorithm performance based on data obtained from a 2D-RGB video camera. A summary of the recognized activities (e.g., duration, frequency) is produced at the end of the event recognition task to be provided to doctors as a basis for the assessment of patient performance on IADL. This approach description and the evaluation results are published in 2013 AVSS Conference (see details in [36]). Figure 36 illustrates an example of a patient being monitored. Blue dots illustrates previous positions of the person in the scene.

The proposed approach has been also evaluated using a RGB-D camera, as this camera increases the robustness of the monitoring system against environment illumination changes and also eases the deployment of the system by providing real 3-D information on the scene. The evaluation of the RGB-D-based activity monitoring system has been published in [38]. A live demonstration of this system has also been presented and applied in the scope of Dem@care project (a FP7 project devoted to multi-sensor older people monitoring) in the exhibition held in November 2013 in conjunction with the 2013 edition of the ICT (Information Communication Technologies) Conference in Vilnius - Lithuania.

6.17. A Framework for Activity Detection of Older People Using Multiple Sensors

Participants: Carlos F. Crispim-Junior, Qiao Ma, Baptiste Fosty, François Brémond, Monique Thonnat.

../../../../projets/stars/IMG/fig_14974.jpg

Figure 33. Illustration of a still group: in the right entrance zone a group of subway security staff is standing still.

../../../../projets/stars/IMG/plotData-ter12.png

Figure 34. Comparison of speed of groups and variation of group's bounding boxes during the day.

../../../../projets/stars/IMG/resulttable.png

Figure 35. Evaluation of Algorithm performance on 38 patient video recordings using a RGB camera

../../../../projets/stars/IMG/trackedpatient.png

Figure 36. Example of a patient been monitored by the described system

keywords: model-based activity recognition, multi-sensor, Dempster-Shafer, Evidence Theory, older people.

We have extended our framework for multi-sensor activity detection by proposing a probabilistic approach for mutually exclusive activity conflict scenario. The proposed approach pre-learned a coefficient of reliability of each sensor with respect to each activity. The combination of the activities detected by multiple sensors is performed using the Dempster-Shafer evidence theory with an adapted combination rule based on runtime data from the sensor and the pre-computed coefficients of reliability. The combination of activities detected by multiple sensors can be performed in an iterative fashion taking into account several sensor contributions (see Fig.37). Tab. 5 presents the early results of the proposed probabilistic method at combining activities detected using RGB and RGB-D cameras. Results are presented individually for each camera and for the proposed approach.

../../../../projets/stars/IMG/fusionapproach.png

Figure 37. Probabilistic Approach Combination Scheme

The proposed fusion scheme performs better than the camera individual process in most of cases even in the presence of noise (see the RGB-D individual result of sensitivity for standing posture detection). The complete evaluation of the proposed approach is published in [43]. The developed probabilistic approach is now integrated into our previously developed framework for multi-sensor activity modeling. The new framework version allows experts to precisely define the sensors which will be used to detect each activity, or to automatically (and then probabilistic) combine multiple instances of (conflicting) activities detected by different sensors. The new framework proposal for multi-sensor activity detection has been published in [37].

6.18. Walking Speed Detection on a Treadmill using an RGB-D Camera

Participants: Baptiste Fosty, François Brémond.

Table 5. Performance of the proposed probabilistic approach on posture detection

Posture	Standing		Sitting	
Sensor	Precision	Sensitivity	Precision	Sensitivity
RGB	84.29	69.41	79.82	91.58
RGB-D	100.00	36.47	86.92	97.89
Fusion	82.35	91.30	91.04	95.31

keywords: RGB-D camera analysis, older people, serious games

Within the context of the Az@Game project, we have studied the potential of the RGB-D camera (Red Green Blue + Depth) for the control of a serious game dedicated to older people suffering from Alzheimer disease. Within this game, the patient is invited to perform some physical and cognitive tasks (walking on a treadmill at different speeds, performing gestures to control his/her character in the game, managing money) in order to assess the evolution of the disease, to stimulate them and improve their abilities. In this context, one of our goal is the computation of the walking speed of a person on a treadmill.

The proposed solution is divided into three distinct steps :

- people detection and tracking using a background subtraction algorithm.
- feet detection : based on the cloud of 3D points of the person, and more particularly on the lower body part, the axis of each leg is computed. The corresponding foot is then defined as the projection of the lowest point of the leg on the leg axis (see Fig. 38 , right picture).
- speed computation : computed from the successive positions of the feet, more precisely from the distances between the feet (see Fig. 38 , left graph representing the distance between the feet). Each time this distance reaches a local maximum (corresponding to each step), the current speed is computed as the maximum over the time since the last step. The speed is then averaged with the previous computed speed to smooth the values (see Fig. 38 , middle graph representing the speed in function of time).

Concerning the results, the first experimentation of the algorithm shows that, qualitatively, the computed walking speed is proportional to the real speed. Future work will focus on experimenting the proposed system on a larger scale (different people, location, etc.) in order to validate the approach. We will also focus on trying to detect arm gestures to have more control on the serious game.
About the Dem@Care project and the previous work on the activity recognition system developed to extract automatically and objectively evidences of early symptoms of Alzheimer's disease for older people, this contribution has been published at ASROB 2013, Workshop on Assistance and Service Robotics in a Human Environment (see [38]).

6.19. Serious Game for older adults with dementia

Participants: Minh Khue Phan Tran, François Brémond.

keywords: human-machine interaction, serious games, Alzheimer, activity recognition

Serious Games is carried out within the framework of the *Az@GAME project*. This project is to create games offering patient-oriented scenarios so as to measure their health progress, improve their physical fitness, stimulate their cognitive abilities, and help maintain their social skills. The main objective is to design a system interacting with older adults suffering from Alzheimer's or Alzheimer's related diseases. The three challenges in designing the system are:

- **perception's precision :** how does the system choose the *"best moments"* to interact with a patient ?
- **attractive-visualization :** how does the system make the patients comfortable?
- **ease of interaction :** how can it optimize the interaction with the patients ? In what way ?

../../../../projets/stars/IMG/walkingSpeed.jpg

Figure 38. Walking speed computation on a treadmill. The left graph is the representation of the distance between the feet as a function of time. The middle graph is the representation of the speed of the person as a function of time. The right picture is the RGB-D camera view with the people detection and current speed. The two yellow circles show the positions of the detected feet.

The first prototype is under development. The system consists of two parts: Recognition and Interaction . Each component requires a 3D camera (Microsoft Kinect for the recognition component and Asus Xtion Pro Live Camera for the interaction component). The recognition part consists in observing the scene and deciding the best time to interact with a patient via the Asus camera. It uses the SUP framework. Afterwards, the interactive system tries to engage the patient via an interface and through Microsoft Kinect, the patient can interact with the interface using voice or gesture. The interface is designed with the Unity 3D game engine.

../../../../projets/stars/IMG/image-front.jpg

Figure 39. Example of a game and its gesture interface.

The first experiment will be conducted in the coffee area. The aim is to test the functionality of the system and measure its accuracy and effectiveness. The system will observe the scene and invite people who are getting coffee or taking a break to play the game. Depending on the interaction with the person, the system will offer different scenarios. Videos will be recorded, with the consent of the subject, in order to evaluate the effectiveness of system. The recorded videos and meta-data provided by SUP will be evaluated to determine the accuracy of the system.

6.20. Unsupervised Activity Learning and Recognition

Participants: Serhan Cosar, Salma Zouaoui-Elloumi, François Brémond.

keywords: Unsupervised activity learning, hierarchical activity models, monitoring older people activities

The aim of this work is to monitor older people activities at hospital or at home environment in an unsupervised manner. We have extended the work in [81] that was initially based on user interface to label activities and proposed a new strongly unsupervised framework. It enables the discovery, modeling, and recognition of activities without user interaction. One advantage of this approach is that the framework learns individual behavioral patterns in unstructured scenes without restraining people to act based on a manually pre-defined model. The Figure 40 -(a) presents the off-line learning steps of this framework. It takes as input a set of videos pre-processed to obtain trajectory information of people in the scene. Using the trajectory information (global position and pixel tracklets of body parts) of each person, zones of interest, where the person performs an activity, are learned. As in [81], we obtain three levels of zones using k-means clustering for different k values. The obtained zones are used to create different levels of events from the coarser to the finer ones. Based on the three levels of events, a hierarchical model of activities is learned to represent each action (Figure 40 -(a). For each new video, an on-line recognition process is performed by using the previously learned zones and models of activities (Figure 40 -(b)).

We have evaluated the performance of the unsupervised algorithm for RGB-D and 2D camera using 8 videos and 10 videos, respectively. Half of the videos are used for learning zones and models of activities. Videos are recorded in CHU Nice hospital while older people are visiting their doctors and include the following actions: “talking on the phone”, “preparing drugs”, “sitting at the table”, “preparing tea”, “looking at the bus map”, “watching TV” and “paying bill”. The trajectory information for 2D camera is obtained using the method in [81]. For RGB-D camera, we have used the person detection algorithm in [79] and tracking algorithm in [33]. The results obtained for both cameras are presented in Table 6 and Table 7 , respectively. We have used the following metrics to evaluate the framework: TP: True positive, FP: False positive, FN: False Negative,

../../../../projets/stars/IMG/flow_diagram.jpg

Figure 40. The flow diagram of the unsupervised activity recognition framework: (a) off-line learning phase and (b) on-line recognition phase

Sensitivity and Precision. According to the trajectory information, sometimes k-means clustering produces zones that are actually union of more than one zones. For such cases, we have combined the actions and presented as one single action.

Table 6. The recognition results obtained by using the 2D camera.

Actions	Instances	TP	FP	FN	Sensitivity(%)	Precision (%)
Paying bill	13	5	0	8	38.46	100
Preparing drugs	7	5	5	2	71.42	50
Looking at bus map+Watching TV	21	6	3	15	28.57	66.66
Sitting at the table	18	6	10	12	33.33	37.5
Talking on the phone	23	17	1	6	73.91	94.44
Preparing tea	23	11	3	12	47.82	78.57

Table 7. The recognition results obtained by using the RGB-D camera.

Actions	Instances	TP	FP	FN	Sensitivity(%)	Precision (%)
Paying bill + Watching TV	13	12	8	1	92.3	60
Preparing drugs	5	5	0	0	100	100
Looking at bus map	9	9	10	0	100	47.36
Sitting at the table	8	4	34	4	50	10.52
Talking on the phone	14	13	1	1	92.85	92.85
Preparing tea	16	9	5	7	56.25	64.28

As it can be seen in the tables, we obtain higher recognition rates by using the information coming from RGB-D camera.

Table 6 shows that for “talking on the phone” and “preparing drugs” actions occurring in two distant zones, using 2D camera gives high recognition rates (higher than 70 %). However, the actions “looking at bus map”, “watching TV” and “sitting at the table” are misclassified (low TP and high FP). Since the zones of these actions are very close to each other, the actions occurring in the borders are not well recognized. The reason of high FN is due to the problems in detection and tracking with 2D video cameras. The process of trajectory extraction described in [81] sometimes fails to track people. Because of the inadequate trajectory information, we have many FNs. Therefore, a better detection can considerably enhance the recognized actions.

By using the information coming from RGB-D camera, except for “sitting at the table” and “preparing tea” actions, we achieve high level of recognition rates (Table 7). However, similar to 2D camera, the recognition of “sitting at the table”, “paying bill” and “watching TV” actions fails because the learned zones in the scene are very close to each other. Hence, we have many false positives (FP) and false negatives (FN) for “sitting at the table” and “preparing tea” actions.

In the light of the preliminary experimental results, we can say that this unsupervised algorithm has a potential to be used for automated learning of behavioral patterns in unstructured scenes, for instance in home care environment for monitoring older people. Since the current framework does not require the user interaction to label activities, an evaluation process on big datasets will be easily performed. The proposed framework gives one action at each zone in an unsupervised way. We are currently focusing on refining the actions for each zone by using the pixel tracklets of the person's body parts. This will be achieved by performing clustering among activity models. As an example, the action of "sitting at the table" will be decomposed to "reading newspaper while sitting at the table" and "distributing cards while sitting at the table".

6.21. Extracting Statistical Information from Videos with Data Mining

Participants: Giuseppe Donatiello, Hervé Falciani, Duc Phu Chau, François Brémond.

keywords: video data mining, activity recognition, clustering techniques

Objective

Manual video observation is becoming less practical due to growing size of data. To tackle this problem, we have built a system to retrieve videos of interest thanks to an index based on activities recognized in an automated manner. We automatically detect activities in videos by combining data mining and computer vision to synthesize, analyze and extract valuable information from video data.

Approach

Our research introduces a new method for extracting statistical information from a video. Specifically, we focus on context modeling by developing an algorithm that automatically learns the zones in a scene where most activities occur by taking as input the trajectories of detected mobiles. Using K-means clustering, we define activity zones characterizing the scene dynamics, we can extract then people activities by relating their trajectories to the learned zones.

Results

To evaluate our system we have extended the OpenJUMP framework, an open source for Geographic Information System (GIS). The end user can have an overview of all activities of a large video, with the possibility of extracting and visualizing activities classified as usual or unusual. We have tested our approach on several videos recorded in subways in Turin (Italy) and Paris, as shown below, some examples of unusual activities (Figures 41 , 42 and 43).The system has been showed in a live demonstration at RATP company in Paris for the European project Vanaheim (http://www.vanaheim-project.eu/).

6.22. SUP

Participants: Julien Gueytat, François Brémond.

keywords: SUP, Software, Video Processing

Presentation

SUP is a Scene Understanding Software Platform writtent in C++ designed for analyzing video content. (see Figure 44) SUP is splitting the workflow into several modules, such as acquisition, segmentation, etc., up to activity recognition. Each module has a specific interface, and different plugins implementing these interfaces can be used for each step of the video processing.

../../../../projets/stars/IMG/figure1.png

Figure 41. Results : three categories of discovered abnormal behaviors.

../../../../projets/stars/IMG/figure2.png

Figure 42. Trajectories and zones from camera M3114 in Paris (the yellow line is the trajectory selected by the user; the numbers represent the weight belonging to each segment : the higher, the more the people go through it.

../../../../projets/stars/IMG/figure3.png ../../../../projets/stars/IMG/figure4.png ../../../../projets/stars/IMG/figure

Figure 43. Left : Person sitting. Middle : Person standing for a long time. Right : Unusual path.

../../../../projets/stars/IMG/SUP-architecture.jpg

Figure 44. SUP workflow

The plugins cover the following research topics:

- algorithms : 2D/3D mobile object detection, camera calibration, reference image updating, 2D/3D mobile object classification, sensor fusion, 3D mobile object classification into physical objects (individual, group of individuals, crowd), posture detection, frame to frame tracking, long-term tracking of individuals, groups of people or crowd, global tacking, basic event detection (for example entering a zone, falling...), human behaviour recognition (for example vandalism, fighting,...) and event fusion;
- languages : scenario description, empty 3D scene model description, video processing and understanding operator description;
- knowledge bases : scenario models and empty 3D scene models;
- algorithms of 2D & 3D visualisation of simulated temporal scenes and of real scene interpretation results;
- algorithms for evaluation of object detection, tracking and event recognition;
- learning techniques for event detection and human behaviour recognition;
- algorithms for image acquisition (RGB and RGBD cameras) and storage;
- algorithms for video processing supervision;
- algorithms for data mining and knowledge discovery;
- algorithms for image/video indexation and retrieval.

The software is already widely disseminated among researchers, universities, and companies:

- PAL Inria partners using ROS PAL Gate as middleware
- Nice University (Informatique Signaux et Systèmes de Sophia), University of Paris Est Créteil (UPEC - LISSI-EA 3956)
- European partners: Lulea University of Technology, Dublin City University,...
- Industrial partners: Toyota, LinkCareServices, Digital Barriers

And new sites are coming: EHPAD Valrose, Institut Claude Pompidou, Delvalle and Biot.

Improvements

Our team focuses on developing a Scene Understanding Platform (SUP). This platform has been designed for analyzing video content. SUP is able to recognize events such as 'falling', 'walking' of a person. We can easily build new analyzing systems thanks to a set of algorithms also called plugins. The order of those plugins and their parameters can be changed at run time and the result visualized on a dedicated GUI. This platform has many more advantages such as easy serialization to save and replay a scene, portability to Mac, Windows or Linux, and easy deployment to quickly setup an experimentation anywhere. All those advantages are available since we are working together with the Inria software developer team SED. Many Inria teams are pushing together to improve a common Inria development toolkit DTK. Our SUP framework is one of the DTK-like framework developed at Inria.

Currently, the OpenCV library is fully integrated with SUP. OpenCV provides standardized dataypes, a lot of video analysis algorhithms and an easy access to OpenNI sensors such as the Kinect or the ASUS Xtion PRO LIVE.

Updates and presentations of our framework can be found on our team website https://team.inria.fr/stars/software . Detailed tips for users are given on our Wiki website http://wiki.inria.fr/stars and sources are hosted thanks to the software developer team SED.

6.23. Model-Driven Engineering for Activity Recognition

Participants: Sabine Moisan, Jean-Paul Rigault, Luis Emiliano Sanchez.

keywords: Feature Model Optimization, Software Metrics, Requirement specification, Component-based system, Dynamic Adaptive Systems, Model-Driven Engineering, Heuristic Search

The domain of video surveillance (VS) offers an ideal training ground for Software Engineering studies, because of the huge variability in both the surveillance tasks, the video analysis algorithms and the context. Such systems require run time adaptation of their architecture to react to external events and changing conditions in their context of execution.

The feature model formalism is widely used to capture variability, commonalities and configuration rules of software systems. We thus use feature modeling to represent the variability of both the specification and component views of video surveillance systems. We also include cross-tree constraints that formalize extra feature dependencies.

Based on this feature model, we can both select an initial system configuration at deployment time and dynamically adapt the current configuration at run time. This year we focused on runtime adaptation, from feature model to running components.

6.23.1. Configuration Adaptation at Run Time

In the continuation of our work on metrics on feature models, we have integrated a configuration selection algorithm in our feature model manager. Context changes or user interactions imply to dynamically reconfigure the model (selecting or deselecting features). Following *model at run time* techniques, we are able to determine the set of valid configurations to apply in a new execution context. Since only one configuration can be applied at a given time, the role of the selection algorithm is to select the "best" one.

To this end we enriched our feature representation with a set of *quality attributes* that correspond to a monotonic quantification of interesting aspects of the system quality. Examples are response time, accuracy, availability, performance, component switching time, etc. The configuration selection algorithm optimises a cost function, a linear weighted combination of the quality attributes. Thus we can rank the possible valid configurations and choose an optimal one. Our algorithm is a variant of the Best-First Search algorithm, a heuristic graph search technique. It starts with the set of valid configurations, which is a feature model where some features are *unselected*. Then it performs a systematic search in a graph where nodes are configurations and edges are selections or deselections of unselected features. The goal is to obtain a full configuration (one without unselected features) optimizing the cost function. The algorithm is parameterized with different strategies and associated heuristics with different optimality and efficiency characteristics.

Search strategies decide which node to visit next. We choose two well-known informed strategies that rely on heuristic functions as choice criteria. First we used a variant of the A* algorithm, BF*, but with a node-cost function instead of a path-cost one; it favors optimality over efficiency. Second, we implemented a Greedy Best-First Search (GBFS) strategy, where the next visited node is the best successor of the current one; it favors efficiency over optimality.

Computing the exact value of the cost function for a partial configuration is too expensive. We thus use heuristics to obtain a quick *estimate*. We have tested two sorts of heuristics. The simplest one, H_A, ignores the differences between the various sorts of groups (AND, OR, XOR) in the feature model and does not considers cross-tree constraints; it is fast but not very accurate. The second one, H_B, just drops the cross-tree constraints; it is thus more accurate, yet at an higher cost.

We have run experiments using large (randomly generated) feature models and compared completeness, optimality and efficiency of the selection algorithm, with different combinations of strategies and heuristics [42]. From our experiments, the GBFS strategy with heuristics H_B appears as the ideal option for real time systems that have to adapt in bounded time. This strategy ensures polynomial time complexity and guarantees optimality over 90%, which is good enough for our purpose (see figure 45). On the other hand, BF* strategy with heuristics H_B is ideal for offline decisions, such as defining the initial configuration of a system. Although this search strategy takes a significant time to compute, this is acceptable at deployment time to obtain the optimal configuration.

6.23.2. Run Time Components

When a configuration has been chosen, we must implement it with real components. We consider a configuration of a video-surveillance processing chain as a set of running components, that can be tuned, removed,

../../../../projets/stars/IMG/experiment_E2C.png

Figure 45. Optimality comparison of heuristics using GBFS strategy with respect to the number of attributes in the cost function

added, or replaced dynamically, in response to events. To apply such configuration changes, we need a way to represent and dynamically manipulate the components themselves.

In a first attempt, we used an OSGi-like C++ framework (SOF, Service Oriented Framework However, SOF did not really fulfilled our needs. First, SOF is the only C++ OSGi framework that we could find and its C++ implmentation deserves some improvement. Moreover, like OSGi, it relies on the notion of “service”, as can be found in Web applications, but which does not really fit our real time requirements. This notion of service is not our concern and makes programming more complicated than necessary.

Thus, we decided to define our own component module and to integrate it in a multi-threaded layer, easy to use for our end-users who are video system developers. Each component runs by default in its own thread and communicates with other components through standardized communication channels. Our goal is to provide end-users with simple patterns to package their video codes into components. Thus we hide as much as possible the technical details such as threading synchronization, data exchange, and mechanisms for component management (replacement, tuning...) ensuring a continuous process.

We are currently setting up this framework on a simple video detection pipeline with OpenCV-based components. Then we shall integrate it within our Model at Run Time architecture.

6.24. Scenario Analysis Module

Participants: Annie Ressouche, Daniel Gaffé, Narjes Ghrairi, Sabine Moisan, Jean-Paul Rigault.

Keywords: Synchronous Modelling, Model checking, Mealy machine, Cognitive systems.

To generate activity recognition systems we supply a scenario analysis module (SAM) to express and recognize complex events from primitive events generated by SUP or other sensors. The purpose of this research axis is to offer a generic tool to express and recognize activities. Genericity means that the tool should accommodate any kind of activities and be easily specialized for a particular framework. In practice, we propose a concrete language to specify activities in the form of a set of scenarios with temporal constraints between scenarios. This language allows domain experts to describe their own scenario models. To recognize instances of these models, we consider the activity descriptions as synchronous reactive systems [80] and we adapt usual techniques of synchronous modelling approach to express scenario behaviours. This approach facilitates scenario validation and allows us to generate a recognizer for each scenario model.

Since last year, we relied on CLEM (see section 6.25) synchronous language to express the automata semantics of scenario models as Boolean equation systems. This year, we continue our research in this direction and we are studying a specific semantics of SAM language operators that translates any SAM program into Boolean equation system. Therefore, we will benefit from CLEM compilation technique to generate recognizer for each scenario model.

This year we focus on the definition of an execution machine able to transform asynchronous events coming from SUP or other devices into synchronous significant events feeding recognition engines generated by SAM. The execution machine can listen three types of asynchronous events: SUP events, Boolean sensors, sampled sensors and pulse train sensors. According to the sampling period of each sensor, the execution machine builds the significant events defining the synchronous logical instants which trigger the reaction of the scenario recognition engine. Thanks to the synchronous approach, scenario recognition engines are able to dynamically express the expected synchronous events of the next step; the execution machine takes into account of this information to filter relevant events. We perform several tests with real SUP data sets and the execution machine has a convincing behaviour (see [55]). To complement this work, we will integrate a notion of incompatible events which will make the execution machine more efficient and robust.

6.25. The Clem Workflow

Participants: Annie Ressouche, Daniel Gaffé, Joel Wanza Weloli.

Keywords: Synchronous languages, Synchronous Modelling, Model checking, Mealy machine.

This research axis concerns the theoretical study of a synchronous language LE with modular compilation and the development of a toolkit (see Figure 9) around the language to design, simulate, verify and generate code for programs. The novelty of the approach is the ability to manage both modularity and causality. This year, we mainly work on the implementation of new theoretical results concerning the foundation of LE semantics. We also design a new simulator for LE programs which integrates our new approach.

First, synchronous language semantics usually characterizes each output and local signal status (as present or absent) according to input signal status. To reach our goal, we defined a semantics that translates LE programs into equation systems. This semantics bears and grows richer the knowledge about signals and is never in contradiction with previous deduction (this property is called constructiveness). In such an approach, causality turns out to be a scheduling evaluation problem. We need to determine all the partial orders of equation systems and to compute them, we consider a 4-valued algebra to characterize the knowledge of signal status (unknown, present, absent, overknown). In [69], we chosen an algebra which is a bilattice and we show that it is well suited to solve our problem. It is a new application of general bilattice theory [70]. This year, this approach has been improved, validated in CLEM compiler and published in [39]. Our compilation technique needs to represent Boolean equation systems with Binary Decision Diagrams (BDD) and we study and design a specific BDD library well suited to ours needs. From a practical point of view, we integrate new operators in LE language (sustain until, no emit, strong abort). We focus on automata extension which can consider now three types of transition: weak transition, strong transition and normal termination transition.

Second, in CLEM, we generate an independent intermediate code (LEC) before specific target generations. This code represents the semantics of programs with 4-valued equation systems. In our design flow, we need to simulate programs at this level. This year, we design the CLES simulator which interprets LEC. The actual version don't integrate the data part of the language and we plan to do this integration.

6.26. Multiple Services for Device Adaptive Platform for Scenario Recognition

Participants: Annie Ressouche, Daniel Gaffé, Mohammed Cherif Bergheul, Jean-Yves Tigli.

Keywords: Synchronous Modelling, Model checking, Mealy machine, Ubiquitous Computing.

The aim of this research axis is to federate the inherent constraints of an activity recognition platform like SUP (see section 5.1) with a service oriented middleware approach dealing with dynamic evolutions of system infrastructure. The Rainbow team (Nice-Sophia Antipolis University) proposes a component-based adaptive middleware (WComp [88], [87], [73]) to dynamically adapt and recompose assemblies of components. These operations must obey the "usage contract" of components. The existing approaches don't really ensure that this usage contract is not violated during application design. Only a formal analysis of the component behaviour models associated with a well sound modelling of composition operation may guarantee the respect of the usage contract.

The approach we adopted introduces in a main assembly, a synchronous component for each sub assembly connected with a critical component. This additional component implements a behavioural model of the critical component and model checking techniques apply to verify safety properties concerning this critical component. Thus, we consider that the critical component is validated.

In [84], [83], we showed that an efficient means to define the synchronous components which allow to validate critical component behaviours, is to specify them with Mealy machines. Previously, we used a classical synchronous language (Lustre) to specify synchronous components, but the integration of the synchronous component code into WComp was not straightforward because Lustre compiler is not opened and cannot integrate new target code needed by WComp. This year, we supply GALAXY automata editor to express Mealy machines and we extend AUTOM2CIRCUIT compiler to generate the internal code of WComp (C#). AUTOM2CIRCUIT is a tool developed by D. Gaffé since several years which compiles an explicit representation of automata into Boolean Mealy machine and generate a large and opened set of targets. This work is a preliminary study to integrate this generation of C# into CLEM.

STEEP Team

6. New Results

6.1. Data Mining for Material Flow Analysis: Application in the Territorial Breakdown of French Regions

One of the major issues for assessment of the long-term sustainability of urban areas is related to the concept of "imported sustainability". In order to produce such an assessment for a given territory, one must first identify and quantify the types of materials used, and the impacts associated to these uses. Material Flow Analysis (MFA) is directly related to how the material circulates and how it is transformed within a territory. In most cases this analysis is performed at national and regional levels, where the statistical data is available. The challenge is to establish such an analysis at smaller scales, e.g. in the case of France, at the department or city level.

We have explored the possibility of applying data analysis at the regional level by generating a mathematical model that can fit well the data at regional scale and estimate well the departmental one. The downscaling procedure relies on the assumption that the obtained model at level 'n' (for example region) will be also true at level 'n+1' (for example department), such that it could properly estimate the unknown data based on a set of chosen drivers (socio-economic data). We have designed and implemented techniques based on parameter optimization and model selection as well as robust estimation, in order to estimate the best drivers for a given set of territories, i.e. the socio-economic data (e.g. employees per type of manufacturing industry, population data, etc.) that best correlate with the production of various types of agricultural or other products [19].

6.2. Calculating indices for urban sprawl

Urban sprawl is a complex concept, that is generally associated with auto-oriented, low-density development. It is the subject of a wide range of research efforts, aiming at understanding and characterizing the underlying driving factors. We have followed up on an effort by Burchfield et al. who proposed a simple measure for urban sprawl, a so-called sprawl index. We proposed several variants of this index with the aim of achieving richer and/or more flexible characterizations of urban sprawl [16]. We also proposed ways of determining metropolitan areas that have similar patterns of urban sprawl, using clustering techniques.

6.3. Computing environmental accounts from the consumer's viewpoint using Input Output Analysis

The Russian-American economist Wassily Leontief introduced Input-Output Analysis (IOA) in the 30's and was awarded the Nobel prize in economy in 1973 for this contribution. IOA is a macro-economic tool which investigates the links and retroactions between the sectors of an economy. It makes it possible to allocate production factors (labor, capital etc.) and environmental externalities of production processes (depletion on resources, emission of pollutants etc.) to final consumption. Our first task was to reproduce the results from various studies on the carbon footprint of France. We couldn't reproduce the results from *Analyse des impacts environnementaux de la consommation des ménages et des marges de manœuvre pour réduire ces impacts* (Ademe, 2012). We underlined a mistake in the mathematical formulas presented in the annex of the paper but couldn't concur it was indeed the source of discrepancy because we were not granted access to the raw data of the study. Our results are however in line with the papers originating from the statistical service of the ministry of ecology (SOeS, J-L Pasquier).

Because working on excel sheets, although widespread in a large number of agencies, proves very inefficient, we started to work on the development of a software called Wassily (see "new results" section) that would automatize the critical calculations. We prepared databases on input-output tables and air emissions of several European countries based on Eurostat data and started to work on the architecture of the software with Julien Alapetite. In parallel, we reviewed the downscaling and regionalizing techniques in the IO literature and looked for the necessary information concerning the Rhône-Alpes region. We concluded that enough data was available in order to carry out downscaling but that data was still too scarce for the finer levels of regionalizing.

6.4. Mapping and land use and land cover change for the ESNET project

The ESNET project (EcoSystem services NETworks) is a collaboration lead by LECA (Laboratoire d'ECologie Alpine, UJF) that aims at characterizing the ecosystem services of the Grenoble urban region (about 2/3 of the Isere département) at the 2030/2040 horizon under various constraints of urban policy planning, changes in agricultural and forest management, and climate change impact on ecosystems. A preliminary task in this research program was the elaboration of very detailed maps (both in terms of land use and of resolution) of the study area at three different dates (1998, 2003 and 2009) based on available satellite and IGN data, in order to characterize past land use patterns as well as agricultural rotation patterns. These have been made and completed at Inria with the hiring of specialized engineers in these tasks, funded by the ESNET program. This exercice informs the next task (land use and land cover change – LUCC – modelling). Hosting this work at Inria was not only logical in terms of the available computer environment, but also useful in terms of visibility of Inria from outside planning agencies.

The LUCC model itself is developed partly at Inria (for modelling expertise) and partly at LECA (for expertise on ecological change drivers). The model development is still underway but in a rather advanced stage. Relevant drivers for urban development have been identified and statistically characterized. The so-called "transition potentials" (which characterize change of land use over a given period of time) are in the process of being calibrated. The next steps involve the completion of this calibration task, the development of relevant scenarios (underway by the whole ESNET collaboration) and projections of land use into the future. Some sensitivity analysis will also be performed in order to characterize the robustness of the model.

6.5. Benchmarking tools for the climate negotiation of GHG emission reduction trajectories

Climate negotiations related to global warming are another important issue of sustainable development. In this framework that is place at international scale we propose a benchmarking tool that is designed to avoid the main limitations of actual negotiation schemes. Our approach is based on the original Soft Landing proposition, made by Criqui and Kouvaritakis in the early 2000. We develop an up to date solution which improves the original idea mainly by introducing common but differentiated emission reduction profiles and by developing a dedicated algorithm for that purpose (called REDEM). To be compatible with global objectives, it is commonly accepted that for most developing regions, the national emission curves should admit a maximum and then should progressively decline. Similarly, we emphasize the fact that, in order to achieve the global objectives, all states will have to entail mitigation efforts, the intensity which may be measured by the rate of variation of the national emissions. At one point, the effort will reach a maximum, when the rate of variation in absolute value is at its maximum, and then decrease. In other words, there will also be a peak in the effort. Then we propose to base the benchmark on this peak of effort. This work has been done in collaboration with EDDEN Laboratory, in particular Patrick Criqui and Constantin Ilasca.

SUMO Team

6. New Results

6.1. Model expressivity and quantitative verification

6.1.1. Diagnosis from scenarios

Participants: Loïc Hélouët, Blaise Genest, Hervé Marchand.

Diagnosis of a system consists in providing explanations to a supervisor from a partial observation of the system and a model of possible executions. This year, we have extended results on diagnosis algorithm from scenarios. Systems are modeled using High-level Message Sequence Charts (HMSCs), and the diagnosis is given as a new HMSC, which behaviors are all explanations of the partial observation. The results published this year are first an offline centralized diagnosis algorithm (a single process in a network collects an observation, and emits a diagnosis) that has then been extended to a decentralized version of this algorithm. This allows us to give a complete diagnosis framework for infinite state systems, with a strong emphasis on concurrency and causal ordering in behaviors. HMSC-based diagnosis showed nice properties w.r.t. compositionality. We have also considered solutions for online diagnosis from scenarios, but came to the conclusion that online solutions are memory consuming, and need too many restrictions to run with finite memory.

The last contribution of this work is an application of diagnosis techniques to anomaly detection, that is a comparison of observation of the system with a model of usual behaviors to detect security attacks. This work is already available online in [25], and will soon be published.

6.1.2. Probabilistic model checking

Participants: Nathalie Bertrand, Blaise Genest, Paulin Fournier.

In [20], we considered the verification of Markov chains against properties talking about distributions of probabilities. Even though a Markov chain is a very simple formalism, by discretizing in a finite number of classes the space of distributions through some symbolics, we proved that the language of trajectories of distribution (one for each initial distribution) is not regular in general, even with 3 states. We then proposed a parametrized algorithm which approximate what happens to infinity, such that each symbolic block in the approximate language is at most ϵ away from the concrete distribution.

With the objective of model checking infinite state probabilistic systems, we proved a general finite-time convergence theorem for fixpoint expressions over a well-quasi-ordered set [22]. This has immediate applications for the verification of well-structured systems, where a main issue is the computability of fixpoint expressions, and in particular for game-theoretical properties and probabilistic systems where nesting and alternation of least and greatest fixpoints are common [35].

Parameterized verification aims at validating a system's model irrespective of the value of a parameter. In [34] we introduced a model for networks of an arbitrary number of probabilistic timed processes, communicating by broadcasting. This model is suitable for distributed protocols, and can be applied to wireless sensor networks or peer-to-peer applications. The number of processes is unknown and either is constant (static case), or evolves over time through random disappearances and creations (dynamic case). On the one hand, most parameterized verification problems turn out to be undecidable in the static case (even for untimed processes). On the other hand, we prove their decidability in the dynamic case.

6.1.3. Distributed timed systems

Participants: Nathalie Bertrand, Amélie Stainer.

We study the reachability problem for communicating timed processes, both in discrete and dense time. Our model comprises automata with local timing constraints communicating over unbounded FIFO channels. Each automaton can only access its set of local clocks; all clocks evolve at the same rate. Our main contribution is a complete characterization of decidable and undecidable communication topologies, for both discrete and dense time. We also obtain complexity results, by showing that communicating timed processes are at least as hard as Petri nets; in the discrete time, we also show equivalence with Petri nets. Our results follow from mutual topology-preserving reductions between timed automata and (untimed) counter automata. To account for urgency of receptions, we also investigate the case where processes can test emptiness of channels. This resut is published in [39] and is a part of Amélie Stainer's PhD manuscript [18]. It also constitutes a contribution to ANR VACSIM.

We also studied a model for distributed systems composed of stochastic and timed processes that interact via broadcasting. For these networks of stochastic timed automata (NSTA), we provided a precise performance evaluation algorithm, without resorting to simulation techniques. The idea is to characterize the general state space Markov chain through transient stochastic state classes that represent the system's state after each action. This yields an algorithmic approach to the transient analysis of NSTA models, with fairly general termination conditions [32].

6.2. Management of large distributed systems

6.2.1. Test generation from Recursive Tile Systems

Participants: Sébastien Chédor, Thierry Jéron, Christophe Morvan.

We explore the generation of conformance test cases for Recursive Tile Systems (RTSs) in the framework of the classical ioco testing theory. The RTS model allows the description of reactive systems with recursion, and is very similar to other models like Pushdown Automata, Hyperedge Replacement Grammars or Recursive State Machines. Test generation for this kind of models is seldom explored in the literature. We first propose an off-line test generation algorithm for Weighted RTSs, a determinizable sub-class of RTSs, and second, an on-line test generation algorithm for the full RTS model. Both algorithms use test purposes to guide test selection through targeted behaviours. Additionally, essential properties relating verdicts produced by generated test cases on an implementation with both the conformance with respect to its specification, and the precision with respect to a test purpose, are proved. This work is published in [51], and a journal version will appear in 2014. It is also a part of Sébastien Chédor's PhD manuscript.

6.2.2. Distributed control

Participants: Blaise Genest, Hervé Marchand.

We focused this year on the control of distributed systems modeled as *asynchronous automata*, that is asynchronous network of automata communicating through peer to peer synchronizations. First, we considered the case where all events are controllable, and the objective is to accept exactly a given language. Here, a famous result is the Zielonka theorem [62], stating that every regular language closed under commutation can be turned into an asynchronous automaton. However, the construction is plagued with deadends and final state of the network are decided by a global controller monitoring every process at the same time and perfectly, which is unrealistic and defeat the distribution idea. This year, we characterized the languages which can be controlled realistically (no deadends, local final states and local decision on each process), and give algorithms to obtain the associated distributed machines in [30]. The case where some events are uncontrollable is reputed very difficult. We made a progress this year in [42], showing that we can decide whether a reachability objective can be ensured, granted that the communication between the processes follow a tree: siblings can not communicate directly together, they need to go through their common parent.

In [27], we consider an alternative model for the control of distributed systems; the aim is to build local controllers that restrict the behavior of a distributed system in order to satisfy a global state avoidance property. We model distributed systems as communicating finite state machines with reliable unbounded FIFO queues between subsystems. Local controllers can only observe the behavior of their proper subsystem and do not see the queue contents. To refine their control policy, controllers can use the FIFO queues to communicate by piggy-backing extra information (some timestamps and their state estimates) to the messages sent by the subsystems. We provide an algorithm that computes, for each local subsystem (and thus for each controller), during the execution of the system, an estimate of the current global state of the distributed system. We then define a synthesis algorithm to compute local controllers. Our method relies on the computation of (co-)reachable states. Since the reachability problem is undecidable in our model, we use abstract interpretation techniques to obtain overapproximations of (co-)reachable states. Similarly, in [46], we have been interested in the control of distributed systems with synchronous communications (called decentralized Discrete Event Systems). We introduced a novel architecture that extends the class of problems that can be solved in decentralized DES control in the absence of communication. In this architecture, unlike previous architectures that use either conjunction or disjunction to fuse local control decisions, the fusion rule is exclusive or. We characterized the new architecture, where controllers take a single decision, with respect to the recently-proposed multi-decision framework of Chakib and Khoumsi. Unlike previous architectures, parity-based controllers cannot predetermine their local control decision based solely on their local observations. Instead, the local control decisions are calculated a priori.

6.2.3. Enforcement of timed and security properties

Participants: Thierry Jéron, Hervé Marchand, Srinivas Pinisetty.

Runtime enforcement is a verification/validation technique aiming at correcting (possibly incorrect) executions of a system of interest. This year, we first consider enforcement monitoring for systems with timing specifications (modeled as timed automata). We consider runtime enforcement of any regular timed property specified by a timed automaton [45]. To ease their design and their correctness-proof, enforcement mechanisms are described at several levels: enforcement functions that specify the input-output behavior, constraints that should be satisfied by such functions, enforcement monitors that implement an enforcement function as a transition system, and enforcement algorithms that describe the implementation of enforcement monitors. The feasibility of enforcement monitoring for timed properties is validated by prototyping the synthesis of enforcement monitors. This work is also a contribution to ANR Vacsim. In [41], we studied an alternative enforcement problem of security properties, namely, the enforcement of K-step opacity at runtime. In K-step opacity, the knowledge of the secret is of interest to the attacker within K steps after the secret occurs and becomes obsolete afterwards. We introduce the mechanism of runtime enforcer that is placed between the output of the system and the attacker and enforces opacity using delays. If an output event from the system violates K-step opacity, the enforcer stores the event in the memory, for the minimal number of system steps until the secret is no longer interesting to the attacker (or, K-step opacity holds again)

6.2.4. Discrete control of computing systems administration

Participants: Hervé Marchand, Nicolas Berthier.

We address the problem of using Discrete Controller Synthesis for the administration of Computing Systems, following an approach supported by a programming language [24]. We present a mixed imperative/declarative programming language, where declarative contracts are enforced upon imperatively described behaviors. Its compilation is based on the notion of supervisory control of discrete event systems. More precisely, our language can serve programming closed-loop adaptation controllers, enabling flexible execution of functionalities w.r.t. changing resource and environment conditions. DCS is integrated into a1 programming language compiler, which facilitates its use by users and programmers, performing executable code generation. The tool is concretely built upon the basis of a reactive programming language compiler, where the nodes describe behaviors that can be modeled in terms of transition systems. Our compiler integrates this with a DCS tool, making it a new environment for formal methods. We apply our method to the problem of coordinating several administration loops in a data center (number of servers, repair, and local processor frequencies) [40].

We formulate this problem as an invariance controller synthesis problem. We are currently working on an extension of the controller synthesis tool so that it can handle the use of numerical variables in order to model both the system and the properties to be ensured by control.

6.2.5. Distributed planning

Participant: Éric Fabre.

Planning problems consist in organizing actions in a system in order to reach one of some target states. The actions consume and produce resources, can of course take place concurrently, and may have costs. We have a collection of results addressing this problem in the setting of distributed systems. This takes the shape of a network of components, each one holding private actions operating over its own resources, and shared/synchronized actions that can only occur in agreement with its neighbors. The goal is to design in a distributed manner a tuple of consistent local plans, one per component, such that their combination forms a global plan of minimal cost.

Our previous solutions to this problem modeled components as weighted automata. In collaboration with Loïg Jezequel (TU Munich) and Victor Khomenko (Univ. of Newcastle), we have extended this approach to the case of components modeled as safe Petri nets [44]. This allows one to benefit from the internal concurrency of actions within a component. Benchmarks have shown that this method can lead to significant time reductions to find feasible plans, in good cases. In the least favorable cases, performances are comparable to those obtained with components modeled as automata. The method does not apply to all situations however, as computations require to perform ϵ-reductions on Petri nets.

6.2.6. Diagnosis based on self-modeling

Participants: Éric Fabre, Carole Hounkonnou.

Model-based approaches have been proved to provide the best results for fault diagnosis in telecommunication networks, with various kinds of models. They suffer however from several difficulties: one has to build a model adequate to the supervised network (and possibly adapt it as the network evolves), one has to find the correct abstraction level for this model, and one has to deal with size issues of such models. In Carole Hounkonnou's thesis [15], we have proposed an approach that addresses these three limitations, under the generic name of self-modeling. It consists modeling a network in a generic manner, through its building rules. The actual instance one has to manage is then discovered on the fly, when some malfunction explanation request is triggered. Starting from the identified malfunction, the network model instance is discovered/revealed progressively, as requested by the needs of the diagnosis procedure. The latter progressively extends a Bayesian network model of the network, in order to collect more information and identify the malfunction rootcause. The model extension is guided by an information theory criterion: it seeks access to the new observations that are be the most informative (on the average) given previous observations taken into account. This approach allows to deal with potentially large models, as the supervised system needs not be entirely modeled before the diagnosis starts. We are currently working on the extension of this setting to model refinement, and to a framework of dynamic systems rather than static systems.

6.2.7. Graceful restart methods for link state routing protocols

Participants: Éric Fabre, Carole Hounkonnou.

Link state routing protocols are ubiquitous in the internet. OSPF (Open Shortest Path First) is one of them within an Autonomous System. In collaboration with Alcatel-Lucent, we have proposed an extension of graceful restart procedures, that allow to shut down the control plane of routers while maintaining the data plane active, and thus the packet forwarding activity. A drawback of existing procedures was that frozen routers had to be removed from the network as soon as topology evolved. We have shown that this pessimistic precaution could be damageable to the network and was not necessary [43]. Frozen routers may still be useful, even if they do not forward packets in an optimal manner. And even if they create routing loops, the latter can be easily detected, and optimally patched, which is often more efficient than declaring these routers as dead. Experiments on classical topologies of the topology zoo, as well as on random topologies, have confirmed these results.

6.3. Data driven systems

6.3.1. Web services

Participants: Blaise Genest, Loïc Hélouët.

This year, we considered transactional properties (ACID) for web services. In particular, we focused on the atomicity (A of ACID) property, obtained in case of a failure inside an atomic block through compensation of the executed actions of the block. To do so, logs need to be kept. We were interested in maintaining the maximal amount of privacy. We proposed modular algorithms [23] which maintain privacy between modules, with minimal information shared among modules, both in the logging and the compensation phases. Furthermore, each module logs a small number of information, such that the sum of all actions logged is guaranteed minimal. Last, modularity allows fast algorithms, as they need to consider only what happens in the module itself, and not the exact structure of its parent module nor of its sub-modules.

We also have extended the *session system* model originally proposed in [55]. We have deisgned a mode for Web-based systems that allows to describe systems running an arbitrary number of transactions over an arbitrary number of agents. For these systems, syntactic restrictions allow to decide coverability properties, and then more elaborated business rules, such as conflict of interest (the fact that a participant to a system can be involved in two exclusive services), or the Chinese Wall Property (that prevents users of a system to use benefits or information right they may have obtained from a privileged role at later instant of any execution of the system. These results were obtained with M. Mukund and S. Akshay within the context of the DISTOL associated team, and should lead to a publication next year.

6.3.2. Implementation of scenarios

Participants: Loïc Hélouët, Rouwaida Abdallah.

We have revisited the problem of program synthesis from specifications described by High-level Message Sequence Charts. The main objective is to obtain a distributed implementation (for instance described with communicating automata) from a global specification given as High-level MSCS. In the general case, synthesis by a simple projection on each component of the system allows more behaviors in the implementation than in the specification. The differences arise from loss of ordering among messages, but we have shown that for a subclass of HMSCs (the *local HMSCs*) behaviors can be preserved by addition of communication controllers, that intercept messages to add stamping information before resending them, and deliver messages to processes in the order described by the specification. This work was published in [19].

The second aspect of our work on scenarios implementability has considered implementation of requirements expressed as non-local HMSCs. We have proposed a new technique to transform an arbitrary HMSC specification into a local HMSC, hence allowing implementation. This transformation can be automated as a constraint optimization problem, and the impact of modifications brought to the original specification minimized w.r.t. a cost function. The approach was evaluated on a large number of randomly generated HMSCs, and the results show an average runtime of a few seconds, which demonstrates applicability of the technique. These results were published in [28]. Both results mentionned in this sections are part of the PhD thesis of Rouwaida Abdallah, defended this year [14].

6.3.3. Attribute grammars

Participant: Éric Badouel.

Evaluation of attributes w.r.t. an attribute grammar can be obtained by inductively computing a function expressing the dependencies of the synthesized attributes on inherited attributes. This higher-order functional approach to attribute evaluation can straightforwardly be implemented in a higher-order lazy functional language like Haskell. The resulting evaluation functions are, however, not easily amenable to optimization when we want to compose two attribute grammars. In [21], we present an alternative first-order functional interpretation of attribute grammars where the input tree is replaced by an extended cyclic tree each node of which is aware of its context viewed as an additional child tree. These cyclic representations of zippers (trees with their context) are natural generalizations of doubly-linked lists to trees over an arbitrary signature.

Then we show that, up to that representation, descriptional composition of attribute grammars reduces to the composition of tree transducers.

TAO Project-Team

6. New Results

6.1. Continuous Optimization

Participants: Ouassim Ait Elhara, Yohei Akimoto, Asma Atamna, Anne Auger, Alexandre Chotard, Nikolaus Hansen, Ilya Loshchilov, Yann Ollivier, Marc Schoenauer, Michèle Sebag, Olivier Teytaud.

Our expertise in continuous optimization is focused on stochastic search algorithms. We address theory, algorithm design, and applications. The methods we investigate are adaptive techniques that are able to learn iteratively the parameters of the distribution used to sample (new) solutions. The Covariance Matrix Adaptation Evolution Strategy (CMA-ES) is nowadays one of the most powerful methods for derivative-free continuous optimization. We work on different variants of the CMA-ES to improve it in various contexts as described below. We have previously proven the convergence of simplified variants of the CMA-ES algorithm using the theory of stochastic approximation, and have provided the first proofs of convergence on composite of twice continuously differentiable functions. More recently, we used Markov chain analysis for analyzing the step-size adaptation rules of evolution strategies related to the CMA-ES algorithm.

Surrogate models for CMA-ES. In the context of his PhD thesis defended in January 2013 [4], Ilya Loshchilov has proposed different surrogate variants of CMA-ES based on ranking-SVM that preserve the invariance to monotonic transformation of the CMA-ES algorithm. As a follow-up, he has proposed an original over-exploitation mechanism in case of accurate surrogate [44]. Several of these models have entered the BBOB-2013 workshop [43]. Further research direction using the KL divergence between successive distributions as a trigger for a new learning phase has been proposed [45].

Step-size adaptive methods. We have proposed a new step-size adaptation mechanism that can loosely be interpreted as a new variant of the 1/5 success rule for comma (non-elitist) strategies and which is applicable with a large population size [21]. The rule uses the success of the median fitness of the current population compared to a (different) fitness percentile from the previous population.

Principles of Stochastic Optimization. Based on the framework of *information geometry* (IGO), theoretical guarantees have been obtained for continuous optimization algorithms: using the *natural gradient* provides improvement guarantees even when using finite step sizes [22]. We have considered the principles of designing effective stochastic optimization algorithms from the bottom-up and the top-down perspective [56]. The top-down perspective takes the information-geometrical view-point and largely confirms the bottom-up construction.

Benchmarking. We have continued our effort for improving standards in benchmarking and pursued the development of the COCO (COmparing Continuous Optimizers) platform (see Section 5.4). We have organized the ACM GECCO 2013 workshop on Black-Box-Optimization Benchmarking [2] and benchmarked different surrogate-based variants of the CMA-ES algorithm [26], [44], [43]. Our new starting ANR project NumBBO, centered on the COCO platform, aims at extending it for large-scale, expensive, constrained and multi-objective optimization.

Theoretical proofs of convergence. We have established the connection between convergence of comparison based step-size adaptive randomized search and the stability analysis of some underlying Markov chains. This connection heavily exploits invariance properties of the algorithm. In a first paper we establish this connection for scaling-invariant functions and prove sufficient conditions for linear convergence expressed in terms of stability conditions [63]. We have proven, using this defined methodology, the linear convergence of a famous algorithm introduced independently by several researchers and known as the (1+1)-ES with one-fifth success rule [62]. In [32], we have proven the linear convergence of a modified evolutionary algorithm without assuming quasi-convexity.

[2] see http://coco.gforge.inria.fr/doku.php?id=bbob-2013

6.2. Optimal Decision Making under Uncertainty

Participants: Olivier Teytaud [correspondent], Jean-Joseph Christophe, Adrien Couëtoux, Jérémie Decock, Nicolas Galichet, Manuel Loth, Marc Schoenauer, Michèle Sebag.

The UCT-SIG works on sequential optimization problems, where a decision has to be made at each time step along a finite time horizon, and the underlying problem involves uncertainties along an either adversarial or stochastic setting.

The most prominent application domain is now energy management, at various time scales, and more generally, planning in uncertain environments. The main advances done this year include:

- A work on metagaming/investment [12], where a macroscopic decision has to be made (e.g., investment decisions, which plants should be built) prior to operational decisions (e.g., unit commitment policy, i.e., the operational management of the system). This is a key part of our activity for 2014.
- Bandit problems with risk [36]. Bandit problems are quite related to metagaming problems (they correspond to the unstructured case).
- A theoretical work on the consistency of Monte Carlo Tree Search / Upper Confidence Tree in continuous domains [27]. A non-trivial extension was necessary for proving such a consistency.
- Noisy optimization is a key part of our work [61], as it is crucial for direct policy search or more generally for dynamic optimization:
  - We have proven lower bounds under "locality assumptions" [33], which are usually informally assumed by some practicionners for justifying the use of evolutionary algorithms.
  - In cases with strong noise (variance not decreasing to zero around the optimum) we proved log-log convergence for simple rules for choosing the number of resamplings [23].
- Several submissions are joint works with Ailab, National Dong Hwa University, Hualien, Taiwan. The drafts can be found at http://www.lri.fr/~teytaud/indema.html.
- In collaboration with Christian Shulte (KTH, Stockholm), one of the main contributors to the well-known general-purpose CP solver *GECODE* (http://www.gecode.org/), and within the Microsoft-Inria joint lab Adapt project, ideas from UCT have been integrated in GECODE for the choice of the variable values during the exploration of the constraint tree. The most critical issue lied in the definition of a meaningful reward for a given node (variable = value) that could cope with the multiple restarts of the search: the deeper the failure, the larger the reward (and hence this work also pertains to the CRI-SIG(Section 6.4). Initial results have been obtained with job-shop scheduling problems [47] and more extensive results have been obtained on 3 benchmarks of the CP community [46].

6.3. Distributed Systems

Participants: Cécile Germain-Renaud [correspondent], Philippe Caillou, Dawei Feng, Cyril Furtlehner, Victorin Martin, Michèle Sebag.

The DIS-SIG explores the issues related to modeling and optimizing distributed systems, ranging from very large scale computational grids to multi-agent systems and large scale traffic management.

Fault management. As Lamport formulated decades ago, fault management in distributed systems exemplifies the unreachability of exact prior knowledge. Real-world large scale systems additionally face the non-stationarity issue.

[20] models the system state and its ruptures (non-stationarity) through the flow of jobs as a stream (scalability), with a traceability goal (interpretability), and addresses a key difficulty in Data Streaming, which is timely detection of a change in the generative process underlying the data stream *drift*. A statistical model based on spatial distance and time frequency is proposed, together with adaptive thresholding. Theoretical and experimental validation show the robustness of the method.

D. Feng's PhD formulates the problem of probe selection for fault prediction based on end-to-end probing as a Collaborative Prediction (CP) problem, based on the reasonable assumption of an underlying factorial model [13]. Monitoring large scale distributed systems differ from CP's usual applications (personalized recommendation), in two major ways. On the brighter side, while users cannot be queried for specific recommendations, probes can be launched at will. On the downside, firstly the distribution of the probe results is highly skewed, faults being a small fraction of the total population, and secondly, some of the faults are transient. Amongst the numerous approaches addressing CP, Minimum Margin Matrix Factorization (MMMF) is easily amenable to active learning, which addresses fault sparsity both at spatial (skewed distributions) and temporal (transients) level. From extensive experiments on real-world data, we have shown that modelling probe-based fault prediction as a CP task and addressing this task through MMMF is an extremely efficient strategy for fault prediction. Comparative analysis and experiments motivate the critical advantage of active learning. It offers a scalable alternative to direct AUC optimization. Similarly, comparison with bias-aware methods (Mixed Membership Matrix Factorization) indicates that the capacity of actively selecting the most informative probes provides the most efficient method to capture the time variability of the system.

Multi-agent and games. Resuming earlier work, our goal is to provide an automated abstract description of simulation results. Data mining methods are used to identify groups in complex simulations [11]. Using activity indicators to identify the most interesting agent groups [17], the groups and their evolution are described through one or several simulations [10]. To facilitate the dissemination of the algorithms, we have participated in the development of a generic multiagent platform (GAMA), in collaboration with IRD, University of Rouen (IDEES), and University of Toulouse (IRIT) [34], [35].

A statistical physics perspective. With motivating applications in large scale traffic congestion inference problems, we have

- Settled a method for encoding real data with pairwise dependencies into an Ising copula [58] suitable to infer real-valued data (travel time) from the computation of its corresponding latent binary state (congested/not congested) probabilities.
- In parallel we have investigated the inverse Ising problem, proposing among other methods a loop analysis based on a duality transformation, leading to a dual belief propagation algorithm runing on the dual graph formed by the network of independent cycles. This aims at finding an MRF to represent pairwise correlated data, close to a dependence tree, able to take into account most important loops [14], [15].

6.4. Designing Criteria

Participants: Jamal Atif, Yoann Isaac, Mostepha Redouane Khouadjia, Hélène Paugam-Moisy, Marc Schoenauer, Michèle Sebag.

This recently created SIG, rooted on the claim that *What matters is the criterion*, aims at defining new learning or optimization objectives reflecting fundamental properties of the model, the problem or the expert prior knowledge.

Image understanding. We continued our effort on the development of model-based image understanding approaches (e.g. [71]). In [18], we have proposed a method for simultaneously segmenting and recognizing objects in images, based on a structural representation of the scene and on a constraint propagation method. Theoretical guaranties have been provided along with a quantitative assessment on healthy and pathological brain structures in magnetic resonance images. Within the ANR project LOGIMA (collaboration with ECP, Telecom PariTech and TU Dresden), our goal is to introduce a new lattice-based representation and reasoning framework suited for dealing with spatial objects in the presence of uncertainty. This framework associates under the aegis of general lattice theory ingredients from mathematical morphology, description logics and formal concept analysis. A first development of this framework can be found in [7] where it has been exploited for the definition of abductive reasoning services and applied to high-level image understanding. Several theoretical issues have been raised in the development of this new framework. Some of them were tackled in [25], [24], [30]. In [25], we have shown how mathematical morphology operators can ben defined on general concept lattices. Explicit join-commuting and meet-commuting operators are defined either from particular valuations on the corresponding lattice or from the decomposition of their elements. In [24], we extended our mathematical morphology based adductive reasoning to multivalued logics, hence allowing us to deal with several uncertainty and imprecision phenomena. In [30], metrics between bipolar information - where the information is represented by a positive/preference part and a negative/constraint part - have been introduced based on particular dilations.

Structured learning. With motivations in bio-informatics and brain computer interfaces, the goal is to take into account priors about the spatio-temporal structure of the underlying phenomenon in order to propose a generative model of the data.

In the context of Yoann Isaac's PhD (Digiteo Unsupervised Brain project), in collaboration with CEA LIST, the goal is to design a representation endowed with appropriate invariance properties. Specifically, within the framework of sparse dictionary coding, we have introduced new priors allowing us to capture both spatial and temporal regularity of multivariate brainwave signals [54]. The learning/optimization criterion, while being multivariate, contains several non-differentiable terms, raising new optimization issues; the proposed approach extends the classical split Bregman iterations algorithm to the multi-dimensional case with several non-differentiable terms [37].

In the context of regulatory gene networks, the challenge is to combine probabilistic inference (does a gene regulate another one) with relational learning (the set of genes is organized in a network). Ensemble learning approaches have been used to cope with the imbalanced nature of the data, e.g., bagging Markov logic networks or boosting operator-valued kernel-based regressors [55], [64]. Another issue, regarding the indeterminacy of the models due to the data sparsity, is addressed through prior-guided regularization beyond model sparsity such as orthogonality [8] or stability [16].

In the domain of medical imaging, the exploitation of computational tomography data to model tumor physiology is hindered by the huge noise level; the multi-task setting is leveraged to provide a better robustness to noise [51].

Robotic value systems. Within the European SYMBRION IP, investigations on the preference-based reinforcement learning were continued in Riad Akrour's PhD, where the robot demonstrations are assessed by the expert and these assessments are used to learn a model of the expert's expectations. In [67], this work had been extended and combined with active learning to yield state-of-the art performances with few binary feedbacks from the expert. The work has first concentrated this year on handling the noise due to expert's mistakes [53], and bridging the reality gap when porting the algorithms on real robots (e-pucks and one Nao robot) – these results will be published in Riad Akrour's PhD dissertation, to be defended in Spring 2014.

Algorithm Selection as Collaborative Filtering. The crucial issue when addressing algorithm selection problems is to be able to come up with features that can describe the problems: with representative features, algorithm selection amounts to supervised learning. However, except for some rare domaines (e.g., SAT, [73]), no satisfying set of features exists. However, algorithm selection can also be viewed as a recommendation problem, and tackled by collaborative filtering: users more or less like movies, and similarly, instances like algorithms as much as these algorithms are efficient in solving it. Applying collaborative filtering leads to designing a latent feature space in which the representation of the problems is highly adapted to the algorithm selection problem. A critical issue in collaborative filtering is the 'cold start' problem, that is making recommendations for a brand new user/problem instance. This issue has been handled by a surrogate model of the latent factors, mapping the initial features onto the latent ones. The *Algorithm Recommender System* has been successfully applied to 3 different domains: Satisfiability, Constraint Programming, and Continuous Black-Box Optimization (data from the COCO platform, see Section 5.4) [59].

Social Networks with insider information. The analysis of social networks based on the contents and structure of information exchanges most often pertains to descriptive learning, e.g., explaining the growth of the communication graph or investigating the sensitivity of existing algorithms to hyper-parameters [31]. In the Modyrum context (coll. SME Augure), a supervised learning perspective is investigated, taking advantage of the fact that experts already know part of the sought results in some specific domains of interest. Based on e.g., Twitter and blogs data, the goal is to define generic features and supervised learning algorithms, enabling to characterize the targets of interest depending on the public relation focus.

Multi-objective AI Planning. Within the ANR project DESCARWIN (http://descarwin.lri.fr), Mostepha-Redouane Kouadjia continued his work on the **multi-objective approach** to AI Planning using the Evolutionary Planner *Divide-and-Evolve* (DaE), that evolves a sequential decomposition of the problem at hand: each sub-problem is then solved in turn by some embedded classical planner [70]. Even though the embedded planner is single-objective, DaE can nevertheless handle multi-objective problems. Current work includes the implementation of the multi-objective version of DaE, the definition of some benchmark suite, and some first numerical experiments, comparing in particular the results of a full Pareto approach to those of the classical aggregation method. These works resulted in 3 conference papers recently accepted, introducing a tunable benchmark test suite [39], demonstrating that the best quality measure for parameter tuning in this multi-objective framework is the hypervolume, even in the case of the aggregation approach [41], and comparing the evolutionary multi-objective approach with the aggregation method, the only method known to the AI Planning community [38]. A sum-up of these recent results have been published at IJCAI [40].

TASC Project-Team

6. New Results

6.1. Solvers

Participants: Nicolas Beldiceanu, Rémi Douence, Narendra Jussien, Xavier Lorca, Eric Monfroy, Charles Prud'Homme.

- [14] presents some research directions wrt sustainable solver development based on the idea that solvers should be based/derived on data bases of combinatorial knowledge.
- [19] and [42] presents a solver independent language dealing both with variable-oriented and constraint-oriented propagation engines to enable the design of propagation engines.
- By observing the resolution process, [35] shows how to dynamically adapt the resolution while propagating constraints.

6.2. Filtering

Participants: Nicolas Beldiceanu, Alban Derrien, Jérémie Du Boisberranger, Jean-Guillaume Fages, Arnaud Letort, Xavier Lorca, Thierry Petit, Charlotte Truchet, Mohamed Wahbi.

- Given a matrix model, with the same constraint defined by a finite-state automaton on each row and a global cardinality constraint on each column, [12] exploits double counting to derive necessary conditions on the cardinality variables of the global cardinality constraints from the automata. (participants: Beldiceanu)
- By using the observation that most global constraints can be reformulated as a conjunction of a total function constraint together with a constraint that can be easily reified (e.g. a linear constraint involving two variables), [13] introduces a simple way for deriving reified global constraints. (participants: Beldiceanu)
- In the context of distributed constraint solving [22], [25] introduce two filtering algorithms that extend Asynchronous Forward Checking (AFC). The last one outperforms AFC specially on sparse problems. (participants: Wahbi)
- We improve the energetic reasoning checker of the cumulative constraint by decreasing the number checked intervals by a factor seven. We prove this approach can be generalized to the ER filtering algorithm. Furthermore, in a context of makespan minimization of hard problems, our experiments demonstrate that associating this checker with a Time-Table propagator is more efficient than using the best state-of-the-art propagators, such as Time-Table Edge-Finding. This work is at the core of Alban Derrien's doctoral research (Alban is a PhD student of TASC). It was published at the doctoral program of CP2013 [28]. (participants: Derrien, Petit)
- [29] introduces a probabilistic model for the bound consistency algorithm of the alldifferent constraint in order to decrease the number of times the constraint is woken without making new deductions during constraint propagation. (participants: Du Boisberranger, Lorca, Truchet).
- Initially motivated by the shift minimisation personal task scheduling problem [30] shows how to integrate difference constraints into the AtMostNValue constraint in order to get a better estimation about the minimum number of distinct values. (participants: Fages, Lorca)
- Motivated by scalability issues, and based on the idea of accelerating the convergence to the fix-point by filtering several cumulative constraints in parallel, [33] and [47] presents a sweep based algorithm for a conjunction of cumulative constraints. (participants: Beldiceanu, Letort)
- [46] come up with a more efficient filtering algorithm than the one introduced for the cost regular constraint for dealing with constraints for which the set of solutions can be represented by an automaton with counters. (participants: Beldiceanu)

6.3. Continuous/discrete

Participants: Nicolas Beldiceanu, Gilles Chabert, Jean-Guillaume Fages, Charles Prud'Homme.

- While convexity (and some of its generalisations) is a key property used for dealing with continuous constraints it was not yet used in the context of discrete global constraints. In [34] we come up with a parametric filtering algorithm based on a form of convexity. It can handle in a uniform way various constraints such as deviation, spread or the conjunction of a linear inequality constraint and count constraint.
- Motivated by hybrid discrete continuous problems we come up in [44] with a simple and efficient interface for connecting a discrete constraint solver (Choco) and a continuous constraint solver (Ibex).

6.4. Learning Constraint Models

Participants: Nicolas Beldiceanu, Naina Razakarison.

- In the context of learning parametrized constraint models for highly structured problems we address in [38] the problem of finding the coefficients of polynomials in several variables from example parameter and function values.
- In the context of semi structured time series where the structural aspect is related to technological constraints we deal in [24] with the problem of extracting functional dependency constraints. The problem is motivated by extracting constraints from electricity production time series and is characterized by a larger set of samples (from 7 years and from 300 plants).

6.5. Meta Heuristics

Participants: Alejandro Reyes Amaro, Eric Monfroy, Florian Richoux, Charlotte Truchet.

- The aim is to develop and implement new algorithmical methods for constraint problems on massively parallel machines. We also conduct more theoretical studies about the parallelization of constraint problems. This year, we proposed a fairly sharp model to predict parallel speed-ups one can expect while parallelizing by a multi-walk parallel scheme any Las Vegas algorithm by just studying the distribution of sequential run-times [41]. This model shows a divergence of only 20% when predicting speed-ups over 256 cores, on very different benchmarks.
- To evaluate the scalability and parallelization of local search algorithms for SAT, [23] presents a statistical method based on the analysis of the runtime behavior of its sequential version.
- [26] and [26] deals with the use of metaheuristics for solving the resource constrained scheduling problem and the set covering problems.

6.6. Search and modelling

Participants: Eric Monfroy, Thierry Petit.

- In the context of autonomous search, [21] deals with the problem of automatically tuning a search strategy (i.e., variable value selection). For this purpose it uses so called *choice functions* which provide an evaluation of a strategy in term of a set of indicators. [36] and [31] go one step further by providing tuning and adaptation facilities at the level of the different components of a constraint solver.
- Using the MiniZinc modeling language, [32] shows how to model and solve the portfolio selection problem with constraint programming. Since more than ten year constraints for which the set of solutions can be matched to the language accepted by an automaton were introduced in many solvers (e.g., Choco, Gecode, SICStus). [40] describes an interface for describing such constraints in a more convenient way.
- Many discrete optimization problems have constraints on the objective function. Being able to represent such constraints is fundamental to deal with many real world industrial problems. In this work, we go one step further in the concept of topologically concentrate high values in a sequence of cost variables. We refine the work we previously published in CP2012 thanks to three generalizations of the focus constraint. We experiment successfully the technique in scheduling, round-robin and musical benchmarks. This work has been published at IJCAI 2013 [37].

6.7. Miscellaneous

Participants: Eric Monfroy, Florian Richoux.

- [15] gives a complete characterization of the complexity of the existential positive first-order logic, that one can interpret as model checking on monotone csp. We exhibit a dichotomy criterion remaining the same on finite domains of every cardinality, as well as countable and uncountable infinite domains.
- We develop an artificial intelligence, AIUR, to play the real time strategy game $StarCraft^{tm}$, using both machine learning and constraint-based techniques. AIUR finished 3rd to $StarCraft^{tm}$ AI competitions organized at the conferences AIIDE 2013 and CIG 2013. [18] presents a survey on AI techniques applied on $StarCraft^{tm}$.

TEXMEX Project-Team

6. New Results

6.1. Description of multimedia content

6.1.1. Multiscale image representations with component trees

Participants: Petra Bosilj, Ewa Kijak.

Joint work with Sébastien Lefevre, IRISA/SEASIDE, France.

The goal of this work is to study deeply the use of component trees, which aim at representing an image by the regions it contains at various scales through a tree-based structure, and their ability in the context of content-based image indexing and retrieval. Their invariance properties and their robustness to noise have motivated recent work in image indexing [83], [97], [98], but their usage in this field stays limited. The first part of this work was mainly dedicated to the study of various existing hierarchical representations. This leads to the presentation of a technique that arranges the elements of hierarchical representations of images according to a coarseness attribute [24]. The transformation is similar to filtering a hierarchy with a non-increasing attribute, and includes the results of multiple simple filterings with an increasing attribute. The transformed hierarchy can be then used for search space reduction prior to the image analysis process because it allows for direct access to the hierarchy elements at the same scale or a narrow range of scales.

6.1.2. Image representation

Participants: Rachid Benmokhtar, Jonathan Delhumeau, Guillaume Gravier, Philippe-Henri Gosselin, Hervé Jégou, Wanlei Zhao.

Partially in collaboration with Patrick Pérez, Technicolor, France.

Recent work on image retrieval have proposed to index images by compact representations encoding powerful local descriptors, such as the closely related vector of aggregated local descriptors (VLAD) and Fisher vector (FV). By combining them with a suitable coding technique, it is possible to encode an image in a few dozen bytes while achieving excellent retrieval results. We have pursed the research on this line of research by proposing two complementary contributions.

In [30], we revisited some assumptions proposed in this context regarding the handling of ”visual burstiness”, and shows that ad-hoc choices are implicitly done which are not desirable. Focusing on VLAD without loss of generality, we propose to modify several steps of the original design. Albeit simple, these modifications significantly improve VLAD and make it compare favorably against the state of the art.

In [65], we proposed a pooling strategy for local descriptors to produce a vector representation that is orientation-invariant yet implicitly incorporates the relative angles between features measured by their dominant orientation. This pooling is associated with a similarity metric that ensures that all the features have undergone a comparable rotation. This approach is especially effective when combined with dense oriented features, in contrast to existing methods that either rely on oriented features extracted on key points or on non-oriented dense features. The interest of our approach in a retrieval scenario is demonstrated on popular benchmarks comprising up to 1 million database images.

In [22], we propose to reduce the dimensionality of visual features for image categorization. We iteratively select sets of projections from an external dataset, using Bagging and feature selection thanks to SVM normals. Features are selected using weights of SVM normals in orthogonalized sets of projections. The bagging strategy is employed to improve the results and provide more stable selection. The overall algorithm linearly scales with the size of features, and is thus able to process large state-of-the-art image representations. Given Spatial Fisher Vectors as input, our method consistently improves the classification accuracy for smaller vector dimensionality, as demonstrated by our results on the popular and challenging PASCAL VOC 2007 benchmark.

6.1.3. Video classification

Participants: Kleber Jacques Ferreira de Souza, Guillaume Gravier, Philippe-Henri Gosselin.

In collaboration with Silvio Jamil F. Guimarães, PUC Minas, Brazil.

Most current motion descriptors for video classification are based on simple video segments, such as rectangular space-time blocks, or more recently rectangular space blocks that follow local trajectories. The aim of this study is to consider more complex video segments that better fit space-time elements of videos, thanks to recent methods for video segmentation proposed by S. Guimarães et al. These methods combine at the same time a fast extraction and stable regions, two essential properties for video indexing. The computation of local motion descriptors on these video segments lead to better video classification for human action recognition, when compared to current video indexing techniques.

6.1.4. Geo-localization of videos with multi-modality

Participants: Jonathan Delhumeau, Guillaume Gravier, Hervé Jégou.

Joint work with Michele Trevisiol, Yahoo! Labs, Spain, who visited the team in 2012.

Geotagging is the process of automatically adding geographical identification metadata to media objects, in particular to images and videos. In [63], we present a strategy to identify the geographic location of videos. First, it relies on a multi-modal cascade pipeline that exploits the available sources of information, namely the user upload history, his social network and a visual-based matching technique. Second, we present a novel divide & conquer strategy to better exploit the tags associated with the input video. It pre-selects one or several geographic area of interest of higher expected relevance and performs a deeper analysis inside the selected area(s) to return the coordinates most likely to be related to the input tags. The experiments were conducted as part of the MediaEval 2012 Placing Task, where we obtained the best results among the competitors when using no external information, i.e. not using any gazetteers nor any other kind of external information.

6.1.5. Violent keysound detection with audio words and Bayesian networks

Participants: Guillaume Gravier, Patrick Gros, Cédric Penet.

Joint work with Claire-Hélène Demarty, Technicolor, France.

We investigated a novel use of the well known audio words representations to detect specific audio events, namely gunshots and explosions, in order to get more robustness towards soundtrack variability in Hollywood movies [51]. An audio stream is processed as a sequence of stationary segments. Each segment is described by one or several audio words obtained by applying product quantization to standard features. Such a representation using multiple audio words constructed via product quantisation is one of the novelties described in this work. Based on this representation, Bayesian networks are used to exploit the contextual information in order to detect audio events. Experiments are performed on a comprehensive set of 15 movies, made publicly available. Results are comparable to the state of the art results obtained on the same dataset but show increased robustness to decision thresholds, however limiting the range of possible operating points in some conditions. Late fusion provides a solution to this issue.

6.2. Large scale indexing and classification

6.2.1. Parallelism and distribution for very large scale content-based image retrieval

Participants: Gylfi Gudmundsson, Diana Moise, Denis Shestakov, Laurent Amsaleg.

Two observations drove the design of the high-dimensional indexing technique developed in the framework of the Ph. D. thesis of Gylfli Gudmundson. Firstly, the collections are so huge, typically several terabytes, that they must be kept on secondary storage. Addressing disk related issues is thus central to our work. Secondly, all CPUs are now multi-core and clusters of machines are a commonplace. Parallelism and distribution are both key for fast indexing and high-throughput batch-oriented searching.

We developed a high-dimensional indexing technique called eCP. Its design includes the constraints associated to using disks, parallelism and distribution. At its core is an non-iterative unstructured vectorial quantization scheme. eCP builds on an existing indexing scheme that is main memory oriented. The first contribution in eCP is a set of extensions for processing very large data collections, reducing indexing costs and best using disks. The second contribution proposes multi-threaded algorithms for both building and searching, harnessing the power of multi-core processors. Datasets for evaluation contain about 25 million images or over 8 billion SIFT descriptors. The third contribution addresses distributed computing. We adapt eCP to the MapReduce programming model and use the Hadoop framework and HDFS for our experiments. This time we evaluate eCP's ability to scale-up with a collection of 100 million images, more than 30 billion SIFT descriptors, and its ability to scale-out by running experiments on more than 100 machines.

6.2.2. Contributions in image indexing

Participants: Hervé Jégou, Giorgos Tolias.

Partially in collaboration with Yannis Avrithis, National Technical University of Athens, Greece, Cai-Zhi Zhu and Shin'ichi Satoh, Japanese National Institute of Informatics, Japan.

In [62], we have considered a framework and its associated family of metrics to compare images based on their local descriptors. It encompasses the VLAD descriptor and matching techniques such as Hamming embedding. Making the bridge between these approaches leads us to propose a match kernel that takes the best of existing techniques by combining an aggregation procedure with a selective match kernel. Finally, the representation underpinning this kernel is approximated, providing a large scale image search both precise and scalable, as shown by our experiments on several benchmarks. We give a Matlab package associated with the paper that allows to reproduce the results of the most interesting variant.

On the same topic, we propose in [78] a query expansion technique for image search that is faster and more precise than the existing ones. An enriched representation of the query is obtained by exploiting the binary representation offered by the Hamming embedding image matching approach: The initial local descriptors are refined by aggregating those of the database, while new descriptors are produced from the images that are deemed relevant. This approach has two computational advantages over other query expansion techniques. First, the size of the enriched representation is comparable to that of the initial query. Second, the technique is effective even without using any geometry, in which case searching a database comprising 105k images typically takes 79 ms on a desktop machine. Overall, our technique significantly outperforms the visual query expansion state of the art on popular benchmarks. It is also the first query expansion technique shown effective on the UKB benchmark, which has few relevant images per query.

Finally, in [67] we considered a problem related to object retrieval, where we aim at retrieving, from a collection of images, all those in which a given query object appears. This problem is inherently asymmetric: the query object is mostly included in the database image, while the converse is not necessarily true. However, existing approaches mostly compare the images with symmetrical measures, without considering the different roles of query and database. This paper first measures the extent of asymmetry on large-scale public datasets reflecting this task. Considering the standard bag-of-words representation, we then propose new asymmetrical dissimilarities accounting for the different inlier ratios associated with query and database images. These asymmetrical measures depend on the query, yet they are compatible with an inverted file structure, without noticeably impacting search efficiency. Our experiments show the benefit of our approach, and show that the visual object retrieval task is better treated asymmetrically, in the spirit of state-of-the-art text retrieval.

6.2.3. Outlier detection applied to content-based image retrieval

Participants: Teddy Furon, Hervé Jégou.

The primary target of content based image retrieval is to return a list of images that are the most similar to a query image, which is usually done by ordering the images based on a similarity score. In most state-of-the-art systems, the magnitude of this score is very different from one query to another. This prevents us from making a proper decision about the correctness of the returned images. Our work [74] considers the applications where a confidence measurement is required, such as in copy detection or when a re-ranking stage is applied on a short-list such as geometrical verification. For this purpose, we formulate image search as an outlier detection problem, and propose a framework derived from extreme values theory. We translate the raw similarity score returned by the system into a relevance score related to the probability that a raw score deviates from the estimated model of scores of random images. The method produces a relevance score which is normalized in the sense that it is more consistent across queries. Experiments performed on several popular image retrieval benchmarks and state-of-the-art image representations show the interest of our approach.

6.2.4. Exploiting motion characteristics for action classification in videos

Participants: Mihir Jain, Hervé Jégou.

In collaboration with Patrick Bouthemy, Inria/Serpico, France.

Several recent studies on action recognition have attested the importance of explicitly integrating motion characteristics in video description. In this work [43], we have re-visited the use of motion in videos, in order to better exploit it and improve action recognition systems. First, we established that adequately decomposing visual motion into dominant and residual motions, both in the extraction of the space-time trajectories and for the computation of descriptors, significantly improves action recognition algorithms. Then, we designed a new motion descriptor, the DCS descriptor, based on differential motion scalar quantities, divergence, curl and shear features. It captures additional information on the local motion patterns enhancing results. Finally, applying the recent VLAD coding technique proposed in image retrieval provides a substantial improvement for action recognition. Our three contributions are complementary and lead to significantly outperform all reported results on three challenging datasets, namely Hollywood 2, HMDB51 and Olympic Sports.

6.2.5. Recognizing events in videos

Participant: Hervé Jégou.

In collaboration with Matthijs Douze, Jérôme Revaud and Cordelia Schmid, Inria/LEAR, France.

We have addressed the problem of event retrieval for large-scale video collection. Given a video clip of a specific event, e.g., the wedding of Prince William and Kate Middleton, the goal is to retrieve other videos representing the same event from a dataset of over 100k videos.

Our first approach [55] encodes the frame descriptors of a video to jointly represent their appearance and temporal order. It exploits the properties of circulant matrices to compare the videos in the frequency domain. This offers a significant gain in complexity and accurately localizes the matching parts of videos. Furthermore, we extend product quantization to complex vectors in order to compress our descriptors, and to compare them in the compressed domain. Our method outperforms the state of the art both in search quality and query time on two large-scale video benchmarks for copy detection, Trecvid and CCweb. The evaluation has also been done on a new challenging dataset for event retrieval that we introduce: EVVE.

In a subsequent paper [39], we have made two other contributions to event retrieval in large collections of videos. First, we propose hyper-pooling strategies that encode the frame descriptors into a representation of the video sequence in a stable manner. Our best choices compare favorably with regular pooling techniques based on k-means quantization. Second, we introduce a technique to improve the ranking. It can be interpreted either as a query expansion method or as a similarity adaptation based on the local context of the query video descriptor. Experiments on public benchmarks show that our methods are complementary and improve event retrieval results, without sacrificing efficiency.

6.2.6. Large-scale SVM image classification

Participants: Thanh Nghi Doan, François Poulet.

Visual recognition remains an extremely challenging problem in computer vision research. Large datasets with millions images for thousands categories poses more challenges. We extend the state-of-the-art large scale linear classifier LIBLINEAR SVM and nonlinear classifier Power Mean SVM in two ways. The first one is to build a balanced bagging classifier with sampling strategy. The second one is to parallelize the training process of all binary classifiers with several multi-core computers [35]. We also applied the same approach to the stochastic gradient descent support vector machines (SVM-SGD) and to both state-of-the-art large linear classifier LIBLINEAR-CDBLOCK and nonlinear classifier Power Mean SVM in an incremental and parallel way [36].

6.2.7. Video copy detection with SNAP, a DNA indexing algorithm

Participants: Laurent Amsaleg, Guillaume Gravier.

In collaboration with Leonardo S. De Oliveira, Zenilton Kleber G. Do Patrocínio Jr. and Silvio Jamil F. Guimarães, PUC Minas, Brazil.

Near-duplicate video sequence identification consists in identifying real positions of a specific video clip in a video stream stored in a database. To address this problem, we proposed a new approach based on a scalable sequence aligner borrowed from proteomics [79]. Sequence alignment is performed on symbolic representations of features extracted from the input videos, based on an algorithm originally applied to bio-informatics. Experimental results demonstrate that our method performance achieved 94 % recall with 100 % precision, with an average searching time of about 1 second.

6.3. Security of multimedia contents and applications

6.3.1. Approximate nearest neighbors search with security and privacy requirements

Participants: Benjamin Mathon, Laurent Amsaleg, Teddy Furon.

In collaboration with Julien Bringer, Morpho, France.

This work presents a moderately secure but highly scalable and fast approximate nearest neighbors search. Our philosophy is to start from a state-of-the-art technique in this field based on approximate metrics: Euclidean distance based search in [47], [70], and cosine similarity based search in [42]. We then analyze the threats, and patch them avoiding as much as possible bricks penalizing too much the scalability and the speed. On the other hand, we do not completely prevent the players to infer some knowledge, but these limitations are well explained and experimentally assessed. The experimental body uses database of size much bigger than what the past secure solutions can handle.

6.3.2. A privacy-preserving framework for large-scale content-based information retrieval

Participants: Ewa Kijak, Laurent Amsaleg, Teddy Furon.

In close cooperation with Stéphane Marchand-Maillet, Li Weng and April Morton, University of Geneva, Switzerland.

We propose a privacy protection framework for large-scale content-based information retrieval. It offers two layers of protection. First, robust hash values are used as queries instead of original content or features. Second, the client can choose to omit certain bits in a hash value to further increase the ambiguity for the server. Due to the reduced information, it is computationally difficult for the server to know the client's interest. The server has to return the hash values of all possible candidates to the client. The client performs a search within the candidate list to find the best match. Since only hash values are exchanged between the client and the server, the privacy of both parties is protected.

We introduce the concept of *tunable privacy*, where the privacy protection level can be adjusted according to a policy. It is realized through hash-based piece-wise inverted indexing. The idea is to divide a feature vector into pieces and index each piece with a sub-hash value. Each sub-hash value is associated with an inverted index list.

The framework has been extensively tested using a large image database. We have evaluated both retrieval performance and privacy-preserving performance for a particular content identification application. Two different constructions of robust hash algorithms are used. One is based on random projections; the other is based on the discrete wavelet transform. Both algorithms exhibit satisfactory performance in comparison with state-of-the-art reference schemes. The results show that the privacy enhancement slightly improves the retrieval performance.

We consider the *majority voting attack* for estimating the query category and ID. Experiment results show that this attack is a threat when there are near-duplicates, but the success rate decreases with the number of omitted bits and the number of distinct items.

6.3.3. Privacy preserving data aggregation and service personalization using highly-scalable indexing techniques

Participants: Raghavendran Balu, Laurent Amsaleg, Hervé Jégou, Teddy Furon.

In collaboration with Armen Aghasaryan, Dimitre Davidov and Makram Bouzid, Alcatel-Lucent, and Sébastien Gambs, Inria/CIDRE, in the framework of the Alcaltel-Lucent / Inria common Lab.

A challenging approach to the problem of privacy preserving data aggregation and service personalization has recently been proposed in Bell Labs, which introduces a privacy-preserving intermediation layer between end-users and service providers. It uses a distributed variant of a Locality Sensitive Hashing (LSH) techniques of doing scalable nearest-neighbor search, adapted in a novel way, to discover similar users while preserving their privacy. This approach faces however several important challenges that will be targeted in the scope of this collaboration. The challenges are:

- *LSH optimization:* Definitions of hash functions as well as various LSH parameters need to be automatically tuned in order to achieve a good quality of generated recommendations with an expected level of the procured user anonymity. An interesting issue is the possibility of supervised machine learning. If some public profiles are available, more efficient clustering methods boost the quality of the recommendation service but their levels of anonymity have never been assessed so far.
- *Irreversibility of anonymization:* This needs to be evaluated for different attack models, e.g. exploiting the knowledge of LSH hashing functions or any other publically available information on users. It is equivalent as being able to define the region of the super high-dimensional space mapped into the same hashing results. This attack is bound to fail as this region is too large to leak information. However, the prior knowledge about the sparseness of the profiles might drastically reduce this region, and hence weaken the privacy.
- *System dynamics:* Dealing with the cold-start problem or controlling the dynamics of a running system when the profiles and the cluster assignments evolve over the time is yet another challenge this approach is confronted with. If these temporal issues are well studied in conventional relational databases, no clear solution is efficient in the recommendation area, and a fortiori in privacy enhancing recommendation systems.

6.4. Structuring multimedia content and summarization

6.4.1. Stream labeling for TV Structuring

Participants: Vincent Claveau, Guillaume Gravier, Patrick Gros, Emmanuelle Martienne, Abir Ncibi.

In this application, we focus on the problem of labeling the segments of TV streams according to their types (eg. programs, commercial breaks, sponsoring...). During this year, following the work initiated in 2012, we have proposed an in-depth analysis of the use of conditional random fields (CRF) for our task [50]. Through several experiments conducted on real TV streams, we have shown that the CRF yields high results compared with state-of-the-art approaches. In particular, CRF offers several ways to efficiently take the sequenciality of our stream labeling problem into account. We also showed that it is robust when dealing with few training data or few features.

6.4.2. Statistical tests for repetition detection in TV streams

Participant: Patrick Gros.

Detecting all repeated sequences in a TV stream is the first step of all techniques of TV stream structuring. We have improved our technique in several ways. First, a statistical hypothesis test with a corrected risk of Bonferroni was used to clean the repetitions of small sequences. Second, a content-based test is used to clean the remaining sequences, but also to complete the repeated sequences to their maximal length. One of our objective is to reduce the number of descriptor needed to achieve this task, given that this computation is the most expensive of the method. As a matter of fact, the method required computing the descriptors of 15.4 % of the images only.

6.4.3. Video summarization with constraint programming

Participants: Mohamed-Haykel Boukadida, Patrick Gros.

Joint work with Sid-Ahmed Berrani, Orange labs.

Up to now, most video summarization methods are based on concepts like saliency and often use a single modality. In order to develop a more general framework, we propose to use a constraint programming approach, where summarizing a video is seen as a constraint resolution problem, which consists in choosing certain excerpts with respect to various criteria. This year we studied several ways to model the problem in order to gain a maximum flexibility in the summary. A first model was based on the selection of shots, the second one on the selection of parts of shots; The third one does not relies on shots and select image sequences directly. The challenge is to express the useful constraints with these models and the limited possibilities of the solver.

6.4.4. Transcript-free spoken content summarization using motif discovery

Participants: Sébastien Campion, Guillaume Gravier.

Joint work with Frédéric Bimbot and Nathan Souviráa-Labastié, Inria/PANAMA, France.

Exploiting previous results on the unsupervised discovery of repeating words in speech signals, we proposed a method dedicated to transcript-free spoken content summarization. Extractive summarization is performed by selecting a small number of segments, typically one or two, which contains most of the repeated fragments [77]. Audio summaries were included in the Texmix demonstration and are currently being evaluated.

6.4.5. TV program structure discovery using grammatical inference

Participants: Guillaume Gravier, Bingqing Qu.

Joint work with Félicien Vallet and Jean Carrive, Institut National de l'Audiovisuel.

Video structuring, in particular applied to TV programs which have strong editing structures, mostly relies on supervised approaches either to retrieve a known structure for which a model has been obtained or to detect key elements from which a known structure is inferred. We investigated an unsupervised approach to recurrent TV program structuring, exploiting the repetitiveness of key structural elements across episodes of the same show. We cast the problem of structure discovery as a grammatical inference problem and show that a suited symbolic representation can be obtained by filtering generic events based on their reoccurring property [92]. The method follows three steps: *i*) generic event detection, *ii*) selection of events relevant to the structure and *iii*) grammatical inference from a symbolic representation. Experimental evaluation is performed on three types of shows, viz., game shows, news and magazines, demonstrating that grammatical inference can be used to discover the structure of recurrent programs with very limited supervision.

6.4.6. Discovering and linking related images in large collections

Participants: Guillaume Gravier, Hervé Jégou, Wanlei Zhao.

We have tackled the problem of image linking. One of the most successful method to link all similar images within a large collection is min-Hash, which is a way to significantly speed-up the comparison of images when the underlying image representation is bag-of-words. However, the quantization step of min-Hash introduces important information loss. In [66], we proposed a generalization of min-Hash, called Sim-min-Hash, to compare sets of real-valued vectors. We demonstrated the effectiveness of our approach when combined with the Hamming embedding similarity. Experiments on large-scale popular benchmarks demonstrated that Sim-min-Hash is more accurate and faster than min-Hash for similar image search. Linking a collection of one million images described by 2 billion local descriptors is done in 7 minutes on a single core machine.

6.5. Natural language processing in multimedia data

6.5.1. Text detection in videos

Participants: Khaoula Elagouni, Pascale Sébillot.

Texts embedded in multimedia documents often provide high level semantic clues that can be used in several applications or services. We thus aim at designing efficient Optical Character Recognition (OCR) systems able to recognize these texts. During the last three years, we have proposed three novel approaches, robust to text variability (different fonts, colors, sizes, etc.) and acquisition conditions (complex background, non-uniform lighting, low resolution, etc.). The first approach relies on a segmentation step and computes nonlinear separations between characters well adapted to the local morphology of images. The two other ones, called segmentation-free approaches, avoid the segmentation step by integrating a multi-scale scanning scheme: The first one relies on a graph model, while the second one uses a particular connectionist recurrent model able to handle spatial constraints between characters. In 2013, a precise evaluation and comparison between these approaches was conducted and published in [16].

6.5.2. Combining lexical cohesion and disruption for topic segmentation

Participants: Guillaume Gravier, Pascale Sébillot, Anca-Roxana Simon.

Topic segmentation classically relies on one of two criteria, either finding areas with coherent vocabulary use or detecting discontinuities. We proposed a segmentation criterion combining both lexical cohesion and disruption, enabling a trade-off between the two [58]. We provide the mathematical formulation of the criterion and an efficient graph based decoding algorithm for topic segmentation. Experimental results on standard textual data sets and on a more challenging corpus of automatically transcribed broadcast news shows demonstrate the benefit of such a combination. Gains were observed in all conditions, with segments of either regular or varying length and abrupt or smooth topic shifts. Long segments benefit more than short segments.However the algorithm has proven robust on automatic transcripts with short segments and limited vocabulary reoccurrences.

6.5.2.1. Information extraction and text mining

Participants: Vincent Claveau, Marie Béatrice Arnulphy.

Following the work initiated in the previous period, we have kept on working on relation extraction. During this year, we have proposed a new prototype that still relies on a supervised machine learning approach but we now rely on the sequence built from the shortest syntactic path between the entities, as it is done in many studies. These paths of lemmas are then used in a kNN whose similarity score is based on language modeling techniques. Based on this new prototype, we have participated to several tracks of the BioNLP challenges concerning the automatic extraction of relations in a specialized corpus. Results obtained with this simple and non-domain specific technique were relatively good, with a second and fourth ranks among the participants for the two tasks concerned [26].

We also pursued previous work on supervised techniques for entity extraction and classification. Instead of working on complex machine learning approaches, we rather use simple methods but the focus is set on clever similarity computing between training examples and candidates for which we make the most of existing information retrieval techniques. Our approach has been evaluated through our participation to BioNLP-ST13 competition, where it has been ranked first [26].

We have also proposed unsupervised techniques for knowledge discovery, more precisely, to bring out coherent groups of entities. Existing techniques are usually based on clustering; the challenge is then to define a notion of similarity between the relevant entities. In this work, we have proposed to divert conditional random fields (CRF) in order to calculate indirectly the similarities among text sequences. Our approach consists in generating artificial labeling problems on the data to be processed to reveal regularities in the labeling of the entities. The good results obtained shows the validity of our approach [27] and opens many research avenues for other knowledge discovery tasks.

6.5.3. Unsupervised approaches to fine-grained morphological analysis

Participants: Vincent Claveau, Ewa Kijak.

Following the work initiated in the previous years, we have kept on studying fine-grained morphological analysis for biomedical information retrieval. In the biomedical field, the key to access information is the use of specialized terms (like *photochemotherapy*). These complex morphological structures may prevent a user querying for *gastrodynia* to retrieve texts containing *stomachalgia*. The original unsupervised technique proposed in 2012 has been further developed and tested. In particular, during this year, we have shown that it largely outperforms state-of-the-art tools (*e.g.*, Morfessor and Derif) for morphological segmentation tasks. It also offers indirect morpho-lexical resources that are more reliable than hand-coded ones used in most state-of-the-art tools [11].

6.5.4. Tree-structured named entities recognition

Participants: Christian Raymond, Davy Weissenbacher.

Many natural language processing tasks needs the production of tree-structured outputs, like syntactic parsing, named entities recognition or language understanding. Currently, only machine learning based systems are robust enough to process the raw and noisy automatic transcribed speech while no machine learning paradigm are able to learn directly the tree structure in a reasonable time. In this work, we studied a solution to tackle the problem of predicting tree structured named entities from speech contents. We investigate a fast and robust decomposition strategy that was implemented and ranked best at the ETAPE NER evaluation campaign with results far better than those of the other participant systems [54].

6.5.5. Fast machine learning algorithm for efficient combination of various features

Participant: Christian Raymond.

Currently, in the field of natural language processing the machine learning algorithm "boosting over decision stumps" is often designed as the best off-the-shell classifier. It's actually widely used for his abilities to work on relatively big dataset, to operate intrinsically feature selection and to produce very good decision rules. We investigated a slight modification of this algorithm where the decision stumps are replaced by bonsai trees. Bonsai trees are small decision trees (with low depth) that can capture some structure in the data that decision stumps can not. This modification allows the boosting algorithm to exhibits better (or in the worst case similar) performances with a lower number of iteration the original algorithm needs. Thus allows in some cases a big improvement in term of performance for a lower cost in term of learning time. An application on image processing (typed/hand classification) exhibited interesting results in [94]

6.6. Competitions and international evaluation benchmarks

6.6.1. FGcomp'2013, in conjunction with Imagenet

Participants: Philippe-Henri Gosselin, Hervé Jégou.

Joint participation with Naila Murray and Florent Perronnin, Xerox Research Center Europe.

We have participated the the FGCOMP'2013 challenge and obtained the best results among all participants, see http://sites.google.com/site/fgcomp2013 Although the proposed system follows most of the standard Fisher classification pipeline, we have evaluated and used several key features and good practices that improve the accuracy when specifically considering fine-grained classification tasks [75]. In particular, we consider the late fusion of two systems both based on Fisher vectors, but that employ drastically different design choices that make them very complementary. Moreover, we show that a simple yet effective filtering strategy significantly boosts the performance for several class domains. The method is described in a technical report.

6.6.2. Hyperlink generation in broadcast videos

Participants: Guillaume Gravier, Pascale Sébillot, Anca-Roxana Simon.

Joint participation with Camille Guinaudeau, Heidelberg Institute of Technology (currently LIMSI-CNRS).

Following up on our 2012 participation, we participated in the Search and hyperlinking task implemented in the framework of the Mediaeval 2013 benchmark initiative. We limited ourselves to hyperlink generation, building on research results in natural language processing, information retrieval and topic segmentation, focusing our contribution on the selection of precise target segments for hyperlinks.

6.6.3. Maurdor campaign

Participant: Christian Raymond.

Joint participation with Yann Ricquebourg, Baptiste Poirriez, Aurélie Lemaitre and Bertrand Coüasnon, IRISA/Intuidoc.

We are participating to the ongoing MAURDOR campaign http://www.maurdor-campaign.org which aims at evaluating systems for automatic processing of written documents. The contribution of TEXMEX comes from the machine learning system based on boosting over bonsai trees we implemented. In the context of this campaign, we investigate the usefulness of this algorithm to combine efficiently features on a relatively big dataset. The very first result shows that this system get state-of-the-art performance while it is much faster than traditional SVM approaches.

6.6.4. Information extraction challenge at BioNLP-ST13

Participant: Vincent Claveau.

BioNLP Shared Task is a community-wide effort to address fine-grained, structural information extraction from biomedical literature. This year, several tasks were proposed and 22 teams participated. TexMex has proposed runs for three main tasks concerning entity extraction and categorization, and relation extraction. The methods proposed by our team are based on machine learning and information retrieval components. Although they do not exploit specialized or domain-specific knowledge, we obtained good results and ranked first, first and third according to the tasks.

TITANE Team

6. New Results

6.1. Analysis

6.1.1. Detecting Parametric Objects in Large Scenes by Monte Carlo Sampling

Participants: Yannick Verdie, Florent Lafarge.

Point processes constitute a natural extension of Markov Random Fields (MRF), designed to handle parametric objects. They have shown efficiency and competitiveness for tackling object extraction problems in vision. Simulating these stochastic models is however a difficult task. The performance of existing samplers are limited in terms of computation time and convergence stability, especially on large scenes. We propose a new sampling procedure based on a Monte Carlo formalism [11]. Our algorithm exploits the Markovian property of point processes to perform the sampling in parallel. This procedure is embedded into a data-driven mechanism so that the points are distributed in the scene as a function of spatial information extracted from the input data. The performance of the sampler is analyzed through a set of experiments on various object detection problems from large scenes, including comparisons to the existing algorithms. The sampler is also evaluated as an optimization algorithm for MRF-based labeling problems (Figure 1).

../../../../projets/titane/IMG/hal-00843022.png

Figure 1. Detection and modeling of trees from large-scale Lidar scans by our parallel Monte Carlo sampler.

6.1.2. Recovering Line-networks in Images by Junction-Point Processes

Participant: Florent Lafarge.

In collaboration with Dengfeng Chai (Zheijiang University, China) and Wolfgang Forstner (University of Bonn, Germany).

We tackle the automatic extraction of line-networks from images. Appearance and shape considerations have been deeply explored in the literature to improve accuracy in presence of occlusions, shadows, and a wide variety of irrelevant objects. However most existing work has ignored the structural aspect of the problem. We present an original method which provides structurally-coherent solutions [13]. Contrary to the pixel-based and object-based methods, our result is a graph in which each node represents either a connection or an ending in the line-network. Based on stochastic geometry, we develop a new family of point processes consisting in sampling junction-points in the input image by using a Monte Carlo mechanism. The quality of a configuration is measured by a probability density which takes into account both image consistency and shape priors. Our experiments on a variety of problems illustrate the potential of our approach in terms of accuracy, flexibility and efficiency (Figure 2).

../../../../projets/titane/IMG/hal-00814262.png

Figure 2. Line-network extraction from images using a junction-point process.

6.2. Approximation

6.2.1. Integer-Grid Maps for Reliable Quad Meshing

Participants: David Bommes, Pierre Alliez.

In collaboration with Leif Kobbelt from RWTH Aachen.

Quadrilateral remeshing approaches based on global parametrization enable many desirable mesh properties. Two of those are (1) high regularity due to explicit control over irregular vertices and (2) smooth distribution of distortion achieved by convex variational formulations. In this work [2] we propose a novel convex Mixed-Integer Quadratic Programming (MIQP) formulation which ensures by construction that the resulting map is within the class of so called Integer-Grid Maps that are guaranteed to imply a quad mesh. In order to overcome the NP-hardness of MIQP we propose two additional optimizations: a complexity reduction algorithm and singularity separating conditions. While the former decouples the dimension of the MIQP search space from the input complexity of the triangle mesh, the latter improves the continuous relaxation, which is crucial for the success of modern MIQP optimizers. Our algorithm also enables the global search for high-quality coarse quad layouts as illustrated in Figure 3 , a difficult task solely tackled by insufficient greedy methodologies before.

../../../../projets/titane/IMG/rmiq.png

Figure 3. Integer-grid maps for reliable quad meshing.

6.2.2. *QEx: Robust Quad Mesh Extraction*

Participant: David Bommes.

In collaboration with Leif Kobbelt from RWTH Aachen.

Among the class of quad remeshing techniques, the ones based on parameterization strive to generate an integer-grid map, i.e., a parametrization of the input surface in 2D such that the canonical grid of integer iso-lines forms a quad mesh when mapped back onto the surface in 3D. An essential, albeit broadly neglected aspect of these methods is the quad extraction step. This step is not a trivial matter: ambiguities induced by numerical inaccuracies and limited solver precision, as well as imperfections in the maps produced by most methods (unless costly countermeasures are taken) pose significant hurdles to the quad extractor. In this work [6] we present a method to sanitize a provided parametrization such that it becomes numerically consistent even with limited precision floating point arithmetic. We also devise a novel strategy to cope with common local fold-overs in the parametrization. We can generate all-quadrilateral meshes where otherwise holes, non-quad polygons or no output at all would have been produced like for the example in Figure 4 .

6.2.3. *Advanced Automatic Hexahedral Mesh Generation from Surface Quad Meshes*

Participant: David Bommes.

In collaboration with Leif Kobbelt (RWTH Aachen).

A purely topological approach for the generation of hexahedral meshes from quadrilateral surface meshes of genus zero has been proposed by M. Müller-Hannemann: in a first stage, the input surface mesh is reduced to a single hexahedron by successively eliminating loops from the dual graph of the quad mesh; in the second stage, the hexahedral mesh is constructed by extruding a layer of hexahedra for each dual loop from the first stage in reverse elimination order. We introduce several techniques to extend the scope of target shapes of the

../../../../projets/titane/IMG/qex.png

Figure 4. QEx: Robust Quad Mesh Extraction.

approach and significantly improve the quality of the generated hexahedral meshes [14]. While the original method can only handle almost-convex objects and requires mesh surgery and remeshing in case of concave geometry, we propose a method to overcome this issue by introducing the notion of concave dual loops in order to handle non-convex objects like the one displayed in Figure 5 . Furthermore, we analyze and improve the heuristic to determine the elimination order for the dual loops such that the inordinate introduction of interior singular edges, i.e., edges of degree other than four in the hexahedral mesh, can be avoided in many cases.

6.2.4. On the Equilibrium of Simplicial Masonry Structures

Participant: Pierre Alliez.

In collaboration with Mathieu Desbrun, Fernando de Goes and Houman Owhadi from Caltech.

We contributed a novel approach for the analysis and design of self-supporting simplicial masonry structures [4]. A finite-dimensional formulation of their compressive stress field is derived, offering a new interpretation of thrust networks through numerical homogenization theory. We further leverage geometric properties of the resulting force diagram to identify a set of reduced coordinates characterizing the equilibrium of simplicial masonry. We finally derive computational form-finding tools that improve over previous work in efficiency, accuracy, and scalability.

6.3. Reconstruction

6.3.1. Noise-Adaptive Shape Reconstruction from Raw Point Sets

Participants: Simon Giraudot, Pierre Alliez.

In collaboration with David Cohen-Steiner (GEOMETRICA project-team)

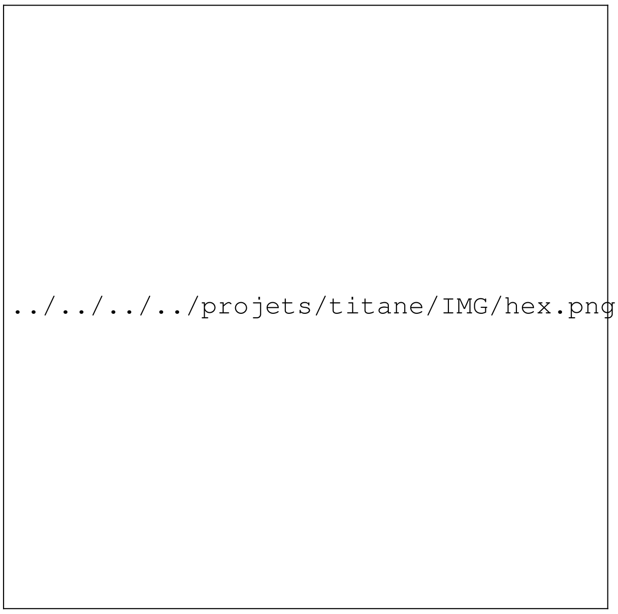


Figure 5. Advanced Automatic Hexahedral Mesh Generation from Surface Quad Meshes.

We devised a noise-adaptive shape reconstruction method specialized to smooth, closed shapes [7]. Our algorithm takes as input a defect-laden point set with variable noise and outliers, and comprises three main steps. First, we compute a novel noise-adaptive distance function to the inferred shape, which relies on the assumption that the inferred shape is a smooth submanifold of known dimension. Second, we estimate the sign and confidence of the function at a set of seed points, through minimizing a quadratic energy expressed on the edges of a uniform random graph. Third, we compute a signed implicit function through a random walker approach with soft constraints chosen as the most confident seed points computed in the previous step.

6.3.2. Surface Reconstruction through Point Set Structuring

Participants: Florent Lafarge, Pierre Alliez.

We present a method for reconstructing surfaces from point sets [8]. The main novelty lies in a structure-preserving approach where the input point set is first consolidated by structuring and resampling the planar components, before reconstructing the surface from both the consolidated components and the unstructured points. Structuring facilitates the surface reconstruction as the point set is substantially reduced and the points are enriched with structural meaning related to adjacency between primitives. Our approach departs from the common dichotomy between smooth/piecewise-smooth and primitive-based representations by gracefully combining canonical parts from detected primitives and free-form parts of the inferred shape (Figure 6).

6.3.3. Hybrid Multi-view Stereo for Modeling Urban Scenes

Participant: Florent Lafarge.

In collaboration with Renaud Keriven (Acute3D), Mathieu Bredif (IGN), and Hiep Vu (Ecole des Ponts ParisTech).

We present an original multi-view stereo reconstruction algorithm which allows the 3D-modeling of urban scenes as a combination of meshes and geometric primitives [9]. The method provides a compact model while preserving details: irregular elements are described by meshes whereas regular structures are described by

Figure 6. Structuring of a church.

canonical geometric primitives. We adopt a two-step strategy consisting first in segmenting the initial mesh-based surface using a multi-label Markov Random Field based model and second, in sampling primitive and mesh components simultaneously on the obtained partition by a Jump-Diffusion process. The quality of a reconstruction is measured by a multi-object energy model which takes into account both photo-consistency and semantic considerations (i.e. geometry and shape layout). The segmentation and sampling steps are embedded into an iterative refinement procedure which provides an increasingly accurate hybrid representation (Figure 7).

6.3.4. Indoor Scene Reconstruction using Primitive-driven Space Partitioning and Graph-cut

Participants: Sven Oesau, Florent Lafarge, Pierre Alliez.

In collaboration with EADS ASTRIUM

We present a method for automatic reconstruction of permanent structures of indoor scenes, such as walls, floors and ceilings, from raw point clouds acquired by laser scanners [15]. Our approach employs graph-cut to solve an inside/outside labeling of a space decomposition. To allow for an accurate reconstruction the space decomposition is aligned with permanent structures. A Hough Transform is applied for extracting the wall directions while allowing a flexible reconstruction of scenes. The graph-cut formulation takes into account data consistency through an inside/outside prediction for the cells of the space decomposition by stochastic ray casting, while favoring low geometric complexity of the model. Our experiments produces watertight reconstructed models of multi-level buildings and complex scenes (Figure 8).

6.3.5. Watertight Scenes from Urban LiDAR and Planar Surfaces

Participant: Thijs Van Lankveld.

In collaboration with Marc Van Kreveld and Remco Veltkamp

../../../../projets/titane/IMG/hal-00759261.png

Figure 7. Reconstruction of a facade from Multi-View Stereo images using a multi-shape hybrid sampler.

../../../../projets/titane/IMG/UDMV13.png

Figure 8. Reconstruction of a multi-floor indoor scene from an input point cloud.

The demand for large geometric models is increasing, especially of urban environments. This has resulted in production of massive point cloud data from images or LiDAR. Visualization and further processing generally require a detailed, yet concise representation of the scene's surfaces. Related work generally either approximates the data with the risk of over-smoothing, or interpolates the data with excessive detail. Many surfaces in urban scenes can be modeled more concisely by planar approximations. We present a method that combines these polygons into a watertight model [10]. The polygon-based shape is closed with free-form meshes based on visibility information. To achieve this, we divide 3-space into inside and outside volumes by combining a constrained Delaunay tetrahedralization with a graph-cut. We compare our method with related work on several large urban LiDAR data sets. We construct similar shapes with a third fewer triangles to model the scenes. Additionally, our results are more visually pleasing and closer to a human modeler's description of urban scenes using simple boxes (Figure 10).

../../../../projets/titane/IMG/lidar.png

Figure 9. Reconstruction of watertight model from LIDAR data sets.

6.3.6. Feature-Preserving Surface Reconstruction and Simplification from Defect-Laden Point Sets

Participant: Pierre Alliez.

In collaboration with David Cohen-Steiner, Julie Digne, Mathieu Desbrun and Fernando de Goes

We introduce a robust and feature-capturing surface reconstruction and simplification method that turns an input point set into a low triangle-count simplicial complex [5]. Our approach starts with a (possibly non-manifold) simplicial complex filtered from a 3D Delaunay triangulation of the input points. This initial approximation is iteratively simplified based on an error metric that measures, through optimal transport, the distance between the input points and the current simplicial complex, both seen as mass distributions. Our approach exhibits both robustness to noise and outliers, as well as preservation of sharp features and boundaries (Figure 10). Our new feature-sensitive metric between point sets and triangle meshes can also be used as a post-processing tool that, from the smooth output of a reconstruction method, recovers sharp features and boundaries present in the initial point set.

../../../../projets/titane/IMG/jmiv.png

Figure 10. Robust reconstruction through optimal transportation.

6.3.7. Splat-based Surface Reconstruction from Defect-Laden Point Sets.

Participant: Pierre Alliez.

In collaboration with Mariette Yvinec (EPI GEOMETRICA), Ricard Campos (University of Girona), Raphael Garcia (University of Girona)

We introduce a method for surface reconstruction from point sets that is able to cope with noise and outliers. First, a splat-based representation is computed from the point set. A robust local 3D RANSAC-based procedure is used to filter the point set for outliers, then a local jet surface – a low-degree surface approximation – is fitted to the inliers. Second, we extract the reconstructed surface in the form of a surface triangle mesh through Delaunay refinement (Figure 11). The Delaunay refinement meshing approach requires computing intersections between line segment queries and the surface to be meshed. In the present case, intersection queries are solved from the set of splats through a 1D RANSAC procedure. [3].

../../../../projets/titane/IMG/splat.png

Figure 11. Splat-based reconstruction.

TOCCATA Team

6. New Results

6.1. Deductive Verification

- F. Bobot, J.-C. Filliâtre, C. Marché, G. Melquiond, and A. Paskevich have presented the proof session mechanism of *Why3* at VSTTE 2013 [23]. It is a technique to maintain a proof session against modification of verification conditions. It was successfully used in developing more than a hundred verified programs and in keeping them up to date along the evolution of *Why3* and its standard library. It also helps out with changes in the environment, *e.g.* prover upgrades.
- M. Clochard, C. Marché, and A. Paskevich developed a general setting for developing programs involving binders, using Why3. This approach was successfully validated on two case studies: a verified implementation of untyped lambda-calculus and a verified tableaux-based theorem prover. This work will be presented at the PLPV conference in January 2014 [29]
- M. Clochard published at the POPL conference a paper presenting a work done during an internship at Rice University (Houston, TX, USA) with S. Chaudhuri and A. Solar-Lezama [28]. It is a new technique for parameter synthesis under boolean and quantitative objectives. The input to the technique is a "sketch" — a program with missing numerical parameters — and a probabilistic assumption about the program's inputs. The goal is to automatically synthesize values for the parameters such that the resulting program satisfies: (1) a boolean specification, which states that the program must meet certain assertions, and (2) a quantitative specification, which assigns a real valued rating to every program and which the synthesizer is expected to optimize.
- J.-C. Filliâtre, L. Gondelman, and A. Paskevich have formalized the notion of ghost code implemented in *Why3*, in a paper *The Spirit of Ghost Code* [49] to be submitted. This is an outcome of L. Gondelman's M2 internship (spring/summer 2013).
- In 2013, two public releases of *Why3* were launched, version 0.81 in March and version 0.82 in December [42]. A first important evolution relies on significant efficiency improvements both in terms of execution speed and of memory usage. The second major evolution is the support for many new provers, including interactive provers PVS 6 (used at NASA) and Isabelle2013-2 (planned to be used in the context of Ada program via Spark), and automated ones: CVC4, Mathematica, Metitarski, Metis, Beagle, Princess, and Yices2. The design of the programming language of *Why3*(WhyML) was presented during a tool demonstration at the ESOP conference [33].

6.2. Floating-Point and Numerical Programs

- S. Boldo, F. Clément, J.-C. Filliâtre, M. Mayero, G. Melquiond, and P. Weis, finished the formal proof of a numerical analysis program: the second order centered finite difference scheme for the one-dimensional acoustic wave [15].
- S. Boldo developed a formal proof of an algorithm for computing the area of a triangle, an improvement of its error bound and new investigations in case of underflow [25].
- S. Boldo, J.-H. Jourdan, X. Leroy, and G. Melquiond, extended CompCert to get the first formally verified compiler that provably preserves the semantics of floating-point programs [26].
- S. Boldo and G. Melquiond wrote a chapter of the book [38] that describes the current state of the Mathematics/Computer science research in France.
- C. Lelay worked on formalizing power series for the Coq proof assistant [35].

- Most 18-year old French students pass an exam called Baccalaureate which ends the high school and is required for attending the university. The idea was to try our Coq library Coquelicot on the 2013 mathematics test of the scientific Baccalaureate. C. Lelay went to the “Parc de Vilgénis” high school in Massy, France and took the 2013 test at the same time as the students, but had to formally prove the answers. There was therefore no possible cheating: the Coq library was already developed and it was tested as is during the four hours of the test. This experiment shows that Coquelicot is able to cope with basic real analysis: it has the necessary definitions and lemmas, and its usability and efficiency have been demonstrated in a test with a limited time [45] (see also https://www.lri.fr/~lelay/).
- D. Ishii and G. Melquiond applied methods of deductive program verification to ensure the safety of hybrid automata [34].
- É. Martin-Dorel, G. Hanrot, M. Mayero, L. Théry, showed how to generate and formally check certificates in the Coq proof assistant to solve myriads of instances of the Integer Small Value Problem (ISValP). This problem is directly related to solving the Table Maker’s Dilemma with hardest-to-round computations [50]. A new version of the formalized library has been released (http://tamadi.gforge.inria.fr/CoqHensel/).
- É. Martin-Dorel, G. Melquiond, and J.-M. Muller, studied issues related to double rounding in the implementation of error-free transformations [16].

6.3. Automated Reasoning

- C. Dross, S. Conchon, J. Kanig, and A. Paskevich have proposed a new approach for handling quantified formulas in SMT solvers. Their framework is based on the notion of instantiation patterns, also known as triggers, that suggest instances which are more likely to be useful in proof search. This framework has been implemented in the Alt-Ergo SMT solver [48].
- S. Conchon, A. Goel, S. Krstic, A. Mebsout, and F. Zaïdi have designed a new model checking algorithm that is able to infer invariants strong enough to prove complex parameterized cache-coherence protocols [30].
- S. Conchon, A. Mebsout, and F. Zaïdi have presented a new SMT library called Alt-Ergo-Zero. This library is tightly integrated to the backward reachability algorithm of the Cubicle model checker [31].
- S. Conchon, M. Iguernelala, and A. Mebsout have designed a collaborative framework for reasoning modulo simple properties of non-linear arithmetic. This framework has been implemented in the Alt-Ergo SMT solver [47].
- J. C. Blanchette and A. Paskevich designed an extension to the TPTP TFF (Typed First-order Form) format of theorem proving problems to support rank-1 polymorphic types (also known as ML-style parametric polymorphism). This extension, named TFF1, was incorporated in the TPTP standard and was presented at the CADE-24 conference [22].

6.4. Certification of Languages, Tools and Systems

- A. Tafat and C. Marché developed a certified VC generator using Why3. The challenge was to formalize the operational semantics of an imperative language, and a corresponding weakest precondition calculus, without the possibility to use *Coq* advanced features such as dependent types nor higher-order functions. The classical issues with local bindings, names and substitutions were solved by identifying appropriate lemmas. It was shown that Why3 can offer a very significantly higher amount of proof automation compared to *Coq* [36]
- A. Charguéraud, together with the other members of the *JsCert* team have developed this year the first complete formalization of the semantics of the JavaScript programming language. This project is joint work with Philippa Gardner, Sergio Maffeis, Gareth Smith, Daniele Filaretti and Daiva Naudziuniene from Imperial College, and Alan Schmitt and Martin Bodin from Inria Rennes – Bretagne Atlantique (see http://jscert.org).

The formalization consists of a set of inductive rules translating the prose from the *ECMAScript Language Specification, version 5*. These rules can be used to formally reason about program behaviors or to establish the correctness of program transformations. In addition to the inductive rules, a reference interpreter has been proved correct. This interpreter may be used to run actual JavaScript program following the rules of the formal semantics. It has been used in particular to validate the formal semantics against official JavaScript test suites.

The formalization of JavaScript has been published at POPL 2014 [24]. A key ingredient in this formalization is the use of the *pretty-big-step semantics*. This technique allows for representing evaluation rules in big-step style without suffering from a duplication of several premises across different rules. The pretty-big-step technique is described in a paper published by A. Charguéraud at ESOP 2013 [27].

- É. Contejean, together with V. Benzaken and their PhD student S. Dumbrava, have proposed a *Coq* formalization of the relational data model which underlies relational database systems [21]. Proposing such a formalization is the first, *essential* step, that will allow to *prove* that existing systems conform to their specifications and to *verify* both production implementations of database systems and database-backed applications. More precisely, they present and formalize the data definition part of the model including integrity constraints, attributes, tuples, relations, schemas and integrity constraints (including the so-called Armstrong's system and the chase). They model two different query language formalisms: relational algebra and conjunctive queries. The former is the basis of the SQL commercial query language and the latter underlies graphical languages, such as Microsoft Access or Query By Example (QBE). They also present logical query optimization and prove the main "database theorems": algebraic equivalences, the homomorphism theorem and conjunctive query minimization.

6.5. Miscellaneous

- R. El Sibaie and J.-C. Filliâtre have developed *Combine*, an OCaml library for combinatorics. It provides two different solutions to the exact matrix cover problem: Knuth's dancing links and ZDDs, a variant of binary decision diagrams [32].

TOSCA Project-Team

6. New Results

6.1. Probabilistic numerical methods, stochastic modelling and applications

Participants: Mireille Bossy, Nicolas Champagnat, Julien Claisse, Madalina Deaconu, Samuel Herrmann, James Inglis, Antoine Lejay, Sylvain Maire, Sebastian Niklitschek Soto, Denis Talay, Etienne Tanré, Denis Villemonais, Laurent Violeau.

6.1.1. Published works and preprints

- M. Bossy and J-F. Jabir (University of Valparaíso) [29], have proved the well-posedness of a conditional McKean Lagrangian stochastic model, endowed with the specular boundary condition, and further the mean no-permeability condition, in a smooth bounded confinement domain $\mathcal{D}$. This result extends their previous work [48], where the confinement domain was the upper-half plane. The extension of the construction to more general confinement domain exhibits difficulties that we handle by combining stochastic calculus and the analysis of kinetic equations. As a prerequisite for the study of the nonlinear case, we construct a Langevin process confined in $\overline{\mathcal{D}}$ and satisfying the specular boundary condition. We then use PDE techniques to construct the time-marginal densities of the nonlinear process from which we are able to exhibit the conditional McKean Lagrangian stochastic model.
- N. Champagnat studied in collaboration with S. Méléard (Ecole Polytechnique, Palaiseau) and P.-E. Jabin (Univ. of Maryland) adaptive dynamics and evolutionary branching in individual-based models of populations competing for resources, where resources consumption is modelled similarly as for chemostat systems of ODEs [13].
- M. Deaconu and S. Herrmann constructed a new procedure for the simulation of the hitting times of nonlinear boundaries for Bessel processes. This method, called the random walk on moving spheres algorithm, is based on two key properties: first, the explicit distribution of the first hitting time of a particular boundary for the Bessel process; second, the connexion between the Bessel process and the Euclidean norm of a Brownian motion having the same dimension. This result can be applied for the hitting time of a given level for the Cox-Ingersoll-Ross process and thus be used in models arising from finance and neurosciences [15].
- J. Inglis and E. Tanré completed their study with F. Delarue and S. Rubenthaler (Univ. Nice – Sophia Antipolis) on the global solvability of a networked system of integrate-and-fire neurons proposed in the neuroscience literature. To do this it was necessary to obtain some general estimates of the first hitting times of barriers by non-homogeneous processes, which have been collected together separately in [40], http://hal.inria.fr/hal-00870991.
- J. Inglis, in collaboration with O. Faugeras (EPI NeuroMathComp), studied the well-posedness of stochastic neural field equations within a rigorous framework. The deterministic versions of these equations have been used to great success for the macroscopic modeling of brain activity. Their stochastic counterparts are non-trivial SPDEs, due to the presence of a nonlocal operator [26], http://hal.inria.fr/hal-00907555.
- A. Lejay and L. Coutin (Université de Toulouse) have continued their work on the sensitivity of the Itô's map in the context of rough paths [37].
- With L. Coutin (Université de Toulouse), A. Lejay has provided a framework for considering linear rough differential equations [49].
- With A. Kohatsu-Higa (Ritsumeikan University) and K. Yasuda (Hosei University), A. Lejay provided bounds on the weak rate of convergence of the Euler scheme when the drift term is discontinuous [41].

- S. Maire and G. Nguyen have developed a Monte Carlo method to deal with Robin and transmission conditions for elliptic diffusion equations in stratified media. It combines walk on spheres techniques and finite differences [44].
- D. Villemonais worked on the empirical distribution of Fleming-Viot type particle systems. Using couplings with reflected diffusion processes, he proved the uniform tightness of such empirical distributions and deduced the non-degeneracy of the law of diffusion processes conditioned not to hit a boundary [19].
- D. Villemonais proved in [18] a general approximation method for Markov processes conditioned not be killed. The method is based on a mean field interacting particles system which is easy to simulate. The study also details the particular case of time/environment dependent diffusion processes.

6.1.2. Other works in progress

- N. Champagnat and B. Henry work on the long-time behaviour of the frequency spectrum for the Splitting Tree models under the infinitly-many alleles model. Specificaly, they want to study the asymptotic behavior of the largests families in the “supercritical clonal” case. Such results could be applied to design statistical methods to detect positive selection of a gene in a growing population.
- N. Champagnat, D. Ritchie (ORPAILLEUR team, Inria Nancy) and B. Henry work on the design of a stochastic model for the evolution of 3D structures of proteins. Using Kpax algorithm [52], which allow to quantify the evolutionary distance between proteins, their goal is to design a statistical method to infer phylogenetic trees with particle systems methods.
- N. Champagnat and D. Villemonais obtained criterions for existence and uniqueness of quasi-stationary distributions and Q-processes for general absorbed Markov processes. A quasi-stationary distribution is a stationary distribution conditionnally on non-absorbtion, and the Q-process is defined as the original Markov process conditionned to never be absorbed. The criterion that they obtain ensures exponential convergence of the conditionned t-marginal of the process conditionned not to be absorbed at time t, to the quasi-stationary distribution and also the exponential ergodicity of the Q-process. This work is currently being written.
- J. Claisse continued his PhD. under the supervision of N. Champagnat and D. Talay on stochastic control of population dynamics. He completed a finite-horizon optimal control problem on branching–diffusion processes. He also created and studied a hybrid model of tumor growth emphasizing the role of acidity. Key therapeutic targets appear in the model to allow investigation of optimal treatment problems.
- M. Deaconu and S. Herrmann are developing a new algorithm for the simulation of Bessel processes hitting times for non-integer dimensions. The idea is to decompose the dimension into its integer part and its fractional part and use the additivity property for squared Bessel processes. Each simulation step is splitted in two parts: one uses the integer dimension case and the other one considers hitting times of a Bessel process starting from zero.
- M. Deaconu in collaboration with L. Beznea (IMAR Bucarest) and O. Lupaşcu (Université Paris 13 and IMAR Bucarest) studies the connexion between the coagulation/fragmentation phenomena and branching processes.
- J. Inglis and D. Talay are developing a mean-field model of a network of neurons, that contains both a spatial element describing the transmission of a signal along dendrites, as well as non-homogenous weights that represent the strength of the synaptic connections. More generally, this leads to the study of the limiting behavior of non-exchangeable mean-field particle systems.
- J. Inglis and E. Tanré are continuing their collaboration with F. Delarue (Univ. Nice – Sophia Antipolis) by developing approximations to a limiting equation describing the behavior of a large network of neurons all behaving according to the integrate-and-fire model. Both a particle system approximation and an approximation involving delays are considered.
- S. Larnier and A. Lejay have worked on nearshore wave analysis and bathymetry identification through the use of a video installed on the shore [42], [43].

- A. Lejay has continued his work with R. Rebolledo (Pontificia Universidad Católica), S. Torres (Universidad de Valparaíso) and E. Mordecki (Universidad de la República) on the parametric estimation of coefficients of diffusion with discontinuous coefficients.
- S. Maire and I. Dimov (Bulgarian academy of sciences) have introduced a new Monte Carlo method to solve real or complex linear systems of equations. Coupled with sequential Monte Carlo this walk on equations method shows a very fast convergence. A similar method is in progress to solve linear integral equations.
- S. Niklitschek Soto and D. Talay have set up and solved a new martingale problem which has allowed them to get a new stochastic representation for solutions of multi-dimensional diffraction parabolic PDEs with general discontinuous coefficients. One of the main difficulties to overcome has been to identify the proper weighted local time process involved in the stochastic dynamics. This work opens the way to innovating Monte Carlo methods for this class of PDEs.
- P. Guiraud (University of Valparaíso) and E. Tanré study the effect of noise in the phenomenon of spontaneous synchronisation in a network of full connected integrate- and-fire neurons. They detail cases in which the phenomenon of synchronization persists in a noisy environment, cases in which noise permits to accelerate synchronization, and cases in which noise permits to observe synchronization while the noiseless model does not show synchronization. (Math Amsud program SIN)
- L. Capietto worked during his internship under the supervision of O. Faugeras (EPI NEUROMATHCOMP) and E. Tanré on extension of [51], in a context with several populations of homogeneous neurons. They study the limit mean field equation of the membrane potential as the number of neurons increase in a network with correlated synaptic weights.
- E. Tanré, in collaboration with O. Faugeras (EPI NEUROMATHCOMP) and the team Inference and Visual Behavior (IViBe) of Institut de Neurosciences de la Timone (INT), studied the motion of eyes, the phenomena of sacades and micro-saccades when monkeys or humans have to fix the center of a picture during a few minutes. They introduce a stochastic model to describe the typical path of the eyes on the picture and evaluate the link between the characteristics of the artificial pictures and the coefficients of the stochastic model.
- L. Violeau continued his PhD. on *Stochastic Lagrangian Models and Applications to Downscaling in Fluid Dynamics* under the supervision of M. Bossy and A. Rousseau (LEMON team, Inria Sophia Antipolis - Méditerranée). Laurent Violeau has obtained a theoretical rate of convergence of the particle approximation of kinetic conditional McKean-Vlasov stochastic models. This result is the first that explicits the complex relationship between the two sources of spacial errors in such kind of algorithm: the smoothing parameter for the conditional expectation estimator and the number of interacting particles. This theoretical convergence rate was confronted with numerical tests in the case of simplified Lagrangian models that confirm the pertinence of the theoretical bound for the error.
- C. Graham and D. Talay are writing the second volume of their series published by Springer on the Mathematical Foundations of Stochastic Simulations.
- In collaboration with N. Touzi (Ecole Polytechnique), D. Talay is studying stochastic differential equations involving local times with stochastic weights, and extensions of classical notions of viscosity solutions to PDEs whose differential operator has discontinuous coefficients and transmission boundary conditions.

6.2. Financial Mathematics

Participants: Mireille Bossy, Nicolas Champagnat, Paul Charton, Madalina Deaconu, Dalia Ibrahim, Antoine Lejay, Khaled Salhi, Denis Talay, Etienne Tanré.

6.2.1. Published works and preprints

- In collaboration with N. Maïzi (CMA - Mines Paristech) and O. Pourtallier (COPRIN team, Inria Sophia Antipolis - Méditerranée), M. Bossy studied the existence result of a Nash equilibrium between electricity producers selling their production on an electricity market and buying CO2 emission allowances on an auction carbon market. The producers' strategies integrate the coupling of the two markets via the cost functions of the electricity production. The authors set out a clear Nash equilibrium that can be used to compute equilibrium prices on both markets as well as the related electricity produced and CO2 emissions covered [30]
- In addition to the internship of K. Salhi, N. Champagnat, M. Deaconu, and A. Lejay have worked on the use of power law to predict risk in financial markets using data from Euronext NSYE stocks exchanges [33].
- P. Charton submitted an article [35] on the optimal operation of a windfarm equipped with a storage unit.

6.2.2. Other works in progress

- D. Ibrahim, D. Talay and E. Tanré worked on a model coming from technical analysis in finance. They study the Bollinger Bands indicator to detect jumps in the volatility in an extension of classical Black and Scholes models. They evaluate the efficiency of such indicators to detect the random time at which the volatility jump from a *small* value to a *large* one. A paper is being written.
- In collaboration with Victor Reutenauer and Christophe Michel (CA-CIB), D. Talay and E. Tanré worked on a model in financial mathematics including bid-ask spread cost. They study the optimal strategy to hedge an interest rate swap that pays a fixed rate against a floating rate. They present a methodology using a stochastic gradient algorithm to optimize strategies. A paper is being submitted.
- In collaboration with J. Bion-Nadal (Ecole Polytechnique and CNRS), D. Talay introduced a new calibration method based on dynamical risk measures and stochastic control PDEs. A paper is being written.

6.3. Stochastic Analysis

Participants: Nicolas Champagnat, Julien Claisse, Denis Talay.

- N. Champagnat studied in collaboration with P.-E. Jabin (Univ. of Maryland) strong existence and pathwise uniqueness for stochastic differential equations driven by a Brownian motion and with rough coefficients [34]. The method is an extension of the one of [50], which studies well-posedness for deterministic dynamical system. Strong existence and pathwise uniqueness can be proved for example if the drift vector is $L^1(W^{1,1})$ and the diffusion matrix is uniformly elliptic and $L^q(W^{1,p})$ with $2/q + d/p = 1$. This improves the previous conditions of [53].
- J. Claisse and D. Talay studied in collaboration with X. Tan (Univ. of Paris Dauphine) a conditioning argument which is often used to prove the dynamic programming principle [36]. Their study of the literature revealed that previous proofs of this argument are incorrect or incomplete. They provided a rigorous and detailed proof by setting up martingale controlled problems in a original way.

TRIO Team

6. New Results

6.1. Probabilistic real-time systems

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

The arrival of complex hardware responding to the increasing demand for computing power in next generation systems exacerbates some of the limitations of static timing analysis for the estimation of the worst-case execution time (WCET) estimation. In particular, the effort of acquiring (1) detail information on the hardware to develop an accurate model of its execution latency as well as (2) knowledge of the timing behaviour of the program in the presence of varying hardware conditions, such as those dependent on the history of previously executed instructions. These problems are also known as the timing analysis walls. The probabilistic timing analysis, a novel approach to the analysis of the timing behaviour of next-generation real-time embedded systems, provides answers to timing analysis walls. In [7], [13], [11] timing analysis attacks the timing analysis walls. We have also presented experimental evidence that shows how probabilistic timing analysis reduces the extent of knowledge about the execution platform required to produce probabilistically-safe and tight WCET estimations.

Based on existing estimations of WCET or minimal inter-arrival time [16], we may propose different probabilistic schedulability analyses [6], [12].

2013 was also the year when through several invited talks [8], [10], [9], we had the opportunity to underline historical misunderstandings on probabilistic real-time systems. The most common is related to the notion of independence that is used with a wrong meaning by different papers.

TRISKELL Project-Team

6. New Results

6.1. Support for Reverse Engineering and Maintaining Feature Models

Feature Models (FMs) are a popular formalism for modelling and reasoning about commonality and variability of a system. In essence, FMs aim to define a set of valid combinations of features, also called configurations. In [35], we tackle the problem of synthesising an FM from a set of configurations. The main challenge is that numerous candidate FMs can be extracted from the same input configurations, yet only a few of them are meaningful and maintainable. We first characterise the different meanings of FMs and identify the key properties allowing to discriminate between them. We then develop a generic synthesis procedure capable of restituting the intended meanings of FMs based on inferred [72] or user-specified knowledge. Using tool support, we show how the integration of knowledge into FM synthesis can be realized in different practical application scenarios that involve reverse engineering and maintaining FMs.

6.2. Feature Model Extraction from Large Collections of Informal Product Descriptions

Feature Models (FMs) are used extensively in software product line engineering to help generate and validate individual product configurations and to provide support for domain analysis. As FM construction can be tedious and time-consuming, researchers have previously developed techniques for extracting FMs from sets of formally specified individual configurations, or from software requirements specifications for families of existing products. However, such artifacts are often not available. In [44] we present a novel, automated approach for constructing FMs from publicly available product descriptions found in online product repositories and marketing websites such as SoftPedia and CNET. While each individual product description provides only a partial view of features in the domain, a large set of descriptions can provide fairly comprehensive coverage. Our approach utilizes hundreds of partial product descriptions to construct an FM and is described and evaluated against antivirus product descriptions mined from SoftPedia.

6.3. On Product Comparison Matrices and Variability Models

Product comparison matrices (PCMs) provide a convenient way to document the discriminant features of a family of related products and now abound on the internet. Despite their apparent simplicity, the information present in existing PCMs can be very heterogeneous, partial, ambiguous, hard to exploit by users who desire to choose an appropriate product. Variability Models (VMs) can be employed to formulate in a more precise way the semantics of PCMs and enable automated reasoning such as assisted configuration. Yet, the gap between PCMs and VMs should be precisely understood and automated techniques should support the transition between the two. We propose variability patterns that describe PCMs content and conduct an empirical analysis of 300+ PCMs mined from Wikipedia [62], we also identify the limits of existing comparators, configurators and PCMs [67], [62]. Our findings are a first step toward better engineering techniques for maintaining and configuring PCMs.

6.4. Generating Counterexamples of Model-based Software Product Lines: An Exploratory Study

Model-based Software Product Line (MSPL) engineering aims at deriving customized models corresponding to individual products of a family. The design space of an MSPL is extremely complex to manage for the engineer, since the number of variants may be exponential and the derived product models have to conform to numerous well-formedness and business rules. We provide a way to generate MSPLs, called counterexamples, that can produce invalid product models despite a valid configuration in the variability model [49]. We provide a systematic and automated process, based on the Common Variability Language (CVL), to randomly search the space of MSPLs for a specific formalism. We validate the effectiveness of this process for three formalisms at different scales (up to 247 metaclasses and 684 rules).

6.5. Composing your Compositions of Variability Models

Modeling and managing variability is a key activity in a growing number of software engineering contexts. Support for composing variability models is arising in many engineering scenarios, for instance, when several subsystems or modeling artifacts, each coming with their own variability and possibly developed by different stakeholders, should be combined together. We consider in [34] the problem of composing feature models (FMs), a widely used formalism for representing and reasoning about a set of variability choices. We show that several composition operators can actually be defined, depending on both matching/merging strategies and semantic properties expected in the composed FM. We present four alternative forms and their implementations. We discuss their relative trade-offs w.r.t. reasoning, customizability, traceability, composability and quality of the resulting feature diagram. We summarize these findings in a reading grid which is validated by revisiting some relevant existing works. Our contribution should assist developers in choosing and implementing the right composition operators.

6.6. Extraction and Evolution of Architectural Variability Models in Plugin-based Systems

Variability management is a key issue when building and evolving software-intensive systems, making it possible to extend, configure, customize and adapt such systems to customers' needs and specific deployment contexts. A wide form of variability can be found in extensible software systems, typically built on top of plugin-based architectures that offer a (large) number of configuration options through plugins. In an ideal world, a software architect should be able to generate a system variant on-demand, corresponding to a particular assembly of plugins. To this end, the variation points and constraints between architectural elements should be properly modeled and maintained over time (i.e., for each version of an architecture). A crucial, yet error-prone and time-consuming, task for a software architect is to build an accurate representation of the variability of an architecture, in order to prevent unsafe architectural variants and reach the highest possible level of flexibility. In [23], we propose a reverse engineering process for producing a variability model (i.e., a feature model) of a plugin-based architecture. We develop automated techniques to extract and combine different variability descriptions, including a hierarchical software architecture model, a plugin dependency model and the software architect knowledge. By computing and reasoning about differences between versions of architectural feature models, software architect can control both the variability extraction and evolution processes. The proposed approach has been applied to a representative, large-scale plugin-based system (FraSCAti), considering different versions of its architecture. We report on our experience in this context.

6.7. FAMILIAR: A Domain-Specific Language for Large Scale Management of Feature Models

The feature model formalism has become the de facto standard for managing variability in software product lines (SPLs). In practice, developing an SPL can involve modeling a large number of features representing di erent viewpoints, sub-systems or concerns of the software system. This activity is generally tedious and error-prone. In [24], we present FAMILIAR a Domain-Specific Language (DSL) that is dedicated to the large scale management of feature models and that complements existing tool support. The language provides a powerful support for separating concerns in feature modeling, through the provision of composition and decomposition operators, reasoning facilities and scripting capabilities with modularization mechanisms. We illustrate how an SPL consisting of medical imaging services can be practically managed using reusable FAMILIAR scripts that implement reasoning mechanisms. We also report on various usages and applications of FAMILIAR and its operators, to demonstrate their applicability to di erent domains and use for di erent purposes.

6.8. Web Configurators

Nowadays, mass customization has been embraced by a large portion of the industry. As a result, the web abounds with sales configurators that help customers tailor all kinds of goods and services to their specific needs. In many cases, configurators have become the single entry point for placing customer orders. As such, they are strategic components of companies' information systems and must meet stringent reliability, usability and evolvability requirements. However, the state of the art lacks guidelines and tools for efficiently engineering web sales configurators. To tackle this problem, empirical data on current practice is required. The paper [51] reports on a systematic study of 111 web sales configurators along three essential dimensions: rendering of configuration options, constraint handling, and configuration process support. Based on this, we highlight good and bad practices in engineering web sales configurator. The reported quantitative and qualitative results open avenues for the elaboration of methodologies to (re-)engineer web sales configurators. In [48] we focus on how to associate product configurations to visual representations in a Web configurator. We present a formal statement of the problem and a model-driven perspective.

6.9. Separating Concerns in Feature Models

Feature models (FMs) are a popular formalism to describe the commonality and variability of a set of assets in a software product line (SPL). SPLs usually involve large and complex FMs that describe thousands of features whose legal combinations are governed by many and often complex rules. The size and complexity of these models is partly explained by the large number of concerns considered by SPL practitioners when managing and configuring FMs. In the chapter [68], we first survey concerns and their separation in FMs, highlighting the need for more modular and scalable techniques. We then revisit the concept of view as a simplified representation of an FM. We finally describe a set of techniques to specify, visualize and verify the coverage of a set of views. These techniques are implemented in complementary tools providing practical support for feature-based configuration and large scale management of FMs.

6.10. Bridging the Chasm between Executable Metamodeling and Models of Computation

The complete and executable definition of a Domain Specific Language (DSL) relies on the specification of two essential facets: a model of the domain-specific concepts with actions and their semantics; a scheduling model that orchestrates the actions of a domain-specific model. Metamodels can capture the former facet, while Models of Computation (MoCs) capture the latter facet. Unfortunately, theories and tools for metamodeling and MoCs have evolved independently, creating a cultural and technical chasm between the two communities. We introduce a new framework to bridge a metamodel and a MoC in a modular fashion [43]. This bridge allows (i) the complete and executable definition of a DSL, (ii) the reuse of MoCs for different domain-specific metamodels, and (iii) the use of different MoCs for a given metamodel, to cope with variation points of a DSL.

6.11. Reifying Concurrency for Executable Metamodeling

Current metamodeling techniques can be used to specify the syntax and semantics of domain specific modeling languages (DSMLs). However, there is currently very little support for explicitly specifying concurrency semantics using metamodels. We reify concurrency as a metamodeling facility, leveraging formalization work from the concurrency theory and models of computation (MoC) community [42]. The essential contribution of this paper is a proposed language workbench for binding domain-specific concepts and models of computation through an explicit event structure at the metamodel level. We illustrate these novel metamodeling facilities for designing two variants of a concurrent and timed final state machine, and provide other experiments to validate the scope of our approach.

6.12. Using Model Types to Support Contract-Aware Model Substitutability

Model typing brings the benefit associated with well-defined type systems to model-driven development (MDD) through the assignment of specific types to models. In particular, model type systems enable reuse of model manipulation operations (e.g., model transformations), where manipulations defined for models of a supertype can be used to manipulate models of subtypes. Existing model typing approaches are limited to structural typing defined in terms of object-oriented metamodels (e.g., MOF) in which the only structural (well-formedness) constraints are those that can be expressed directly in metamodeling notations (e.g., multiplicity and element containment constraints). We propose an extension to model typing that takes into consideration structural invariants, other than those that can be expressed directly in metamodeling notation, and specifications of behaviors associated with model types [64]. The approach supports contract-aware substitutability, where contracts are defined in terms of invariants and pre-/postconditions expressed using OCL. Support for behavioral typing paves the way for behavioral substitutability. We also describe a technique to rigorously reason about model type substitutability as supported by contracts and apply the technique in use cases from the optimizing compiler community.

6.13. Variability Support in Domain-Specific Language Development

Domain Specific Languages (DSLs) are widely adopted to capitalize on business domain experiences. Consequently, DSL development is becoming a recurring activity. Unfortunately, even though it has its benefits, language development is a complex and time-consuming task. Languages are commonly realized from scratch, even when they share some concepts and even though they could share bits of tool support. This cost can be reduced by employing modern modular programming techniques that foster code reuse. However, selecting and composing these modules is often only within the reach of a skilled DSL developer. We propose to combine modular language development and variability management, with the objective of capitalizing on existing assets [63]. This approach explicitly models the dependencies between language components, thereby allowing a domain expert to configure a desired DSL, and automatically derive its implementation. The approach is tool supported, using Neverlang to implement language components, and the Common Variability Language (CVL) for managing the variability and automating the configuration. We illustrate our approach with the help of different case studies, including the implementation of a family of DSLs to describe variants of state machines.

6.14. Automatically Searching for Metamodel Well-Formedness Rules in Examples and Counter-Examples

Current metamodeling formalisms support the definition of a metamodel with two views: classes and relations, that form the core of the metamodel, and well-formedness rules, that constraints the set of valid models. While a safe application of automatic operations on models requires a precise definition of the domain using the two views, most metamodels currently present in repositories have only the first one part. We propose in [47] to start from valid and invalid model examples in order to automatically retrieve well-formedness rules in OCL using Genetic Programming. The approach is evaluated on metamodels for state machines and features diagrams. The experiments aim at demonstrating the feasibility of the approach and at illustrating some important design decisions that must be considered when using this technique.

6.15. Building Modular and Efficient DSLs: Mashup of Meta-Languages and its Implementation in the Kermeta Language Workbench

With the growing use of domain-specific languages (DSL) in industry, DSL design and implementation goes far beyond an activity for a few experts only and becomes a challenging task for thousands of software engineers. DSL implementation indeed requires engineers to care for various concerns, from abstract syntax, static semantics, behavioral semantics, to extra-functional issues such as run-time performance. We propose an approach that uses one meta-language per language implementation concern [27] in the new version (v2) of the Kermeta workbench. We show that the usage and combination of those meta-languages is simple and intuitive enough to deserve the term "mashup". We evaluate the approach by completely implementing the non trivial fUML modeling language, a semantically sound and executable subset of the Unified Modeling Language (UML) ; Kompren, a DSL for designing and implementing model slicers ; and KCVL, the Commun Variability Language dedicated to variability management in software design models [65].

6.16. On the Globalization of Modeling Languages

In the software and systems modeling community, research on domain-specific modeling languages (DSMLs) is focused on providing technologies for developing languages and tools that allow domain experts to develop system solutions efficiently in a particular domain. Unfortunately, the current lack of support for explicitly relating concepts expressed in different DSMLs makes it very difficult for software and system engineers to reason about information spread across models describing different system aspects. Supporting coordinated use of DSMLs leads to what we call the globalization of modeling languages. We present a research initiative that broadens the current DSML research focus beyond the development of independent DSMLs to one that provides support for globalized DSMLs, that is, DSMLs that facilitate coordination of work across different domains of expertise [31]. We explore this new grand challenge in recent workshops, *e.g.,* GlobalDSL'13 at ECSA, ECMFA and ECOOP 2013 [69], and GEMOC'13 at MODELS 2013 [70].

6.17. Automating the Maintenance of Non-functional System Properties using Demonstration-based Model Transformation

Given a base model with functional components, maintaining the non-functional properties that crosscut the base model has become an essential modeling task when using DSMLs. We present a demonstration-based approach to automate the maintenance of non-functional properties in DSMLs [29]. Instead of writing model transformation rules explicitly, users demonstrate how to apply the non-functional properties by directly editing the concrete model instances and simulating a single case of the maintenance process. An inference engine generates generic model transformation patterns, which can be refined by users and then reused to automate the same evolution and maintenance task in other models. Our demonstration-based approach has been applied to several scenarios, such as auto-scaling and model layout.

6.18. Improving Reusability and Automation in Software Process Lines

Software processes orchestrate manual or automatic tasks to create new software products that meet the requirements of specific projects. While most of the tasks are about inventiveness, modern developments also require recurrent, boring and time-consuming tasks (e.g., the IDE configuration, or the continuous integration setup). Such tasks struggle to be automated due to their various execution contexts according to the requirements of specific projects. We propose a methodology that benefits from an explicit modeling of a family of processes to identify the possible reuse of automated tasks in software processes [60]. Then, we propose a tool-supported approach that integrates both reuse and automation [61]. It consists of reusing processes from an SPL according to projects' requirements. The processes are bound to components that automate their execution. When the variability of a process to execute is not fully resolved, our approach consists of resolving this variability during the execution of this process. We illustrate our approach on industrial projects in a software company, as well as on a family of processes for designing and implementing modeling languages. Our approach promoted the identification of possible automated tasks for configuring IDEs and continuous integration, their reuse in various projects of the company, and the automation of their execution, while enabling to resolve process variability during the execution.

6.19. Towards Trust-Aware and Self-Adaptive Systems

The dynamic conditions under which Future Internet (FI) applications must execute call for self-adaptive software to cope with unforeseeable changes in the application environment. Software engineering currently provides frameworks to develop reasoning engines that support the runtime adaptation of distributed, heterogeneous applications. However, these frameworks have very limited support to address security concerns of these application, hindering their usage for FI scenarios. We address this challenge by enhancing self-adaptive systems with the concepts of trust and reputation [58]. Trust improves decision-making processes under risk and uncertainty, in turn improving security of self-adaptive FI applications.

6.20. SOA Antipatterns: an Approach for their Specification and Detection

The changes resulting from the evolution of Service Based Systems (SBSs) may degrade their design and quality of service (QoS) and may often cause the appearance of common poor solutions in their architecture, called antipatterns. We introduce a novel and innovative approach supported by a framework for specifying and detecting antipatterns in SBSs [25]. We specify ten well-known and common antipatterns, including Multi Service and Tiny Service, and automatically generate their detection algorithms. We validate the detection algorithms in terms of precision and recall on two systems developed independently. This validation demonstrates that our approach enables the specification and detection of SOA antipatterns with an average precision of 90% and recall of 97.5%.

6.21. Automated Measurement of Models of Requirements

One way to formalize system requirements is to express them using the object-oriented paradigm. In this case, the class model representing the structure of requirements is called a requirements metamodel, and requirements themselves are object-based models of natural-language requirements. We show that such object-oriented requirements are well-suited to support a large class of requirements metrics[28]. We define a requirements metamodel and use an automated measurement approach proposed in our previous work to specify requirements metrics. We show that it is possible to integrate 78 metrics from 11 different papers in the proposed framework. The software that computes the requirements metric values is fully generated from the specification of metrics.

6.22. Empirical Evidence of Large-Scale Diversity in API Usage of Object-Oriented Software

In this paper, we study how object-oriented classes are used across thousands of software packages. We concentrate on "usage diversity’", defined as the different statically observable combinations of methods called on the same object. We present empirical evidence that there is a significant usage diversity for many classes. For instance, we observe in our dataset that Java’s String is used in 2460 manners. We discuss the reasons of this observed diversity and the consequences on software engineering knowledge and research [56].

6.23. Efficient high-level abstractions for web programming

Writing large Web applications is known to be difficult. One challenge comes from the fact that the application’s logic is scattered into heterogeneous clients and servers, making it difficult to share code between both sides or to move code from one side to the other. Another challenge is performance: while Web applications rely on ever more code on the client-side, they may run on smart phones with limited hardware capabilities. These two challenges raise the following problem: how to benefit from high-level languages and libraries making code complexity easier to manage and abstracting over the clients and servers differences without trading this ease of engineering for performance? In [59], we present high-level abstractions defined as deep embedded DSLs in Scala that can generate efficient code leveraging the characteristics of both client and server environments. We compare performance on client-side against other candidate technologies and against hand written low-level JavaScript code. Though code written with our DSL has a high level of abstraction, our benchmark on a real world application reports that it runs as fast as hand tuned low-level JavaScript code.

6.24. Exploring Optimal Service Compositions in Highly Heterogeneous and Dynamic Service-Based Systems

Service-oriented pervasive systems, composed of a large number of devices with heterogeneous capabilities where devices’ resources are abstracted as software services, challenge the creation of high-quality composite applications. Resource heterogeneity, dynamic network connectivity, and a large number of highly distributed service providers complicate the process of creating applications with specific QoS requirements. Existing approaches to service composition control the QoS of an application solely by changing the set of participating concrete services which is not suitable for ad-hoc service-based systems characterised by high intermittent connectivity and resource heterogeneity. In [46], we propose a flexible way of formulating composition configurations suitable for such service-based systems. Our formulation proposes the combined consideration of the following factors that affect the QoS of a composed service: (a) service selection, (b) orchestration partitioning, and (c) orchestrator node selection. We show that the proposed formulation enables the definition of service composition configurations with 49% lower response time, 28% lower network latency, 36% lower energy consumption, and 13% higher success ratio compared to those defined with the traditional approach. In [45], we present the problem of efficiently exploring at runtime the serach space of possible configurations for a service orchestration with various Quality of Services.

TYREX Team

6. New Results

6.1. Multimedia Models and Formats

Modeling and authoring web content including rich media and interactions is still an open problem. We have drawn a reference state of the art of this area in [17]. We have also contributed to the cultural heritage domain through the experimentation of a multimedia production chain for digitized theatre performances based on semantic annotations [12].

In the context of the Claire project (see section 7.1.1), the results we obtained in 2013 in this area are:

- providing a data model which is format agnostic to cope with existing and future rendering systems. More specifically, we specified a chaptering component that includes the structuration and navigation features for continuous media such as video.
- prototyping a web environment for authoring such rich media (see section 5.2 . This authoring services are built as a contribution of the Mozilla Popcorn Maker project (Popcorn Maker).
- experimenting this environment for the production of the multimedia part of the first MOOC developed by OpenClassrooms, our main partner in this project.

6.2. XML Processing

In the area of XML processing, we obtained new results in several directions:

- We showed how to translate Schematron descriptions into the tree logic [15];
- We built the first IDE equipped with path reasoning capabilities [13];
- We showed that a whole class of logical combinators (or "macros") can be used as an intermediate language between the query language and the logical language [20]. This provides a gain in terms of succinctness for the logical formalism.
- We continued our work on a novel technique and a tool for the static type-checking of XQuery programs, using backward type inference.
- We made preliminary investigations on how to support backward navigation axes in the static type checking for XQuery [18].
- In a joint work with the Exmo team, we benchmarked solvers for deciding the problem of query containment for fragments of SPARQL [14].

We briefly review these results below.

6.2.1. Rule-Based Validation à la Schematron

One major concept in web development using XML is validation: checking whether some document instance fulfills structural constraints described by some schema. Over the last few years, there has been a growing debate about XML validation, and two main schools of thought emerged about the way it should be done. On the one hand, some advocate the use of validation with respect to complete grammar-based descriptions such as DTDs and XML Schemas. On the other hand, motivated by a need for greater flexibility, others argue for no validation at all, or prefer the use of lightweight constraint languages such as Schematron with the aim of validating only required constraints, while making schema descriptions more compositional and more reusable.

We built a compiler for Schematron [15]. This compiler takes a Schematron description as input and generates the corresponding constraints as a logical formula. We showed that validators used in each of these approaches share the same theoretical foundations, meaning that the two approaches are far from being incompatible. Our findings include that modal logic can be seen as a unifying formal ground for the construction of robust and efficient validators and static analyzers using any of these schema description techniques. This reconciles the two approaches from both a theoretical and a practical perspective, therefore facilitating any combination of them.

6.2.2. Integrated Development Environments with Path Reasoning Capabilities

One of the challenges in web development is to help achieving a good level of quality in terms of code size and runtime performance, for popular domain-specific languages such as XQuery, XSLT, and XML Schema. We presented the first IDE augmented with static detection of inconsistent XPath expressions that assists the programmer for simplifying the development and debugging of any application involving XPath expressions [13]. The tool is based on newly developed formal verification techniques based on expressive modal logics, which are now mature enough to be introduced in the process of software development. We further develop this idea in the context of XQuery for which we introduce an analysis for identifying and eliminating dead code automatically. This proof of concept aims at illustrating the benefits of equipping modern IDEs with reasoning capabilities.

6.2.3. Logical Combinators for Rich Type Systems

A popular technique in the static analysis for query languages relies on the construction of compilers that effectively translate queries into logical formulas. These formulas are then solved for satisfiability using an off-the-shelf satisfiability solver. A critical aspect in this approach is the size of the obtained logical formula, since it constitutes a factor that affects the combined complexity of the global approach.

We showed that a whole class of logical combinators (or “macros”) can be used as an intermediate language between the query language and the logical language [20]. Those logical combinators provide an exponential gain in succinctness over the corresponding explicit logical representation, yet preserve the typical exponential time complexity of the subsequent logical decision procedure. This opens the way for solving a wide range of problems such as satisfiability and containment for expressive query languages in exponential-time, even though their direct formulation into the underlying logic results in an exponential blowup of the formula size, yielding an incorrectly presumed two-exponential time complexity. We illustrated this from a very practical point of view on a few examples such as numerical occurrence constraints and tree frontier properties, which are concrete problems found in the XML world.

6.2.4. Backward type inference for XQuery

We have continued our work on the design of a novel technique for static type-checking of XQuery programs based on backward type inference. The tool looks for errors in the program by jointly analyzing the source code of the program, input and output schemas that respectively describe the sets of documents admissible as input and as output of the program. The crux and the novelty of our results reside in the joint use of backward type inference and a two-way logic to represent inferred tree type portions. This allowed us to design and implement a type-checker for XQuery which is more precise and supports a larger fragment of XQuery than the approaches previously proposed in the literature; in particular compared to the only few actually implemented static type-checkers such as the one in Galax. The whole system uses compilers and a satisfiability solver for deciding containment for two-way regular tree expressions. Our tool takes an XQuery program and two schemas S_{in} and S_{out} as input. If the program is found incorrect, then it automatically generates a counter-example valid w.r.t. S_{in} and such that the program produces an invalid output w.r.t S_{out}. This counter-example can be used by the programmer to fix the program.

6.2.5. XQuery and Static Typing: Tackling the Problem of Backward Axes

XQuery is a functional language dedicated to XML data querying and manipulation. As opposed to other W3C-standardized languages for XML (e.g. XSLT), it has been intended to feature strong static typing.

Currently, however, some expressions of the language cannot be statically typed with any precision. We argue that this is due to a discrepancy between the semantics of the language and its type algebra: namely, the values of the language are (possibly inner) tree nodes, which may have siblings and ancestors in the data. The types on the other hand are regular tree types, as usual in the XML world: they describe sets of trees. The type associated to a node then corresponds to the subtree whose root is that node and contains no information about the rest of the data. This makes navigational expressions using 'backward axes,' which return e.g. the siblings of a node, impossible to type.

In [18], we discussed how to solve this discrepancy and proposed a compromise: to use extended types representing possibly inner tree nodes in some key parts of a program, and to cut out the subtrees from their original context in the rest.

6.2.6. Semantic Web queries and μ-calculus

Querying the semantic web is mainly done through the SPARQL language or its extensions through paths and entailment regimes. Query containment is the problem of deciding if the answers to a query are included in those of another query for any queried database [4], [3]. This problem is very important for query optimization purposes. In the SPARQL context, it can be equally useful for distributing federated queries or for implementing schema-based access control. In order to experimentally assess implementation strengths and limitations, we provided a first SPARQL containment test benchmark. We studied the query demographics on DBPEDIA logs to design benchmarks for relevant query containment solvers. We tested available solvers on their domain of applicability on three different benchmark suites [14]. (i) tested solutions are overall functionally correct, (ii) in spite of its complexity, SPARQL query containment is practicable for acyclic queries, (iii) state-of-the-art solvers are at an early stage both in terms of capability and implementation.

This work has been developed in collaboration with the EXMO team. The benchmarks, results and software are available at http://sparql-qc-bench.inrialpes.fr.

6.3. Mixed Reality Environment

The concept of Mixed Reality comes from the fact that the real-virtual dichotomy is not sharp. Augmented Reality (AR) mode refers to all cases in which the auditory or visual display of a real environment is augmented by virtual sound or graphic objects. Pedestrian navigation is one of the numerous applications that fit into this field. Depending on the real speed of the user and on the real environment in which he moves (inside a building, ...), the system is augmented with synthetic audio instructions and points of interest. OpenStreetMap format has been extended to support navigation authoring and information related to the various passive or active location providers supported by IXE such as PDR, GPS and NFC.

6.3.1. Navigation Authoring

We defined a cue-based XML language (A2ML, for Advanced Audio Markup Language) using SMIL for internal and external synchronization of sound objects. A2ML is specified by a RELAX-NG grammar. A rule-based selector mechanism allows defining style sheets for OpenStreetMap (OSM) elements. This auditory display together with TTS makes our IXE browser accessible to visually impaired people. Format and Delivery for Mixed Reality Content IXE is based on an extended OSM data format with triggering zones, relations or groups with specific semantics and nodes or POIs whose URIs refer to content expressed in HTML5 and A2ML. Content delivery can be of two types, push or pull. Push content is coming from POIs which trigger when the user enters a new zone. This kind of content is very useful for navigation. We support it through a triggering specification that is inserted in the OpenStreetMap document. We use style sheets with rules to specify both the audio and visual rendering of the various types of OSM nodes. Pull content allows users to search detailed information about the artifacts that are located in the content referenced by the POI. Most of the time, this content is described using HTML5 and A2ML.

6.3.2. Location Provider Fusion

Pedestrian navigation can be done with several sensors. GPS locations are better for outside locations, PDR is useful to guide people indoor, but we can also use NFC tags, user proprioception, wifi... Our researches focus on a smart fusion of providers depending on sensor accuracy and on the context in which the person moves. We start by using Kalman Filter to smooth locations and disable jumps during the walk. These algorithms have been successfully tested during Venturi Y2 demo.

6.3.3. Map Rendering

We worked on offline map rendering around two solutions. The first one is based on an open source Android project called Mapsforge; it provides a tile generation mechanism from a given OpenStreetMap file and a tile caching system for fast rendering on mobile devices. We mainly enhanced the open source project by increasing the zoom level limitation (21 by default) to 24 for displaying indoor maps. The other solution on which we worked is SVG-oriented and based on OpenLayers (dedicated to web browsers). As the rendering uses SVG we are no longer limited by a maximum zoom level. On the other hand, the SVG drawing has to be fully designed by the author, as we don't support, for the time being, SVG file generation from an OpenStreetMap document. These two approaches are different and their uses depend on the desired level of customization of the rendering (generated automatically or manually).

URBANET Team

6. New Results

6.1. Characterizing and measuring urban networks

Participants: Marco Fiore, Diala Naboulsi, Razvan Stanica, Sandesh Uppoor

6.1.1. Properties of urban vehicular traffic and implications on mobile networking.

The goal of Sandesh Uppoor's PhD thesis [4] was to model and understand the mobility dynamics of high-speed vehicular users and their effect on wireless network architectures in an urban environment. Given the importance of developing the study on a realistic representation of vehicular mobility, we first survey the most popular approaches for the generation of synthetic road traffic and discuss the features of publicly available vehicular mobility datasets. Using original travel demand information of the population of a metropolitan area (Cologne area, Germany), detailed road network data and realistic microscopic driving models, we propose a novel state-of-art vehicular mobility dataset that closely mimics the real-world road traffic dynamics in both time and space [25]. We then study the impact of such mobility dynamics from the perspective of wireless cellular network architecture in presence of a real-world base station deployment. In addition, by discussing the effects of vehicular mobility on autonomous network architecture, we hint at the opportunities for future heterogeneous network paradigms and demonstrate how incomplete representations of vehicular mobility may result in over-optimistic network connectivity and protocol performance [8].

Motivated by the time-evolving mobility dynamics observed in our original dataset, we also propose an on line approach to predict near-future macroscopic traffic flows. We analyze the parameters affecting the mobility prediction in an urban environment and unveil when and where network resource management is more crucial to accommodate the traffic generated by users on-board. Such studies unveil multiple opportunities in transportation management either for building new roads, installing electric charging points, or for designing intelligent traffic light systems, thereby contributing to urban planning.

6.1.2. Feasibility of multi-hop vehicular communications in an urban environment.

Despite the growing interest in a real-world deployment of vehicle- to-vehicle communication, many topological features of the resulting vehicular network remain largely unknown. We still lack a clear understanding of the level of connectivity achievable in large-scale urban scenarios, of the availability and reliability of connected multi-hop paths, and of the evolution of such features over daytime. In [14], we investigate how the instantaneous topology of the vehicular network would look like in the case of a typical middle-sized European city, using the example of the Cologne mobility trace. Through a complex network analysis, we unveil the low connectivity, availability, reliability and navigability of the network, and exploit our findings to derive network design and usage guidelines.

6.1.3. Investigating the accuracy of mobile urban sensing.

Community urban sensing is one of the emerging applications enabled by the growing popularity of mobile user devices, like smartphones and in-vehicle monitoring systems. Such devices feature sensing and wireless communication capabilities, which enable them to sample large-scale phenomena, like air pollution and vehicular traffic congestion, and upload these data to the Internet. In [10], we focus on the above scenario and investigate the level of accuracy that can be achieved in estimating the phenomenon of interest through a mobile crowdsourcing application. Specifically, we take a signal processing-based approach and leverage results on signal reconstruction from sets of irregularly spaced samples. We apply such results to a realistic scenario where samples are collected by vehicular and pedestrian users, and study the accuracy level of the phenomenon estimation as the penetration rate of the sensing application varies.

6.1.4. Analysis of mobile network call detail records.

The growing ubiquity of mobile communications has offered researchers new possibilities to understand human mobility over the last few years. In [22], we analyze Call Detail Records (CDR) made available within the context of the Orange D4D Challenge, focusing on calls of individuals in the city of Abidjan, Ivory Coast, over a period of five months. Our results illustrate how aggregated CDR can be used to tell apart typical and special mobility behaviors, and demonstrate how macroscopic mobility flows extracted from these cellular network data reflect the daily dynamics of a highly populated city. We discuss how these macroscopic mobility flows can help solve problems in developing urban areas.

6.2. Scalable solutions for capillary networks

Participants: Isabelle Augé-Blum, Jin Cui, Marco Fiore, Ochirkhand Erdene-Ochir, Alexandre Mouradian, Hervé Rivano, Razvan Stanica, Fabrice Valois

6.2.1. Real-time wireless sensor networks.

Critical applications for WSNs are emerging, with real-time and reliability requirements. Critical applications are applications on which depend human lives and the environment: a failure of a critical application can thus have dramatic consequences. We are especially interested in anomaly detection applications (forest fire detection, landslide detection, intrusion detection, etc), which require bounded end to end delays and high delivery ratio. Few WSNs protocols of the literature allow to bound end to end delays. Among the proposed solutions, some allow to effectively bound the end to end delays, but do not take into account the characteristics of WSNs (limited energy, large scale, etc). Others take into account those aspects, but do not give strict guaranties on the end to end delays. In this sense, the PhD thesis of Alexandre Mouradian [2] proposes a real-time anomaly detection solution composed of:

- A virtual coordinate system which allows to discriminate nodes in a 2-hop neighborhood and to bound the number of hops between any source and the sink.
- A cross-layer protocol for WSNs (named RTXP) based on the proposed virtual coordinate system. Thanks to these coordinates it is possible to introduce determinism in the accesses to the medium and to bound the hop-count, this allows to bound the end to end delay. RTXP adapts its duty-cycle to the traffic loads and uses an opportunistic routing scheme to increase its delivery ratio. We show, by simulation, that RTXP outperforms real-time protocols of the literature for anomaly detection in WSNs under harsh radio conditions.
- A real-time aggregation scheme to mitigate the alarm storm problem which causes collisions and congestion and thus limit the network lifetime. This scheme is also based on the virtual coordinate system and is used before RTXP in order to reduce the number of similar alarms converging toward the sink.

6.2.2. Formal verification of wireless sensor networks protocols.

WSN protocols used by critical applications must be formally verified in order to provide the strongest possible guaranties: simulations and tests are not sufficient in this context, formal proofs of compliance with the specifications of the application have to be provided.. Unfortunately the radio link is unreliable and it is thus difficult to give hard guarantees on the temporal behavior of the protocols. Indeed, a message may experience a very high number of retransmissions and the temporal guarantee can only be given with a certain probability. This probability must meet the requirements of the application. Network protocols have been successfully verified on a given network topology without taking into account unreliable links. Nevertheless, the probabilistic nature of radio links may change the topology (links which appear and disappear). Thus instead of a single topology we have a set of possible topologies, each topology having a probability to exist. In [12], we propose a method that produces the set of topologies, checks the property on every topology, and gives the probability that the property is verified. This technique is independent from the verification technique, i.e. each topology can be verified using any formal method which can give a “yes” or “no” answer to the question: “Does the model of the protocol respect the property?”. We apply this method on the f-MAC protocol. We use

UPPAAL model checker as verification tool. We implement a tool that automatizes the process and thus show the feasibility of our proposition. We compare the results of the verification with simulation results. It appears that the verification is, as expected, conservative but not overly pessimistic compared to the simulated worst case. Besides we show that f-MAC is a reliable real-time protocol for WSNs (for up to 6 nodes), as we were not able to detect faults.

Moreover, in [2], a verification technique which mixes Network Calculus and Model Checking is proposed, in order to be both scalable and exhaustive. This technique consists in modeling the interaction of each node with the rest of the network with arrival curves and then to verify with UPPAAL that each node is capable of handling these interactions while meeting the deadlines. We apply this methodology in order to formally verify our pervious proposition, RTXP.

6.2.3. Reliability in wireless sensor networks.

WSN critical applications require the respect of time and reliability constraints. In [13], we provide a theoretical study of the reliability in WSNs. We define the reliability as the probability of success of an end-to-end transmission in the WSN. In this work, we use two radio propagation models : a basic model where the nodes have a set of neighbors they can communicate with, with a given probability, and the log-normal shadowing model, where probability of reception depends on the emitter-receiver distance. We determine the reliability of two routing schemes : unicast-based routing (classical routing) and broadcast-based routing (opportunistic routing). We conclude that the broadcast-based routing allows to reach a higher reliability than the unicast case. The main result is that we show the existence of a reliability bottleneck at the sink node in the case of the broadcast-based routing. We show that the addition of another sink improves the reliability of the network in this case.

6.2.4. Resiliency in wireless sensor networks.

Because of their open and unattended deployment, in possibly hostile environments, powerful adversaries can easily launch Denial-of-Service (Dos) attacks on wireless sensor networks, cause physical damage to sensors, or even capture them to extract sensitive information (encryption keys, identities, addresses, etc.). Consequently, the compromised node poses severe security and reliability concerns, since it allows an adversary to be considered as a legitimate node inside the network. To cope with these "insider" attacks, stemming from node compromise, "beyond cryptography" algorithmic solutions must be envisaged to complement the traditional cryptographic solutions. In this sense, in [1], we first propose the resiliency concept. Our goal is to propose a definition of the resiliency in our context (security of WSNs routing protocols) and a new metric to compare routing protocols. The originality of this metric is that we combine the graphical representation (qualitative information) with the aggregation method (quantitative information). We introduce a two dimensional graphical representation with multiple axes forming an equiangular polygon surface. This method allows to aggregate meaningfully several parameters and makes it easier to visually discern various trade-offs, thus greatly simplifying the process of protocol comparison.Secondly, we propose the protocol behaviors enhancing resiliency. Our proposition consists in three elements: (i) introduce random behaviors (ii) limit route length (iii) introduce data replication. Random behaviors increase uncertainty for an adversary, making the protocols unpredictable. Data replication allows route diversification between the sources and the sink, thus improving the delivery success and fairness. Limitation of the route length is necessary to reduce the probability of a data packet to meet a malicious insider along the route. The quantitative metric enables to propose a new resiliency taxonomy of WSNs routing protocols. According to this taxonomy, the gradient based routing is the most resilient when it is combined with the proposed behaviors. Thirdly, several variants of the gradient-based routing (classical and randomized) under more complex and realistic adversary model (several combined attacks) are considered to extend our simulations. Several values of bias are introduced to the randomized variants and two data replication methods (uniform and adaptive) are considered. Without attacks, the most biased variants without replications are the most efficient. However, under moderate attacks, the replication uniform is the most adapted, while under intense attacks, the replication adaptive is the most suitable. Finally, a theoretical study of the resiliency is introduced. We present an analytical study of the biased random walk routing under attacks. The influence of bias is evaluated and two replication methods that previously evaluated by simulations are considered. After presenting the delivery success and the energy consumption of all scenarios, we evaluate them with our resiliency metric. This study permits to confirm the results obtained with simulations and it shows that the bias is essential to enhance the resiliency of random routing.

6.2.5. Data aggregation in wireless sensor networks.

Data aggregation is a crucial problem in wireless sensor networks due to their constrained-energy and constrained-bandwidth nature. In [26], we highlight the aggregation benefits at the Network layer and MAC layer by modeling the energy consumption for some energy-efficient routing protocols and MAC protocols. Besides, we define two parameters, the aggregation ratio and the packet size coefficient to evaluate the efficiency of an aggregation method, and to discuss the trade-off. Additionally, we investigate the differences between time series and compressive sensing, which are representative state-of-the-art solutions for forecasting aggregation and compressing aggregation respectively.

6.2.6. Routing in delay-tolerant networks.

Delay-Tolerant Networks (DTN) model systems that are characterized by intermittent connectivity and frequent partitioning. Routing in DTNs has drawn much research effort recently. Since very different kinds of networks fall in the DTN category, many routing approaches have been proposed. In particular, the routing layer in some DTNs has information about the schedules of contacts between nodes and about data traffic demand. Such systems can benefit from a previously proposed routing algorithm based on linear programming that minimizes the average message delay. This algorithm, however, is known to have performance issues that limit its applicability to very simple scenarios. In [9], we propose an alternative linear programming approach for routing in Delay-Tolerant Networks. We show that our formulation is equivalent to that presented in a seminal work in this area, but it contains fewer LP constraints and has a structure suitable to the application of Column Generation (CG). Simulation shows that our CG implementation arrives at an optimal solution up to three orders of magnitude faster than the original linear program in the considered DTN examples.

6.2.7. Performance evaluation of vehicular communications.

Wireless vehicular networks face different problems and challenges, especially in a dense urban environment. In [23], we first characterize the different types of loss in vehicular networks: radio propagation problems, expired security messages, collision with one hop neighbor and collisions with hidden terminals. In a second step, we give the architecture of the wireless vehicular network and describe the Medium Access Control (MAC) quality of service mechanisms proposed by vehicular environment standards that aim at meeting the road drivers' expectation and increasing road safety. To complete this image, in [24], we provide a literature survey that covers the solutions proposed in order to enable critical dissemination of urgent messages and surpass the challenging vehicular dynamic topology. More particularly, we detail the following techniques: beaconing frequency reduction, transmit rate control, power control, adaptation of the contention window and adaptation of the carrier sense threshold.

6.2.8. Secure node localization in mobile ad-hoc networks.

A growing number of ad hoc networking protocols and location-aware services require that mobile nodes learn the position of their neighbors. However, such a process can be easily abused or disrupted by adversarial nodes. In absence of a priori trusted nodes, the discovery and verification of neighbor positions presents challenges that have been scarcely investigated in the literature. In [6], we address this open issue by proposing a fully distributed cooperative solution that is robust against independent and colluding adversaries, and can be impaired only by an overwhelming presence of adversaries. Results show that our protocol can thwart more than 99% of the attacks under the best possible conditions for the adversaries, with minimal false positive rates.

In a vehicular context, knowledge of the location of vehicles and tracking of the routes they follow are a requirement for a number of applications. However, public disclosure of the identity and position of drivers jeopardizes user privacy, and securing the tracking through asymmetric cryptography may have an exceedingly high computational cost. In [11], we address all of the issues above by introducing A-VIP, a lightweight privacy-preserving framework for tracking of vehicles. A-VIP leverages anonymous position beacons from vehicles, and the cooperation of nearby cars collecting and reporting the beacons they hear. Such information allows an authority to verify the locations announced by vehicles, or to infer the actual ones if needed. We assess the effectiveness of A-VIP through testbed implementation results.

6.3. Cellular network solutions

Participants: Marco Fiore, Anis Ouni, Hervé Rivano, Razvan Stanica, Fabrice Valois

6.3.1. Optimizing capacity and energy consumption in wireless mesh networks.

Wireless mesh networks (WMN) are a promising solution to support high data rate and increase the capacity provided to users, e.g. for meeting the requirements of mobile multimedia applications. However, the rapid growth of traffic load generated by the terminals is accompanied by an unsustainable increase of energy consumption, which becomes a hot societal and economical challenges. This thesis relates to the problem of the optimization of network capacity and energy consumption of wireless mesh networks. The network capacity is defined as the maximum achievable total traffic in the network per unit time.

The thesis of Anis Ouni [3] addresses this issue and is divided into four main parts. First, we address the problem of improvement of the capacity of 802.11 wireless mesh networks. We highlight some insensible properties and deterministic factors of the capacity, while it is directly related to a bottleneck problem. Then, we propose a joint TDMA/CSMA scheduling strategy for solving the bottleneck issue in the network.

Second, we focus on broadband wireless mesh networks based on time-frequency resource management. In order to get theoretical bounds on the network performances, we formulate optimization models based on linear programming and column generation algorithm. These models lead to compute an optimal offline configuration which maximizes the network capacity with low energy consumption. A realistic SINR model of the physical layer allows the nodes to perform continuous power control and use a discrete set of data rates.

Third, we use the optimization models to provide practical engineering insights on WMN. We briefly study the tradeoff between network capacity and energy consumption using a realistic physical layer and SINR interference model [27]. In particular, we show that power control and multi-rate functionalities allow to reach optimal throughput with lower energy consumption using a mix of single hop and multi- hop routes.

Finally, we focus on capacity and energy optimization for heterogeneous cellular networks. We develop optimization tools to calculate an optimal configuration of the network that maximizes the network capacity with low energy consumption. We then propose a heuristic algorithm that calculates a scheduling and partial sleeping of base stations in two different strategies, called LAFS and MAFS.

6.3.2. Sleep protocols for heterogeneous LTE networks.

The tremendous increase of the traffic demand in cellular networks imposes a massive densification of the traditional cellular infrastructure. The network architecture becomes heterogeneous, in particular 4G networks where LTE micro-eNodeBs are deployed to strengthen the coverage of macro-eNodeBs. This densification yields major issues related to the energy consumption of the infrastructure. Indeed, there is fixed and significant amount of energy required to run each additional node, whatever the traffic load of the network. Mitigating this fixed energy consumption is therefore a major challenge from a societal and economical viewpoint. Extensive researches about energy-saving highlight that to save energy the better strategy is to switch off the radio part of nodes. This is the heart of wireless sensor networks energy-saving strategies, even though the objective for WSN is to maximize the battery life of each individual nodes. In [18], we develop a parallel between the principles of WSN protocols and the requirements of cellular infrastructures. We then propose a distributed and localized algorithm to dynamically switch off and on the micro-eNodeBs of an LTE heterogeneous network following the traffic demand evolution in time and analyze it in terms of energy savings. We show that one can expect energy savings of approximately 12% when implementing sleep modes whereas the energy cost for sending the traffic decreases by 24%.

6.3.3. Content downloading through a vehicular network.

The focus of the work we present in [7] is twofold: information dissemination from infrastructure nodes deployed along the roads, the so-called Road-Side Units (RSUs), to passing-by vehicles, and content downloading by vehicular users through nearby RSUs. In particular, in order to ensure good performance for both content dissemination and downloading, the presented study addresses the problem of RSU deployment and reviews previous work that has dealt with such an issue. The RSU deployment problem is then formulated as an optimization problem, where the number of vehicles that come in contact with any RSU is maximized, possibly considering a minimum contact time to be guaranteed. Since such optimization problems turn out to be NP-hard, heuristics are proposed to efficiently approximate the optimal solution. The RSU deployment obtained through such heuristics is then used to investigate the performance of content dissemination and downloading through ns2 simulations. Simulation tests are carried out under various real-world vehicular environments, including a realistic mobility model, and considering that the IEEE 802.11p standard is used at the physical and medium access control layers. The performance obtained in realistic conditions is discussed with respect to the results obtained under the same RSU deployment, but in ideal conditions and protocol message exchange. Based on the obtained results, some useful hints on the network system design are provided.

6.3.4. Offloading Floating Car Data.

Floating Car Data (FCD) is currently collected by moving vehicles and uploaded to Internet-based processing centers through the cellular access infrastructure. As FCD is foreseen to rapidly become a pervasive technology, the present network paradigm risks not to scale well in the future, when a vast majority of automobiles will be constantly sensing their operation as well as the external environment and transmitting such information towards the Internet. In order to relieve the cellular network from the additional load that widespread FCD can induce, we study [16] a local gathering and fusion paradigm, based on vehicle-to-vehicle (V2V) communication. We show how this approach can lead to significant gain, especially when and where the cellular network is stressed the most. Moreover, we propose several distributed schemes to FCD offloading based on the principle above that, despite their simplicity, are extremely efficient and can reduce the FCD capacity demand at the access network by up to 95%.

6.3.5. Mobile malware propagation in vehicular networks.

The large-scale adoption of vehicle-to-vehicle (V2V) communication technologies risks to significantly widen the attack surface available to mobile malware targeting critical automobile operations. Given that outbreaks of vehicular computer worms self-propagating through V2V links could pose a significant threat to road traffic safety, it is important to understand the dynamics of such epidemics and to prepare adequate countermeasures. In [17], we perform a comprehensive characterization of the infection process of variously behaving vehicular worms on a road traffic scenario of unprecedented scale and heterogeneity. We then propose a simple yet effective data-driven model of the worm epidemics, and we show how it can be leveraged for smart patching infected vehicles through the cellular network in presence of a vehicular worm outbreak.

VEGAS Project-Team

5. New Results

5.1. Classical and probabilistic computational geometry

Participants: Xavier Goaoc, Guillaume Moroz, Sylvain Lazard, Marc Pouget.

5.1.1. Probabilistic complexity analysis of random geometric structures

Average-case analysis of data-structures or algorithms is commonly used in computational geometry when the more classical worst-case analysis is deemed overly pessimistic. Since these analyses are often intricate, the models of random geometric data that can be handled are often simplistic and far from "realistic inputs".

Complexity analysis of random geometric structures made simpler. In a joint work with Olivier Devillers and Marc Glisse (Inria Geometrica), we presented a new simple scheme for the analysis of geometric structures. While this scheme only produces results up to a polylog factor, it is much simpler to apply than the classical techniques and therefore succeeds in analyzing new input distributions related to smoothed complexity analysis. We illustrated our method on two classical structures: convex hulls and Delaunay triangulations. Specifically, we gave short and elementary proofs of the classical results that n points uniformly distributed in a ball in R^d have a convex hull and a Delaunay triangulation of respective expected complexities $\widetilde{\Theta}(n^{((d+1)/(d-1))})$ and $\widetilde{\Theta}(n)$. We then prove that if we start with n points well-spread on a sphere, e.g. an (ϵ, κ)-sample of that sphere, and perturb that sample by moving each point randomly and uniformly within distance at most δ of its initial position, then the expected complexity of the convex hull of the resulting point set is $\widetilde{\Theta}(\sqrt{(}n)^{(1-1/d)}\delta^{-(d-1)/(4d)})$. We presented these results in the *Symposium on Computational Geometry* 2013 [20].

Monotonicity of the number of facets of random polytopes. We also proved a result on the size of the convex hull K_n of n points sampled uniformly in a convex set K. More precisely, let $u_n^{K,i}$ be the expected number of facets of dimension i of the convex hull. We proved that, in the plane, $u_n^{K,0}$ is an increasing sequence. In higher dimension, if K is a convex, smooth, compact body, then we showed that the sequence $u_n^{K,d-1}$ is asymptotically increasing. This result, published in the *Electronic Communications in Probability* [13], was obtained in collaboration with Olivier Devillers and Marc Glisse (Inria Geometrica) and Matthias Reitzner (Osnabruck Univ.).

Worst-case silhouette size of random polytopes. Finally, we studied from a probabilistic point of view the size of the silhouette of a polyhedron. While the silhouette size of a polyhedron with n vertices may be linear for some view points, several experimental and theoretical studies show a sublinear behavior for a wide range of constraints. The latest result on the subject proves a bound in $\Theta(\sqrt{n})$ on the size of the silhouette from a random view point of polyhedra of size n approximating non-convex surfaces in a reasonable way [9]. This result considers the polyhedron given and average the sizes of the silhouettes over all view points. This year, we addressed the problem of bounding the worst-case size of the silhouette where the average is taken over a set of polyhedra. Namely, we consider random polytopes defined as the convex hull of a Poisson point process on a sphere in $\mathbb{R}^3$ such that its average number of points is n. We show that the expectation over all such random polytopes of the maximum size of their silhouettes viewed from infinity is $\Theta(\sqrt{n})$. This work was done in collaboration with Marc Glisse (Inria Geometrica) and Julien Michel (Université de Poitiers) [24].

5.1.2. Embedding geometric structures

We continued working this year on the problem of embedding geometric objects on a grid of $\mathbb{R}^3$. Essentially all industrial applications take, as input, models defined with a fixed-precision floating-point arithmetic, typically doubles. As a consequence, geometric objects constructed using exact arithmetic must be embedded on a fixed-precision grid before they can be used as input in other software. More precisely, the problem is, given a geometric object, to find a similar object representable with fixed-precision floating-point arithmetic, where similar means topologically equivalent, close according to some distance function, etc. We are working on the problem of rounding polyhedral subdivisions on a grid of $\mathbb{R}^3$, where the only known method, due to Fortune in 1999, considers a grid whose refinement depends on the combinatorial complexity of the input, which does not solve the problem at hand. This project is joint work with Olivier Devillers (Inria Geometrica) and William Lenhart (Williams College, USA).

5.1.3. Bounded-Curvature Shortest Paths

We considered the problem of computing shortest paths having curvature at most one almost everywhere and visiting a sequence of n points in the plane in a given order. This problem is a sub-problem of the Dubins Traveling Salesman Problem and also arises naturally in path planning for point car-like robots in the presence of polygonal obstacles. We showed that when consecutive waypoints are distance at least four apart, this question reduces to a family of convex optimization problems over polyhedra in $\mathbb{R}^n$. This result, done in collaboration with Hyo-Sil Kim (KAIST) was published in the *SIAM Journal on Computing* [15].

5.1.4. Approximating Geodesics in Meshes

A standard way to approximate the distance between any two vertices p and q on a mesh is to compute, in the associated graph, a shortest path from p to q that goes through one of k sources, which are well-chosen vertices. Precomputing the distance between each of the k sources to all vertices of the graph yields an efficient computation of approximate distances between any two vertices. One standard method for choosing k sources, which has been used extensively and successfully for isometry-invariant surface processing, is the so-called *Farthest Point Sampling* (FPS), which starts with a random vertex as the first source, and iteratively selects the farthest vertex from the already selected sources.

We analyzed the stretch factor $\mathcal{F}_{FPS}$ of approximate geodesics computed using FPS, which is the maximum, over all pairs of distinct vertices, of their approximated distance over their geodesic distance in the graph. We show that $\mathcal{F}_{FPS}$ can be bounded in terms of the minimal value $\mathcal{F}^*$ of the stretch factor obtained using an optimal placement of k sources as $\mathcal{F}_{FPS} \leq 2r_e^2\mathcal{F}^* + 2r_e^2 + 8r_e + 1$, where r_e is the ratio of the lengths of the longest and the shortest edges of the graph. This provides some evidence explaining why farthest point sampling has been used successfully for isometry-invariant shape processing. Furthermore, we showed that it is NP-complete to find k sources that minimize the stretch factor [25].

5.1.5. On Point-sets that Support Planar Graphs

A set of points is said universal if it supports a crossing-free drawing of any planar graph. For a planar graph with n vertices, if bends on edges of the drawing are permitted, universal point-sets of size n are known, but only if the bend-points are in arbitrary positions. If the locations of the bend-points must also be specified as part of the point-set, no result was known, and we prove that any planar graph with n vertices can be drawn on a universal set $\mathcal{S}$ of $O(n^2/\log n)$ points with at most one bend per edge and with the vertices and the bend points in $\mathcal{S}$. If two bends per edge are allowed, we show that $O(n \log n)$ points are sufficient, and if three bends per edge are allowed, $\Theta(n)$ points are sufficient. When no bends on edges are permitted, no universal point-set of size $o(n^2)$ is known for the class of planar graphs. We show that a set of n points in balanced biconvex position supports the class of maximum-degree-3 series-parallel lattices. These results were published this year in the journal *Computational Geometry: Theory and Application* [14].

We also considered the setting in which graphs are drawn with curved edges. We proved that, surprisingly, there exists a universal set of n points in the plane for which every n-vertex planar graph admits a planar drawing in which the edges are drawn as a circular arc. This result was presented in the *Canadian Conference on Computational Geometry* [17].

5.2. Non-linear computational geometry

Participants: Guillaume Moroz, Sylvain Lazard, Marc Pouget, Yacine Bouzidi, Laurent Dupont.

5.2.1. Solving bivariate systems and topology of algebraic curves

In the context of our algorithm Isotop for computing the topology of algebraic curves [4], we work on the problem of solving a system of two bivariate polynomials. We focus on the problem of computing a Rational Univariate Representation (RUR) of the solutions, that is, roughly speaking, a univariate polynomial and two rational functions which map the roots of the polynomial to the two coordinates of the solutions of the system.

Separating linear forms. We first presented an algorithm for computing a separating linear form of a system of bivariate polynomials with integer coefficients, that is a linear combination of the variables that takes different values when evaluated at distinct (complex) solutions of the system. In other words, a separating linear form defines a shear of the coordinate system that sends the algebraic system in generic position, in the sense that no two distinct solutions are vertically aligned. The computation of such linear forms is at the core of most algorithms that solve algebraic systems by computing rational parameterizations of the solutions and, moreover, the computation of a separating linear form is the bottleneck of these algorithms, in terms of worst-case bit complexity. Given two bivariate polynomials of total degree at most d with integer coefficients of bitsize at most τ, our algorithm computes a separating linear form in $\widetilde{O}_B(d^8 + d^7\tau)$ bit operations in the worst case, which decreases by a factor d^2 the best known complexity for this problem ($\widetilde{O}_B$ refers to the complexity where polylogarithmic factors are omitted and O_B refers to the bit complexity). This result was presented at the *International Symposium on Symbolic and Algebraic Computation* in 2013 [19] and submitted to a journal [23].

Solving bivariate systems & RURs. Given such a separating linear form, we also presented an algorithm for computing a RUR with worst-case bit complexity in $\widetilde{O}_B(d^7 + d^6\tau)$ and a bound on the bitsize of its coefficients in $\widetilde{O}(d^2 + d\tau)$. We showed in addition that isolating boxes of the solutions of the system can be computed from the RUR with $\widetilde{O}_B(d^8 + d^7\tau)$ bit operations. Finally, we showed how a RUR can be used to evaluate the sign of a bivariate polynomial (of degree at most d and bitsize at most τ) at one real solution of the system in $\widetilde{O}_B(d^8 + d^7\tau)$ bit operations and at all the $\Theta(d^2)$ real solutions in only $O(d)$ times that for one solution. These results were also presented at the *International Symposium on Symbolic and Algebraic Computation* in 2013 [18] and submitted to a journal [22].

This work is done in collaboration with Fabrice Rouillier (project-team Ouragan at Inria Paris-Rocquencourt).

5.2.2. Reflection through quadric mirror surfaces

We addressed the problem of finding the reflection point on a quadric mirror surfaces of a light ray emanating from a 3D point source P_1 and going through another 3D point P_2, the camera center of projection. This is a classical problem known as Alhazen's problem dating from around 1000 A.D. and based on the work of Ptolomy around 150 A.D. [31], [33]. We proposed a new algorithm for this problem based on our algorithm for the computation of the intersection of quadrics [7], [30] and using a characterization the reflection point as the tangential intersection point between the mirror and an ellipsoid with foci P_1 and P_2. The implementation is in progress. This work is done in collaboration with Nuno Gonçalves, University of Coimbra (Portugal).

5.2.3. Fast polynomial evaluation and composition

Evaluating a polynomial can be done with different evaluation schemes. The Hörner scheme for example allows to evaluate a polynomial of degree n in $O(n)$ arithmetic operations. When the cost of the arithmetic operations is constant, such as in floating point arithmetic, this leads to $O(n)$ binary operations. However, with integers, the size of the elements grows linearly after each multiplication and this may lead to $O(n^2)$ binary operations. This problem arises also with polynomial composition.

The best way to handle these cases is to use divide-and-conquer algorithms to keep a linear complexity in the degree up to logarithmic factors. State-of-the-art algorithms split at the highest pure power of 2 lower or equal to $\frac{n}{2}$. However when n is not a pure power of 2, this strategy might not be optimal.

We developed the library *fast_polynomial* to explore different divide-and-conquer schemes and observed notably that splitting at $\lfloor \frac{n}{2} \rfloor$ is more efficient in some cases. In particular, this evaluation scheme does not suffer the staircase effect observed in state-of-the-art evaluations. Experimentally, it is always faster than our own implementation of the classical divide-and-conquer scheme, and faster than the state of the art library *Flint 2* when the degree of the input polynomial is between 2^k and $2^k + 2^{k-1}$. These results are presented in the technical report [26].

5.3. Combinatorics and combinatorial geometry

Participant: Xavier Goaoc.

5.3.1. Simplifying inclusion-exclusion formulas

In a joint work with Jiří Matoušek, Pavel Paták, Zuzana Safernová, Martin Tancer (Charles University, Prague, Czech republic), we worked on computing simplified inclusion-exclusion formulas. Let $\mathcal{F} = \{F_1, F_2, ..., F_n\}$ be a family of n sets on a ground set S, such as a family of balls in $\mathbb{R}^d$. For every finite measure μ on S, such that the sets of $\mathcal{F}$ are measurable, the classical *inclusion-exclusion formula* asserts that $\mu(F_1 \cup F_2 \cup \cdots \cup F_n) = \sum_{I:\varnothing \neq I \subseteq [n]} (-1)^{|I|+1} \mu \left(\bigcap_{i \in I} F_i \right)$; that is, the measure of the union is expressed using measures of various intersections. The number of terms in this formula is exponential in n, and a significant amount of research, originating in applied areas, has been devoted to constructing simpler formulas for particular families $\mathcal{F}$. We provide an upper bound valid for an arbitrary $\mathcal{F}$: we show that every system $\mathcal{F}$ of n sets with m nonempty fields in the Venn diagram admits an inclusion-exclusion formula with $m^{O(\log^2 n)}$ terms and with ± 1 coefficients, and that such a formula can be computed in $m^{O(\log^2 n)}$ expected time. We also construct systems of n sets on n points for which every valid inclusion-exclusion formula has the sum of absolute values of the coefficients at least $\Omega(n^{3/2})$. This work was presented at the EUROCOMB conference [21] in September 2013.

5.3.2. Helly numbers of acyclic families

In a joint work with Éric Colin de Verdière (CNRS-ENS) and Grégory Ginot (IMJ-UPMC), we worked on applications of algebraic topology to combinatorial geometry, and more precisely on extending classical results on nerve complexes. The nerve complex of a family is an abstract simplicial complex that encode its intersection patterns. Nerves are widely used in computational geometry and topology, in particular in reconstruction problems where one aims at inferring the geometry of an object from a point sample while guaranteeing that the topology is correct. Indeed, the *nerve theorem* ensures that the nerve of a family of geometric objects has the same “topology” (formally: homotopy type) as the union of the objects whenever they form a “good cover”, that is, when any subset of the objects has an empty or contractible intersection. We relaxed this “good cover” condition to allow for families of non-connected sets. We defined an analogue of the nerve, called the *multinerve*, that is suitable for general acyclic families, and we proved that this combinatorial structure enjoys an analogue of the nerve theorem. Using multinerve, we could derive a new *topological Helly-type theorem* for acyclic families that generalizes previous results of Amenta, Kalai and Meshulam, and Matoušek. We finally used this new Helly-type theorem to (re)prove, in a unified way, bounds on transversal Helly numbers in *geometric transversal theory*. This article was submitted to the journal *Advances in mathematics* in 2012; it was accepted in 2013 and will appear in 2014 [16].

5.3.3. Set systems and families of permutations with small traces

In a joint work with Otfried Cheong (KAIST, South Korea) and Cyril Nicaud (Univ. Marne-La-Vallée), we studied two problems of the following flavor: how large can a family of combinatorial objects defined on a finite set be if its number of distinct “projections” on any small subset is bounded? We first consider set systems, where the “projections” is the standard notion of trace, and for which we generalized Sauer’s Lemma on the size of set systems with bounded VC-dimension. We then studied families of permutations, where the “projections” corresponds to the notion of containment used in the study of permutations with excluded patterns, and for which we delineated the main growth rates ensured by projection conditions. One of our motivations for considering these questions is the “geometric permutation problem” in geometric transversal theory, a question that has been open for two decades. This work was submitted to the European Journal of Combinatorics in 2012 and published in 2013 [12].

VERIDIS Project-Team

6. New Results

6.1. Automated and Interactive Theorem Proving

6.1.1. Using symmetries in SMT

Participants: David Déharbe, Pascal Fontaine, Stephan Merz.

Joint work with Carlos Areces, Raúl Fervari, Guillaume Hoffmann, and Ezequiel Orbe at Universidad Nacional de Córdoba (see also section 8.2).

Methods exploiting problem symmetries have been very successful in several areas including constraint programming and SAT solving. We proposed similar techniques for enhancing the performance of SMT-solvers by detecting symmetries in the input formulas and using them to prune the search space of the SMT algorithm. These techniques are based on the concept of (syntactic) invariance by permutation of symbols. In 2011, we presented a technique restricted to constants but which exhibited impressive results for some categories of formulas [4]; this technique was quickly implemented in major SMT solvers, including CVC4 and Z3.

In 2013, we proposed, together with our colleagues at the University of Córdoba, Argentina, a more general approach to detect symmetries in an SMT context. These techniques are based on graph isomorphisms, and the Schreier-Sims algorithm for improving the presentation of the symmetries. This work was published at the SMT workshop 2013 [21].

6.1.2. Computing minimal models (prime implicants)

Participants: David Déharbe, Pascal Fontaine.

Joint work with Daniel Le Berre and Bertrand Mazure from the CRIL laboratory in Lens, France.

Model checking and counter-example guided abstraction refinement are examples of applications of SAT solving that require the production of models for satisfiable formulas. Instead of giving a truth value to every variable, it is usually preferable to provide an implicant, i.e. a partial assignment of the variables such that every full extension is a model for the formula. An implicant is *prime* if every assignment is necessary. Since prime implicants contain no literal irrelevant for the satisfiability of the formula, they are considered as highly refined information.

In 2013, we proposed a novel algorithm that uses data structures found in modern CDCL SAT solvers for efficiently computing prime implicants starting from an existing model. The original aspects are (1) the algorithm is based on watched literals and a form of propagation of required literals, adapted to CDCL solvers, (2) the algorithm works not only on clauses, but also on generalized constraints, and (3) for clauses (and more generally, for cardinality constraints) the complexity of the algorithm is linear in the size of the constraints. We implemented and evaluated the algorithm with the Sat4j library. This work gave rise to a publication at the FMCAD 2013 international conference [13].

6.1.3. Encoding TLA+ proof obligations for SMT solvers

Participants: Stephan Merz, Hernán Vanzetto.

The TLA$^+$ proof system TLAPS (see section 5.2) is being developed within a project at the MSR-Inria Joint Centre to which we contribute. Typical proof obligations that arise during the verification of TLA$^+$ specifications mix reasoning about sets, functions, arithmetic, tuples, and records. In previous work [47], we have developed translations from TLA$^+$ set theory to SMT-Lib, the standard input language of SMT solvers. The main challenge has been to design a sound translation from untyped TLA$^+$ to the multi-sorted first-order logic that underlies SMT-Lib. Our solution is based on an incomplete type inference based on "typing hypotheses" present in TLA$^+$ proof obligations. When type inference fails, we fall back to an "untyped" encoding where interpreted sorts such as integers are injected into a designated sort of TLA$^+$ values, and proof obligations corresponding to well-sortedness conditions must be discharged during the proof.

In 2013, we have stabilized and extended the type inference, based on a more expressive type system that includes dependent types, predicate types, and subtyping. The new type system is able to solve many more typing conditions during the translation of proof obligations and thus improves both the scope and the efficiency of the SMT backend. It has been implemented as part of the SMT backend of TLAPS, and an article describing the type system has been submitted. A full description will appear in the PhD thesis of Hernán Vanzetto, expected to be defended in early 2014.

6.1.4. Formalization of stuttering invariance in temporal logic

Participant: Stephan Merz.

Extending our previous formalization in the interactive proof assistant Isabelle/HOL of the concept of stuttering invariance, we formally proved that a property expressible in propositional temporal logic is stuttering invariant if and only if it is equivalent to a formula using only the *until* temporal operator (and in particular not the *next-time* operator). The formalization follows the proof in the classical paper by Peled and Wilke [49]. It allowed us to uncover and correct an error in the proof that had previously not been known. The corresponding extended version of the Isabelle proof development has been accepted at the Archive of Formal Proofs.

6.1.5. Superposition modulo theories

Participants: Noran Azmy, Christoph Weidenbach.

We are currently in a transition phase moving SPASS from a first-order logic prover to a first-order logic prover over theories SPASS(T), in particular arithmetic. Our experience in combining SPASS with interactive verification systems such as TLAPS or Isabelle shows that this is a mandatory step in improving automation [46], [34]. Meanwhile we have built the theoretical foundations [41], [40], [43] for combining superposition with theories which we now turn into algorithmic solutions. This makes an overall reimplementation of SPASS necessary. As a first step we reimplemented and improved our clause normal form transformation [11].

In particular, we want to support integer theories and modulo reasoning [15], as it is often used in distributed algorithms [46]. We have built first implementations of arithmetic modules which we want to combine in 2014 to a first version of SPASS(T).

6.1.6. Presburger Arithmetic in Compiler Optimization

Participants: Marek Košta, Thomas Sturm.

One of our focuses in 2013 was the application of SMT-solvers in new and different problem areas. We started a fruitful cooperation with the Compiler Lab at the Saarland University, Germany on compilation of data-parallel languages.

Data-parallel languages like OpenCL and CUDA are an important means to exploit the parallel computational capabilities of today's computing devices. However, the historical development of data-parallel languages stemming from GPUs plays a crucial role when compiling them for a SIMD (Single Instruction Multiple Data) CPU: on the CPU, one has to emulate dynamic features that on GPU are implemented in the hardware. This difference gives rise to several problems that have to be dealt with during the compilation process.

Our work [15] considers compilation of OpenCL programs for CPUs with SIMD instruction sets. It turns out that SMT-solvers can be used to generate more efficient CPU code. The lack of some dynamic features on CPU implies that one wants to statically decide whether or not certain memory operations access consecutive addresses. Our approach formalizes the notion of consecutivity and algorithmically reduces the static decision to satisfiability problems in Presburger Arithmetic. This is where SMT-solvers come into play. To make an application of an off-the-shelf SMT solver feasible, a preprocessing technique on the SMT problems was introduced. Combining three different systems (computer algebra system REDLOG, SMT-solver Z3, and an OpenCL driver developed in the Compiler Lab), a proof-of-concept system based on our approach was developed. The system generated more efficient code than any other state-of-the-art OpenCL compiler.

Further development is needed to turn the proof-of-concept system mentioned above into one integrated software system. To achieve this, the redundant combination of three heterogeneous systems needs to be replaced by a coherent library offering the same functionality. The work [23] presents the development of such a novel library. The library provides functions to fully automatize the approach proposed in the previous work. It is capable of parallel computations by means of threads and processes and uses an SMT-solver library to carry out the needed computations. To create the final system, the integration of the library with the OpenCL driver needs to be done. This final step is left for future work.

6.1.7. Non-Linear SMT-Solving

Participants: Marek Košta, Thomas Sturm.

In [42] de Moura and Jovanović give a novel satisfiability procedure for the theory of the reals. The procedure uses DPLL-style techniques to search for a satisfying assignment. In case of a conflict, cylindrical algebraic decomposition (CAD) [38] is used to guide the search away from the conflicting state: on the basis of one conflicting point, the procedure learns to avoid in the future an entire CAD cell containing the point. The function realizing this learning is the crucial ingredient that makes the DPLL-style search possible at all. Unfortunately, it is the main computational bottleneck of the whole procedure.

The work of Brown [35] develops a more efficient learning function for the case when the cell to-be learned is full-dimensional. In collaboration with Prof. Brown (United States Naval Academy, USA), we extend this to the general case. While restricting to one cell is quite straightforward for the base and lifting phases of a CAD algorithm, our approach is able to optimize the projection phase as well. This requires a thorough analysis of available geometric infomation and properties of the involved projection operator. Our cell construction algorithm is able to produce bigger cells and it is faster than the approach used in [42]. Both of these are benefits, because a bigger cell means a better generalization of the conflicting assignment. Prototypical implementation of our cell construction algorithm gives very promising results on various kinds of problems. Its elaborate implementation and integration with an DPLL engine within the computer algebra system REDLOG is left for future work. A publication has been submitted to the Journal of Symbolic Computation.

6.1.8. Towards Tropical Decision for NLA

Participant: Thomas Sturm.

Inspired by problems related to stability analysis of chemical reaction networks we have developed an incomplete decision procedure for satisfiability in nonlinear real arithmetic. A first implemented version focuses on specific situations where all variables are known to be stricly positive, which naturally occurs in many scientific contexts. Furthermore, only one single equation is considered. The principal *tropical* approach is, after reducing the problem to finding a point with positive value for f in the considered equation $f = 0$, to consider instead of f only the exponent tuples of the contained summands as points in $\mathbb{Z}^n$. On that basis dominating summands can be identified using LP techniques.

In our particular application discussed in [14], we were able to solve problems, which are intractable even by numerical methods: Typical input equations had around 6000 summands and up to seven variables of degrees between 4 and 9. The methods failed in only 3 percent of the 496 considered input problems.

We are currently generalizing the approach to the general case where variables can have arbitrary values. Furthermore, as it is well known that every existential decision problems over the reals can be equi-satisfiably encoded into one equation, we are aiming at a corresponding general procedure as a long-term research goal.

6.1.9. Hierarchical superposition for arithmetic

Participant: Uwe Waldmann.

Many applications of automated deduction require reasoning in first-order logic modulo background theories, in particular some form of integer arithmetic. A major unsolved research challenge is to design theorem provers that are “reasonably complete” even in the presence of free function symbols ranging into a background theory sort. The hierarchic superposition calculus of Bachmair, Ganzinger, and Waldmann already supports such symbols, but not optimally. We have introduced a novel form of clause abstraction, a core component in the hierarchic superposition calculus for transforming clauses into a form needed for internal operation. We have also demonstrated that hierarchic superposition is refutationally complete for linear integer or rational arithmetic, even if one considers the standard model semantics rather than the first-order semantics, provided that all background-sorted terms in the input are either ground or variables (variables with integer offsets can be permitted in certain positions).

6.2. Proved development of algorithms and systems

6.2.1. Incremental development of distributed algorithms

Participants: Dominique Méry, Manamiary Andriamiarina.

Joint work with Mohammed Mosbah and Mohammed Tounsi from the LABRI laboratory in Bordeaux, France.

The development of distributed algorithms and, more generally, of distributed systems, is a complex, delicate, and challenging process. The approach based on refinement helps to gain formality by using a proof assistant, and proposes to apply a design methodology that starts from the most abstract model and leads, in an incremental way, to the most concrete model, for producing a distributed solution. Our work helps formalizing pre-existing algorithms, developing new algorithms, as well as developing models for distributed systems.

Our research was initially supported by the ANR project RIMEL (see http://rimel.loria.fr). More concretely, we aim at an integration of the correct-by-construction refinement-based approach into the *local computation* programming model. The team of LABRI develops an environment called VISIDIA (http://visidia.labri.fr) that provides a toolset for developing distributed algorithms expressed as a set of rewriting rules of graph structures. The simulation of rewriting rules is based on synchronization algorithms, and we have developed these algorithms by refinement [20].

In particular, we show how state-based models can be developed for specific problems and how they can be simply reused by controlling the composition of state-based models through the refinement relationship. Traditionally, distributed algorithms are supposed to run on a fixed network, whereas we consider a network with a changing topology.

The contribution is related to the development of proof-based patterns providing effective help to the developer of formal models of applications [10]. Our patterns simplify the development of distributed systems using refinement and temporal logic. Moreover, we have especially evaluated the extension of the scope of Event B by proposing a technique for integrating fairness in the development of distributed algorithms [17].

6.2.2. Modeling Medical Devices

Participant: Dominique Méry.

Formal modelling techniques and tools have attained sufficient maturity for formalizing highly critical systems in view of improving their quality and reliability, and the development of such methods has attracted the interest of industrial partners and academic research institutions. Building high quality and zero-defect medical software-based devices is a particular domain where formal modelling techniques can be applied effectively. Medical devices are very prone to showing unexpected system behaviour in operation when traditional methods are used for system testing. Device-related problems have been responsible for a large number of serious injuries. Officials of the US Food and Drug Administration (FDA) found that many deaths and injuries related to these devices are caused by flaws in product design and engineering. Cardiac pacemakers and implantable cardioverter-defibrillators (ICDs) are among the most critical medical devices and require closed-loop modelling (integrated system and environment modelling) for verification purposes before obtaining a certificate from the certification bodies.

Clinical guidelines systematically assist practitioners in providing appropriate health care in specific clinical circumstances. Today, a significant number of guidelines and protocols are lacking in quality. Indeed, ambiguity and incompleteness are likely anomalies in medical practice. The analysis of guidelines using formal methods is a promising approach for improving them.

In [9], we propose a refinement-based methodology for complex medical systems design, which possesses all the required key features. A refinement-based combined approach of formal verification, model validation using a model-checker and refinement chart is proposed in this methodology for designing a high-confidence medical device. Furthermore, we show the effectiveness of this methodology for the design of a cardiac pacemaker system.

Inappropriate mode transitions can be a common cause of mishaps in complex health-care systems. In [19], we present an approach for formalizing and reasoning about optimal mode transition in a health-care system that uses several operating modes in various operating states. Modes are formalized and their relation to a state-based formalism is established through a refinement approach. The efficiency of this approach is presented by formalizing an ideal operating mode transition of a cardiac pacemaker case study. An incremental approach is used to develop the system and its detailed design is verified through a series of refinements. In this way, we show how to improve system structuring, elicitation of system assumptions and expected functionality, as well as requirement traceability using modes in state-based modeling. Models are expressed in the Event B [25] modeling language, and they are validated by the model checker ProB.

Finally, in a joint work with colleagues of the CRAN laboratory in Nancy, we have completed a joint project with Airbus on the integration of physiological features in the development of systems like maintenance systems.

6.2.3. Analysis of real-time Java programs

Participants: Jingshu Chen, Marie Duflot-Kremer, Pascal Fontaine, Stephan Merz.

Joint work with Nadezhda Baklanova, Jan-Georg Smaus, Wilmer Ricciotti, and Martin Strecker at IRIT Toulouse, France, funded by EADS Foundation (see also section 7.1).

We investigate techniques for the formal verification of programs written in a dialect of Java that includes real-time annotations. Inspired by Safety-Critical Java [36], our partners in Toulouse developed a formal semantics for that dialect in Isabelle/HOL. In joint work, we have designed translations of programs to respectively timed automata and to SMT-Lib for analysis with the Uppaal model checker and with SMT solvers. We are evaluating the features and the scalability of the two approaches, and also plan to formally prove the soundness of the translations based on the semantics formalized in Isabelle.

6.2.4. Fundamentals of Network Calculus in Isabelle/HOL

Participant: Stephan Merz.

Joint work with Marc Boyer from ONERA (Toulouse, France) and Loïc Fejoz, Etienne Mabille and Nicolas Navet from RealTime at Work (RTaW, Nancy).

Network Calculus [45] is a well-established theory for the design and analysis of embedded networks. Based on the $(\min, +)$ dioid, it allows a network designer to compute upper bounds for delay and buffer sizes in networks. The theory is supported by several commercial and open-source tools and has been used in major industrial applications, such as the design and certification of the Airbus A380 AFDX backbone. Nevertheless, it is difficult for certification authorities to assess the correctness of the computations carried out by the tools supporting Network Calculus, and we propose the use of *result certification* techniques for increasing the confidence in the Network Calculus toolchain. We have formalized parts of the theory underlying Network Calculus in the proof assistant Isabelle/HOL. We have also developed a prototype analyzer that outputs traces of its computations so that they can be certified using Isabelle. Our work has been published at the conferences EUCASS and ITP [16], [24], and we have submitted a project proposal to ANR together with ONERA, RTaW, Kalray, Eurocopter, and Astrium. Unfortunately, the project was not granted, and future work on this promising subject is on hold.

6.2.5. Modeling and verifying the Pastry routing protocol

Participants: Tianxiang Lu, Stephan Merz, Christoph Weidenbach.

As a significant case study for the techniques that we are developing within VeriDis, we are modeling and verifying the routing protocol of the Pastry algorithm [37] for maintaining a distributed hash table in a peer-to-peer network. As part of his PhD work, Tianxiang Lu developed a TLA$^+$ model of the Pastry routing protocol, and has uncovered several problems in the existing presentations of the protocol in the literature that could lead to network partitioning.

He proposed a novel variant of the protocol and proved its correctness under the strong assumption that no nodes leave the network, using TLAPS (see section 5.2). He also demonstrated that the protocol could not work if arbitrary nodes are allowed to leave; it is not clear at this point under what reasonable assumptions the protocol can be made to work. The correctness proofs contain almost 15000 interactions and constitutes the largest case study carried out so far using TLAPS. Tianxiang Lu defended his thesis at the end of November 2013; a journal publication describing this work is in preparation.

6.2.6. Bounding message length in attacks against security protocols

Participant: Marie Duflot-Kremer.

Joint work with Myrto Arapinis from the University of Birmingham, UK.

Security protocols are short programs that describe communication between two or more parties in order to achieve security goals. Despite the apparent simplicity of such protocols, their verification is a difficult problem and has been shown to be undecidable in general. This undecidability comes from the fact that the set of executions to be considered is of infinite depth (an infinite number of protocol sessions can be run) and infinitely branching (the intruder can generate an unbounded number of distinct messages). Several attempts have been made to tackle each of these sources of undecidability. We have shown [30] that, under a syntactic and reasonable condition of "well-formedness" on the protocol, we can get rid of the infinitely branching part. Following this conference publication, we have submitted a journal version of this result extending the set of security properties to which the result is applicable, in particular including authentication properties.

6.2.7. Evaluating and verifying probabilistic systems

Participant: Marie Duflot-Kremer.

Joint work with colleagues at ENS Cachan and University Paris Est Créteil.

Since its introduction in the 1980s, model checking has become a prominent technique for the verification of complex systems. The aim was to decide whether or not a system was fulfilling its specification. With the rise of probabilistic systems, new techniques have been designed to verify this new type of systems, and appropriate logics have been proposed to describe more subtle properties to be verified. However, some characteristics of such systems cannot fall in the field of model checking. The aim is thus not to tell wether a property is satisfied but how well the system performs with respect to a certain measure. We have designed a statistical tool for tackling both performance and verification issues. Following several conference talks, two journal papers have been written. The first one presents the approach in details with a few illustrative applications. The second one focuses on biological application, and more precisely the use of statistical model checking to detect and measure several indicators of oscillating biological systems.

VIRTUAL PLANTS Project-Team

5. New Results

5.1. Analysis of structures resulting from meristem activity

5.1.1. Acquisition and design of plant geometry

Participants: Chakkrit Preuksakarn [Univ Kasertsart, Thaïland], Mathilde Balduzzi, Frédéric Boudon, Christophe Pradal, Christophe Godin, Christian Fournier.

Virtual 3D model of plants are required in many areas of plant modeling. They can be used for instance to simulate physical interaction of real plant structures with their environment (ligth, rain, wind, pests, ...), to set up initial conditions of growth models or to assess their output against real data. In the past decade, methods have been developed to digitize plant architectures in 3D [59], [53]. These methods are based on direct measurements of position and shape of every plant organ in space. Although they provide accurate results, they are particularly time consuming. More rapid and automated methods are now required in order to collect plant architecture data of various types and sizes in a systematic way. In this aim, we explore the use of pictures, laser scanner, video and direct sketching.

- *Reconstruction of plant architecture from 3D laser scanner data.* (Chakkrit Preuksakarn, Mathilde Balduzzi, Frédéric Boudon, Christophe Godin, Pascal Ferraro [Labri, Bordeaux])

  We investigate the possibility to use 3D laser scanners to automate plant digitizing. We are developping algorithms to reconstruct branching systems without leaves or foliage from scanner data or from scan simulated on plant mock-up obtained using different digitizing method.

  For branching system, a number of automatic methods were proposed in the literature to reconstruct plausible branching structures from laser scanner data. The question of their assessment and accuracy is however critical for further exploitation in biological applications. For this, we developed an evaluation pipeline to assess reconstructions accuracy. A laser scan database on which experts built reference reconstructions is used as a basis of the evaluation. The pipeline is given two structures and compares both their elements and their organization. Similar elements are identified based on geometric criteria using an optimization algorithm. The organization of these elements is then compared and their similarity quantified. Two indices of geometrical and structural similarities are defined, and automatic reconstructions can thus be compared to reference structures to assess their accuracy. The method is successful at capturing the variation of similarities between two structures as different levels of noise are introduced. A first comparative evaluation of the different methods of the literature has been designed and conducted. This work has been presented at the FSPM conference and submitted at Annals of Botany for its special issue.

  We also investigated the reconstruction of tree foliage from 3D scans. Such elements are crucial to study the interaction of the plant with its environment. However, laser scans contain outliers on the silhouette of the scans that make the meshing of the pointset difficult. New generation of laser scanners provide intensity of the laser reflected on the surface of scanned objects. This intensity depends of the distance to the object, its optical property and the incidence angle. A first work on this topic showed that after correcting the distance effect, the incidence angle can be deduced from the intensity. From this result, we developed a reconstruction pipeline using the scan intensities and based on Shape-From-Shading approaches. This work has also been presented at the FSPM conference.

- *Reconstruction from video.* (Frédéric Boudon, Jerome Guenard [IRIT, Toulouse], Geraldine Morin [IRIT, Toulouse], Pierre Gurdjos [IRIT, Toulouse], Vincent Charvillat [IRIT, Toulouse])

../../../../projets/virtual_plants/IMG/cherry_tree.png

Figure 2. Reconstruction of a cherry tree. Left: photograph of the original tree. Right: 3D reconstruction from a laser scan rendered and integrated on the same background.

Even if mature computer vision techniques allow the reconstruction of challenging 3D objects from images, due to high complexity of plant topology, dedicated methods for generating 3D plant models must be devised. In collaboration with our colleagues from IRIT, Toulouse, we developed an analysis-by-synthesis method which generates 3D models of a plant from both images and a priori knowledge of the plant species.

Our method is based on a skeletonisation algorithm which allows to generate a possible skeleton from a foliage segmentation. Then, a 3D generative model, based on a parametric model of branching systems that takes into account botanical knowledge is built. This method extends previous works by constraining the resulting skeleton to follow hierarchical organization of natural branching structure. 3D models are then generated. A reprojection of the models can be compared with the original image to assess the visual accuracy of the reconstruction. We optimise the parameter values of the generative model based on the reprojection criterion. Realistic results are obtained on different species of plants, in particular vineyards. This work has been presented at the ISVC conference and published in LNCS [46].

- *Reconstruction of virtual fruits from pictures.* (Mik Cieslak, Nadia Bertin [Inra, Avignon], Frédéric Boudon, Christophe Godin, Michel Genard [Inra, Avignon], Christophe Goz-Bac [Université Montpellier 2])

  This research theme is supported by the Agropolis project Fruit3D.

  The aim of this work is to provide methods for generating fruit structure that can be integrated with models of fruit function and used to investigate such effects. To this end, we have developed a modeling pipeline in the OpenAlea platform that involves two steps: (1) generating a 3D volumetric mesh representation of the entire fruit, and (2) generating a complex network of vasculature that is embedded within this mesh. To create the 3D volumetric mesh, we use reconstruction algorithms from the 3D mesh generation package of the Computational Geometry Algorithms Library (CGAL). To generate the pattern of vasculature within this volumetric mesh, we use a Space Colonisation Algorithm that populates the volume of the fruit by simulating competition for space of the vasculature. We have applied our modeling pipeline to generate the internal and external geometry of a cherry tomato fruit using Magnetic Resonance Imaging data as input. These studies demonstrate the possibility to create species-specific models of fruit structure with relatively low effort [31]. These volumetric meshes are then combined with models of function to form integrative computational fruit models, which will help to investigate the effects of fruit structure on quality (see section 5.3.2).

- *Reconstruction of root structures.* (Julien Diener, Fredéric Boudon, Christophe Pradal, Christophe Godin, Philippe Nacry [BPMP, INRA], Christophe Périn [AGAP, CIRAD], Anne Dievart [AGAP, CIRAD], Xavier Draye [UCL, Belgium])

  Similarly to aerial part of plants, some needs for reconstruction procedure of root systems emerges. Most existing methods focus only on semi-automatic approaches. This does not support the high-throughput capabilities of acquisition systems. Within the Rhizopolis project (Agropolis foundation), we have designed an automatic analysis pipeline to extract root system architecture from images. This pipeline provides i) a model based segmentation method of the scanned image content (Petri plate, seeds, leaves and root pixels), ii) the extraction of a graph representation of the root system, and iii) a novel method to identifying the root axes organization.

5.1.2. Modeling the plant ontogenic programme

Participants: Christophe Godin, Yann Guédon, Evelyne Costes, Jean-Baptiste Durand, Anaëlle Ambreville, Pierre Fernique, Christophe Pradal, Jean Peyhardi, Catherine Trottier, Yassin Refahi, Etienne Farcot.

This research theme is supported by two PhD programmes.

The remarkable organization of plants at macroscopic scales may be used to infer particular aspects of meristem functioning. The fact that plants are made up of the repetition of many similar components at different scales, and the presence of morphological gradients, e.g. [48], [54], [55], [52], provides macroscopic evidence for the existence of regularities and identities in processes that drive meristem activity at microscopic scales. Different concepts have been proposed to explain these specific organizations such as "morphogenetic programme" [57], "age state" [51] or "physiological age" [50]. All these concepts state that meristem fate changes according to position within the plant structure and during its development. Even though these changes in meristem fate are specific to each species and lead to the differentiation of axes, general rules can be highlighted [51], [50]. Here we develop computational methods to decipher these rules.

- *Relating branching structure to the shoot properties* (Jean Peyhardi, Yann Guédon, Evelyne Coste, Catherine Trottier, Yves Caraglio [AMAP], Pierre-Eric Lauri [AGAP, AFEF team])

  Shoot branching structures often take the form of a succession of homogeneous branching zones and have been analyzed using segmentation models such as hidden semi-Markov chains. Axillary meristem fates are influenced by local properties of the parent shoot such as for instance its growth rate or local curvature. The objective of this work, which is part of the PhD subject of Jean Peyhardi, is to develop statistical models that generalize hidden semi-Markov chains with the capability to incorporate explanatory variables that vary along the parent shoot (e.g. maximum growth rate of the leaf, surface of the leaf, length of the internode, local curvature of the parent shoot). More precisely, the simple multinomial distributions that represent the axillary productions observed in the different branching zones are replaced by multinomial generalized linear models (GLMs). Since the two classical categories of multinomial GLMs that correspond either to nominal or ordinal categorical response variables were not appropriate, we chose to develop a new family of multinomial GLMs called multi-step multinomial GLMs that enable to tackle partially ordered categorical response variables. Typically, we need to distinguish different timing of branching (e.g. immediate shoot, one-year-delayed shoot and latent bud), different categories of offspring shoots (e.g. among one-year-delayed shoots, vegetative short shoot, vegetative long shoot and flowering shoot) and to specialize the explanatory variables for certain categories of offspring shoots (e.g. the growth of the parent shoot influence the immediate offspring shoots but not the one-year-delayed offspring shoots). The resulting integrative models are called semi-Markov switching multi-step multinomial GLMs and are applied to different data sets corresponding mainly to fruit tree branching structures.

- *Genetic determinisms of the alternation of flowering in apple tree progenies.* (Jean-Baptiste Durand, Jean Peyhardi, Baptiste Guitton [AGAP, AFEF team], Yan Holtz [AGAP, AFEF team] Catherine Trottier, Evelyne Costes, Yann Guédon)

  A first study was performed to characterize genetic determinisms of the alternation of flowering in apple tree progenies [34], [17]. Data were collected at two scales: at whole tree scale (with annual time step) and a local scale (annual shoot or AS, which is the portions of stem that were grown during the same year). Two replications of each genotype were available.

  Indices were proposed for early detection of alternation during the juvenile phase. They were based on a trend model and a quantification of the deviation amplitudes and dependency, with respect to the trend. This allows to quantify alternation from the yearly numbers of inflorescences at tree scale.

  However, phenotyping subsamples of AS sequences is more realistic in the framework of genotype selection. To model alternation of flowering at AS scale, a second-order Markov tree model was built. Its transition probabilities were modeled as generalized linear mixed models, to incorporate the effects of genotypes, year and memory of flowering for the Markovian part, with interactions between these components. Asynchronism of flowering at AS scale was also assessed using an entropy-based criterion.

  This work started during the PhD's work of Baptiste Guitton. It was then extended in 2012 by Yan Holtz during this Master 2 internship, supervised by Evelyne Costes and Jean-Baptiste Durand. New progenies were considered, as well as the performance of approximating the descriptors at whole tree scale with those at AS scale. These descriptors allowed the identification of QTL zones involved in the control of flowering in apple trees.

  As a perspective of this work, patterns in the production of children ASs (numbers of flowering and vegetative children) depending on the type of the parent AS must be analyzed using branching processes and different types of Markov trees, in the context of Pierre Fernique's PhD Thesis (see next item in Section 5.1.2).

- *Modeling branching patterns in fruit tree shoots through the characterization of their demographic properties* (Pierre Fernique, Jean-Baptiste Durand, Yann Guédon).

  To test the effect of some properties of a given parent shoot on the properties of its children shoots, statistical models based on multitype branching processes were developed. This kind of dependence between parent and children shoots is frequently at stake in fruit trees, for which the number of flowering or vegetative children of a parent shoot depends on its nature, with potential interactions with other factors. Thus, controlling demographic patterns of the shoots (through varietal selection or crop management strategies) is expected to bring substantial improvements in the quantity and quality of yields.

  Formally, the shoot properties are summed up using the notion of shoot state. The number of children shoots in each state is modeled through discrete multivariate distributions. Model selection procedures are necessary to specify parsimonious distributions. We developed an approach based on probabilistic graphical models to identify and exploit properties of conditional independence between numbers of children in different states, so as to simplify the specification of their joint distribution. The graph building stage was based on exploring the space of possible chain graph models, which required defining a notion of neighbourhood of these graphs. A parametric distribution was associated with each graph. It was obtained by combining families of univariate and multivariate distributions or regression models, and their mixtures. These families were chosen by selection model procedures among different parametric families [45], [32].

This work was carried out in the context of Pierre Fernique's first year of PhD (Montpellier 2 University and CIRAD). It was applied to model dependencies between short or long, vegetative or flowering shoots in apple trees. The results highlighted contrasted patterns related to the parent shoot state, with interpretation in terms of alternation of flowering (see previous item in Section 5.1.2). It was also applied to the analysis of the connections between cyclic growth and flowering of mango trees (see next item in Section 5.1.2). This work will be continued during Pierre Fernique's PhD thesis, with extensions to other fruit tree species and other parametric discrete multivariate families of distributions, including covariates and mixed effects.

- *Analyzing fruit tree phenology* (Anaëlle Dambreville, Jean-Baptiste Durand, Pierre Fernique, Yann Guédon, Christophe Pradal, Pierre-Eric Lauri [AFEF team, AGAP], Frédéric Normand, Catherine Trottier)

  Mango is a tropical tree characterized by strong asynchronisms within and between trees. Causation networks explaining the vegetative and reproductive growths within and between growing cycles were studied on the basis of generalized linear models [16]. We highlighted in this way marked interplays between structural and temporal components of tree structure development at three scales. At growth unit scale, a growth unit which appeared early in the growing cycle had higher rate of burst compared to late appeared growth units. At growing cycle scale, a flowering growth unit delayed its future vegetative growth compared to a vegetative growth unit. At tree scale, a fruiting tree delayed further vegetative growth and flowering compared to a non-fruiting tree. These results evidenced that tree phenology is strongly affected by structural components and not only by the environment.

  To study more precisely the interplays between all these structural components, we started during the internship of A. Jestin to build an integrative model to simulate the development of mango tree based on the L-system formalism and GLM to model the dependencies between events.

  We are also investigating jointly structure development and phenology of mango, and characterizing the specific spatio-temporal patterns leading to patches of vegetative or flowering growth units. Our approach is based on statistical models for trees; particularly hidden Markov tree models and multitype branching processes [32].

- *Integrative developmental growth stages of shoots* (Anaëlle Dambreville, Yann Guédon, Pierre-Eric Lauri [AFEF team, AGAP], Frédéric Normand)

  Growth and development are often studied as two separated processes. Our aim is to investigate the coordination between growth and development in mango shoots. We considered three types of organ, namely the shoot axis, its attached leaves and the inflorescence. Two types of data were collected during the shoot and inflorescence follow-up: developmental stages determined in an expert way and organs sizes determined from measurements. To give an integrative view of the shoot and inflorescence growth and development, we adopted the following strategy. For a given cultivar, we first built a multi-state model on the basis of absolute growth rate sequences deduced from the measurements. Using these models, we computed growth stages. These growth stages highlighted growth asynchronisms between two topologically-connected organs: the axis and its leaves. Then, we compared these growth stages with the developmental ones and we obtained strong matches between them. The integrated developmental growth stages emphasized that the developmental stages are markedly related to growth rates an can be interpreted in terms of physiological (hydraulics, carbohydrates partitioning) and developmental (organs preformation versus neoformation) processes.

- *Self-nested structure of plants.* (Christophe Godin, Romain Azais, Farah Ben Naoum, Jean-Baptiste Durand, Alain Jean-Marie)

  In a previous work [5], we designed a method to compress tree structures and to quantify their degree of self-nestedness. This method is based on the detection of isomorphic subtrees in a given tree and on the construction of a DAG, equivalent to the original tree, where a given subtree class is represented only once (compression is based on the suppression of structural redundancies in the original tree). In the compressed graph, every node representing a particular subtree in the original tree has exactly the same height as its corresponding node in the original tree. This method thus compresses a tree in width, but not in height. In this new work, we designed an extension of this compression method in which a tree is compressed in both width and height. The method is based on the detection of so-called *quasi-isomorphic paths* in a tree and on the compression of these paths in height. A paper describing the corresponding algorithms is being written.

We are currently studying how to generalize the previous approach using stochastic trees. The idea is to generalize the previously defined equivalence relation so that equivalent trees are identical in distribution now and are no longer strictly isomorphic. Algorithms to estimate the different distributions from tree samples rely on combinatorics arguments that are used to estimate the probability that an observed tree is produced by a particular stochastic tree model for which a conjecture has been proposed (still to be demonstrated). The asymptotic analysis of similar compression rates on ordered tree-graphs has been carried out by Flajolet et al. (1990) for different types of distributions (uniform, multi-type branching processes). The work is developed in the context of the Post-doc of Romain Azais.

5.1.3. Analyzing the influence of the environment on the plant ontogenic programme

Participants: Frédéric Boudon, Jean-Baptiste Durand, Christophe Godin, Yann Guédon, Jean Peyhardi, Pierre Fernique, Maryline Lièvre, Christine Granier, Evelyne Costes, Pascal Ferraro, Catherine Trottier.

This research theme is supported by three PhD programs.

The ontogenetic programme of a plant is actually sensitive to environmental changes. If, in particular cases, we can make the assumption that the environment is a fixed control variable (see section 5.1.2), in general the structure produced by meristem results from a tight interaction between the plant and its environment, throughout its lifetime. Based on observations, we thus aim to trace back to the different components of the growth (ontogenetic development and its modulation by the environment). This is made using two types of approaches. On the one hand, we develop a statistical approach in which stochastic models are augmented with additional time-varying explanatory variables that represent the environment variations. The design of estimation procedures for these models make it possible to separate the plant ontogenetic programme from its modulation by the environment. On the other hand, we build reactive models that make it possible to simulate in a mechanistic way the interaction between the plant development and its environment.

- *Influence of environment conditions and horticultural practices on the branching and axillary flowering structures of fruit tree shoots.* (Yann Guédon, Evelyne Costes [AFEF Team, AGAP], Ted DeJong [UC Davis], Claudia Negron [UC Davis]).

  In the context of a collaboration with Claudia Negron and Ted DeJong, we studied the influence of water availability and pruning practices on the branching and axillary flowering structures of different categories of almond shoots Stochastic models (hidden semi-Markov chains) were built for the branching and axillary flowering structures of different categories of almond shoots corresponding to different genetic backgrounds, levels of irrigation and pruning practices.

- *Analyzing growth components in trees.* (Yann Guédon, Yves Caraglio [AMAP], Olivier Taugourdeau [AMAP])

  In a forest ecology context, we identified robust indicators that summarize the balance between tree ontogeny and environmental constraints (mainly related to light environment). In this context, tree growth data typically correspond to the retrospective measurement of annual shoot characteristics (e.g. length, number of branches) along the main stem. We applied segmentation models (hidden Markov and semi-Markov chains) that enable to identify tree growth phases. This statistical modeling approach was applied to both deciduous (sessile oak and Persian walnut) and evergreen (Corsican pine and silver fir) tree species growing in contrasted conditions ranging from managed forest stands to unmanaged understoreys. The growth phase duration distributions estimated within these segmentation models characterize the balance between tree ontogeny and the environmental constraints in tree development at the population scale. These distributions had very contrasted characteristics in terms of shape and relative dispersion between ontogeny-driven and environment-driven tree development. The characteristics of growth phase duration distributions may change over tree life reflecting changes in tree competition.

- *Investigating how architectural development interfer with epidemics and epidemic control* (Christian Fournier, Corinne Robert [EGC], Guillaume Garin [ITK], Bruno Andrieu [EGC], Christophe Pradal)

  Recent considerations towards sustainable agriculture require identifying new natural strategies of crop protection. In this perspective, a better identification of major interactions inside pathosystems between the plants, the pathogens and their environment is crucial. These multiscale biological systems are complex: multiple relationships stand out with various dynamics and at various locations in the canopy, related to its architectural development. The purpose of this research is to provide a framework to study the influence of architectural development on pathosystems with modeling. A first generic framework was designed and implemented in the platform OpenAlea [36]. It allows implementing pathogens of different kind using the same concepts, and a re-use of plant models available in Openalea, thus simplifiying the development of pathosystem models based on 3D plants models. A second action was to develop a modular integrated model coupling architectural canopy development, disease dynamics, pesticide application, pesticide decay and effect of pesticide on disease dynamics [35]. This model is currently being assessed against data for validation, and aim at designing new strategies that reduce pesticide applications by increasing natural resistance linked to canopy architecture.

5.2. Meristem functioning and development

In axis 2 work focuses on the creation of a *virtual meristem*, at cell resolution, able to integrate the recent results in developmental biology and to simulate the feedback loops between physiology and growth. The approach is subdivided into several sub-areas of research.

5.2.1. Data acquisition and design of meristem models

Participants: Frédéric Boudon, Christophe Godin, Christophe Pradal, Léo Guignard, Vincent Mirabet [RDP, ENS], Jan Traas, Grégoire Malandain, Jean-Luc Verdeil [PHIV, AGAP].

This research theme is supported by the iSam and Morphogenetics projects.

- *Improvement of the MARS-ALT pipeline robustness* Meristem, laser microscopy, image reconstruction, cell segmentation, automatic lineaging
  Participants: Léo Guignard, Christophe Godin, Christophe Pradal, Grégoire Malandain, Guillaume Baty, Jan Traas, Patrick Lemaire, Pradeep Das [RDP, ENS], Yassin Refahi [RDP, ENS].

  The MARS-ALT (Multi-Angles Registration and Segmentation - Automatic Lineage Tracking) software pipeline automatically performs a segmentation at cell resolution from 3D or 2D voxel images where the membranes/walls are marked (by a die for example) and makes it possible to follow the lineage of these cells through time [4]. A new version of this pipeline is currently being developed. MARS-ALT Version 2 is based on the same algorithms and methods and is intended to improve the overall robustness of the pipeline (protocol, noise in the input image) and automate completely the process. To test the new pipeline, we use different acquisition protocols and different organisms (floral and apical meristems and the early stages of development of a marine animal *Phallusia mammillata*). The segmentation is corrected a posteriori to deal with imaging artifacts due to uncertainties of acquisition. The image data set on which we develop the methods consists of :

  – *Arabidopsis thaliana* shoot apical meristem and primordia with around 6000 cells. The organ is captured from three different angles every 4 hours during 2 or 3 days with a confocal microscope (Collaboration Sainsbury lab, Cambridge)

  – *Arabidopsis thaliana* flower meristems with around 2000 cells. The organ is also captured from three different angles with a confocal microscope (Collaboration RDP Lyon and Sainsbury lab)

– *Phallusia mammillata* and *Ciona intestinalis* embryos with from 32 cells to around 1000 cells. The organism is captured from four different angles every minute during 2 to 3 hours with a SPIM (Single Plane Illumination Microscope) (Collaboration CRBM Montpellier / EMBL Heidelberg). This work is developed in the context of the PhD work of Léo Guignard.

The pipeline provides as an output segmented images on which metrics for each cells can be extracted such as volume, principal components, convex hull and more. A new non-linear registration algorithm developed by G. Malandain (MORPHEME team, Inria Sophia-Antipolis) is now available and will lead to an improvement of ALT algorithm. Redesign and improvement of the lineage tracking pipeline will be the next step.

../../../../projets/virtual_plants/IMG/Arabes_fused_seg.png

Figure 3. Superimposition of an automatic cell segmentation of an arabidopsis flower meristem using the new MARS pipeline with the original confocal image stack where the membranes are marked.

- *Design of 3D virtual atlases for specifying gene expression patterns* (Yassin Refahi, Christophe Godin, Jan Traas, Patrick Lemaire, Grégoire Malandain, Françoise Monéger [RDP, ENS])

  This research theme is supported the ANR GeneShape and iSam projects.

  To organize the various genetic, physiological, physical, temporal and positional informations, we build a spatialized and dynamic database. This database makes it possible to store all the collected information on a virtual 3D structure representing a typical organ. Each piece of information has to be located spatially and temporally in the database. Tools to visually retrieve and manipulate the information, quantitatively through space and time are being developed. For this, the 3D structure of a typical organ has been created at the different stages of development of the flower bud. This virtual structure contains spatial and temporal information on mean cell numbers, cell size, cell lineages, possible cell polarization (transporters, microtubules), and gene expression patterns. Such 3D virtual atlas is mainly descriptive. However, like for classical databases, specific tools make it possible to explore the virtual atlas according to main index keys, in particular spatial and temporal keys. Both a dedicated language and a 3D user interface are being designed to investigate and query the 3D virtual atlas.

A prototype version of the 3D virtual atlas was built 2 years ago [6]. Further developments of this tool will rely on the segmented images produced from microscopy, as presented in the previous section. In particular, a common underlying data structure has to be developed transversally to these two scientific developments. The definition of this data structure has been initiated last year through several team meetings, and should lead to a revised implementation next year.

5.2.2. Shape analysis of meristems

(Jonathan Legrand, Frédéric Boudon, Christophe Godin, Yann Guedon, Pradeep Das [ENS Lyon])
At cellular resolution, we studied the organization of cells in the meristems. The MARS-ALT pipeline provides rich spatio-temporal data sets for analyzing the development of meristems. A first step consisted of designing a dedicated graph for efficiently representing the spatial (adjacency between cells) and temporal (cell division) relationships between cells. Various variables can be attached either to the vertices (e.g. cell volume, inertia axes) or the edges (e.g. wall surface, distance between cell centroids). This graph may be augmented by new variables resulting from various spatial or temporal filtering (e.g. cell volumetric growth). We are now designing models and algorithms for finding patterns in time courses of meristems. In particular, we are investigating spectral clustering methods in order to define homogeneous regions in terms of cell identities in the context of the PhD Work of Jonathan Legrand.

5.2.3. Transport models

Participants: Michael Walker, Christophe Godin, Etienne Farcot, Jan Traas, Yuan Yuan [University of Newfoundland, Canada].

This research theme is supported by the ANR GeneShape and ERASysBio+ iSAM projects and Morphogenetics.

Active transport of the plant hormone auxin has been shown to play a key role in the initiation of organs at the shoot apex, and vein formation in both leaves and the shoot apical meristem. Polar localized membrane proteins of the PIN1 and AUX/LAX family facilitate this transport and observations and models suggest that the coherent organization of these proteins in the L1 layer is responsible for the creation of auxin maxima (surrounded by a depletion zone), which in turn triggers organ initiation close to the meristem center [58] [1]. Furthermore, canalized PIN allocations are thought to play a crucial role in vein formation in the leaf and in the L2. Previous studies have typically modeled the L1 and L2 with different models to explain different patterns of PIN allocations. In the last two years, we developed a unifying model showing that a unique flux-based model could be sufficient to explain PIN patterns in both L1 and L2 [27]. Contrary to our previous study [9], here no change in the model parameters is needed for this. Our approach is based on inherent topological and geometrical differences between the L1 and L2, specifically their dimensionality and the distribution of sources and sinks.

In a different perspective, another study on auxin transport models have been submitted this year. In this work, a generic, adimensional flux-based model of auxin transport was studied using a combination of analytic and numeric approach. The steady-states with uniform auxin distribution were characterised for arbitrary tissues, and some of their bifurcations (loss of stability and Hopf) were described [18]. This work, initiated during an "Explorateur" project funded by Inria during the period October 2012-January 2013, results from the collaboration between E. Farcot and Y. Yuan (Memorial University of Newfoundland, Canada).

5.2.4. Mechanical model

Participants: Olivier Ali, Christophe Godin, Benjamin Gilles, Frédéric Boudon, Jan Traas, Olivier Hamant [ENS-Lyon], Arezki Boudaoud [ENS-Lyon], Jérôme Chopard [University of Western Australia, Perth].

This research theme is supported by the ANR VirtualFlower and Geneshape projects together with the Inria project Morphogenetics and the ERC from Jan Traas.

The rigid cell walls that surround plant cells are responsible for their shape. These structures are under constraint due to turgor pressure inside the cell. To study the changes of shape in plant tissues during organogenesis, we need a mechanical model of tissue development at cellular resolution. We developed such a model, in which walls are characterized by their mechanical properties like the Young modulus which describes the elasticity of the material. Wall deformation results from forces due to turgor pressure. Growth results from cell wall synthesis that is triggered when wall deformation exceeds a particular threshold. The final shape of the tissue integrates mechanically all the local deformations of each cell.

To model this process, we used a tensorial approach to describe both tissue deformation and stresses. Deformations were decomposed into elementary transformations that can be related to underlying biological processes. However, we showed that the observed deformations does not map directly local growth instructions given by genes and physiology in each cell. Instead, the growth is a two-stage process where genes are specifying how cell walls should yield to mechanical stresses. In this way, different regions in the tissue with different cell identities can have different growth properties. The final shape of the tissue results from the integration of all these mechanical properties and stresses at organ level under the growth force due to turgor pressure at tissue scale.

A paper describing the mechanical model and its application to model primorium formation in the shoot apical meristem has been submitted to PNAS in December. Additionally, a redesign of our mechanical model using the SOFA framework is being finalized.

5.2.5. Gene regulatory networks

Modeling gene activities within cells is of primary importance since cell identities correspond to stable combination of gene expression [25].

- *Complex dynamics and spatial interactions in gene networks* (Yassin Refahi, Etienne Farcot, Christophe Godin)

  Complex computational and mathematical questions arise in the study of gene networks at two levels: (i) the single cell level, due to complex, nonlinear interactions, (ii) the tissue level, where multiple cells interact through molecular signals and growth, so that even simple local rules can challenge our intuition at higher scales.
  At the single cell level, new results were obtained in the framework of piecewise-linear models, it is in general very difficult to entirely characterize the attractors of a given system. In an attempt to improve our ability on this question, a probabilistic approach has been proposed in [14], in which it is shown that a Markov chain can built as an approximation of a given piecewise-linear system, and actually used to make predictions about its periodic attractors.

  At a higher scale, we have also continued the study of gene regulation in meristematic tissues. In the context of Y. Refahi's post-doc between Virtual Plants and the group of Henrik Jönsson in Cambridge (Sainsbury Laboratory), we have continued a work that was initiated in Y. Refahi's thesis. This work is motivated by recent biological results, indicating that gradient-like patterns originating from the external layers of meristems may play a decisive role in the specification of the pool of stem cells in a central position. Using the methods in [4], and their on-going improvements, we have acquired new 3D and 4D images that were then segmented. These structures will be used in the next few months to investigate generic patterning properties of gradient like morphogen patterns. This will require a thorough analysis of free diffusion in realistic geometries, as made possible by the newly acquired images. As a preliminary work, we are also currently investigating the formation of gradient patterns in idealised tissues, allowing for deeper analytic treatment than the complex structures obtained by microscopy.

5.2.6. *Model integration*

Participants: Frédéric Boudon, Christophe Godin, Eugenio Azpeitia, Laurent Laplaze, Jan Traas, François Parcy.

This research theme is supported by the ANR/BBSRC project iSam.

Our approach consists of building a programmable tissue which is able to accept different modeling components. This includes a central data structure representing the tissue in either 2-D or 3-D, which is able to grow in time, models of gene activity and regulation, models of signal exchange (physical and chemical) between cells and models of cell cycle (which includes cell division). For each modeling component, one or several approaches are investigated in depth, possibly at different temporal and spatial scales, using the data available from the partners (imaging, gene networks, and expression patterns). Approaches are compared and assessed on the same data. The objective of each submodel component will be to provide plugin components, corresponding to simplified versions of their models if necessary, that can be injected in the programmable tissue platform. This work is developed in collaboration with the RDP group at ENS-Lyon [56] and the CPIB group in Nottingham, UK [49].

- *Development of a computer platform for the 'programmable tissue'.* (Frédéric Boudon, Christophe Godin)

  One key aspect of our approach is the development of a computer platform dedicated to programming virtual tissue development. This platform will be used to carry out integration of the different models developed in this research axis. The platform is based on *OpenAlea*. Partner models can be integrated in the platform in a non-intrusive way (the code of their model need not be rewritten). In this context, model integration will i) consist of designing adequate data-structures at different levels that will be exchanged and reused among the different plug-in models and ii) defining control flows at adequate levels to avoid the burden of excessive interaction between components. In the past year, progress has been made in defining a generic tissue data structure that could be used in this platform, through several group meetings along the year. A redesign of the structure is in progress.

- *Design of a genetic model of inflorescence development.* (Etienne Farcot, Eugenio Azpeitia, Christophe Godin, François Parcy)

  We studied the regulatory network that control flower development during morphogenesis. To overcome the network complexity and integrate this regulation during ontogenesis, we have developed a first model of the control of floral initiation by genes, and in particular the situation of cauliflower mutants, in which the repeatedly meristem fails in making a complete transition to the flower. This work couples models at different scales, since gene regulation is described by a minimal gene network, which is used as a decision module in an L-system model of the inflorescence architecture. This mixed model has led us to make different hypotheses about gene interactions and hormonal regulation. First predictions about gene actors controling the passage to flower could be verified. A first integrated picture of flower development could be reached in the context of the internship of Eugenio Azpeitia (PhD Student).

5.3. Multi-scale models and analysis: from cells to plant architecture (and back)

5.3.1. *Transport model in roots*

Participants: Mikaël Lucas [IRD], Christophe Pradal, Christophe Godin, Christophe Maurel [BPMP].

This research theme is supported by the ANR project HydroRoot.

A model of Arabidopsis thaliana root hydraulics at the cellular level was developped in the OpenAlea modeling platform. The model relies on the integration throughout root architecture of elementary hydraulic components. Each component integrates local radial and axial water flows. Axial hydraulic conductivity is calculated according to Poiseuille's law, based on local size of xylem vessels. Radial hydraulic conductivity is determined in part by aquaporin activity and was set constant throughout root architecture in the first model versions. In its current state, the model is parameterized using architectural, tissular and physiological data that were experimentally determined in the Aquaporin group at BPMP. The architectural reconstruction of the root system is based on a tridimensional multi-scale tree graph (MTG). The current model is capable of predicting the water flow that is transported by a root system in the standard experimental conditions used in the Aquaporin group. This model was used to perform sensitivity analyses and determine the respective contributions to root hydraulic dynamics of various biological parameters (axial and radial hydraulic conductivites, root architecture). One major finding is that the root hydraulic conductivity (Lpr) computed from the model is highly dependent on root architecture. This is due to the limiting role of axial (xylem) conductance, one feature that had been neglected in previous representations of root water transport. The radial hydraulic conductivity may primarily be limiting in conditions of Lpr inhibition, since its increase from values in control roots has marginal effects on Lpr. A new set of experimental data including root diameter repartitions in wild-type plants, and xylem vessel diameters in mutants with altered xylem morphology (irx3, esk1) will be used to implement the model. Root cell hydraulic conductivities will also be measured in these and aquaporin mutant phenotypes. Our aim is to check whether, based on anatomical and morphological data, the model can properly predict the radial hydraulic conductivity of these genotypes.

5.3.2. Transport in fruits

Participants: Mik Cieslak, Nadia Bertin [Inra, Avignon], Frédéric Boudon, Christophe Godin, Michel Genard [Inra, Avignon], Christophe Goz-Bac [Université Montpellier 2].

This research theme is supported by the Agropolis project Fruit3D.

Understanding the controlling factors of fruit quality development is challenging, because fruit quality results from the interplay between physical and physiological processes that are under the control of genes and the environment. Although process-based models have been used to make significant progress in understanding these factors, they ignored to a large extent the shape and internal structure of the fruit.

To help characterizing effects of fruit shape and internal structure on quality, the creation of a 3D virtual fruit model that integrates fruit structure and function with growth governed by environmental inputs has been investigated. For this, a modeling pipeline has been developed that includes the following steps: creation of a 3D volumetric mesh of the internal fruit structure, including vasculature (see section 3). Based on previous compartment models of fruit physiology developed at Avignon, we have then developed models of water and carbon transport that have been coupled with the 3D model of fruit. In the 3D model, different equations are describing the transport between adjacent regions of the fruit represented as a 3D mesh. The integration through space and time is carried out using a standard integration scheme (Runge-Kutta of order 4).

This approach has been applied to study tomato fruit (Solanum lycopersicum) by constructing 3D volumetric meshes from different sources (images of perpendicular fruit slices and MRI data), and integrating water and carbon transport processes into these meshes. To illustrate the tomato model, a simulation of one season of the fruit's growth has been performed and its results compared with an already published process-based tomato fruit model. We first showed that our spatialized model is compliant with classical results of the abstract process-based models but also provides additional information on the internal heterogeneity of the fruit, such as a gradient in sugar concentration. Once the model is calibrated and evaluated, our approach will be suitable for studying the effects of internal fruit heterogeneity and overall shape on fruit quality development.

5.3.3. Analyzing root growth and branching

Participants: Beatriz Moreno Ortega, Sixtine Passot, Yann Guédon, Laurent Laplaze [IRD, DIADE], Mikaël Lucas [IRD, DIADE], Bertrand Muller [INRA, LEPSE].

This research theme is supported by two PhD programmes.

../../../../projets/virtual_plants/IMG/virtualfruit.png

Figure 4. Virtual models of peaches reconstructed from images with simulated vasculatur to simulate carbon and water transport in the fruit

New 2D and 3D root phenotyping plateforms are emerging with associated image analysis toolbox (e.g. SmartRoot). The analysis of complex root phenotyping data is thus a new challenge in developmental biology.

We aim at developing a pipeline of methods for analyzing root systems at three scales:

1. tissular scale to identify and characterize the meristem, elongation and mature zones along a root using piecewise heteroscedastic linear models.
2. individual root scale to analyze the dynamics of root elongation
3. root system scale to analyze the branching structure.

This pipeline of analysis methods will be applied to different species (maize, millet and *arabidopsis*) and for different biological objectives (study of genetic diversity for millet and of metabolic and hormonal controls of morphogenesis for maize).

5.3.4. Analyzing shoot and leaf elongation

Participants: Maryline Lièvre, Yann Guédon, Christine Granier [INRA, LEPSE].

This research theme is supported by one PhD programme.

The analysis of phenotyping data coming from automated platforms such as PHENOPSIS often focuses on the growth of a leaf at a given rank along the stem. We aim at developing a pipeline of methods for analyzing the growth of *arabidopsis* shoot at three scales:

1. tissular scale using a probabilistic model of endoreduplication for modeling the distribution of the leaf epidermis cell surfaces. Endoreduplication, which is a replication of the nuclear genome in the absence of cell division that leads to elevated nuclear gene content, strongly affects the leaf epidermis cells of *arabidopsis*.
2. organ scale using nonlinear regression model for analyzing the growth of each successive leaf.

3. shoot scale: The outputs of the analyses at the tissular and organ scales will be summarized as multivariate sequences along the shoots characterizing each successive leaf. These sequences will be augmented by supplementary morphological variables characterizing leaf shape and properties (e.g. presence/absence of trichomes). These sequences will be globally analyzed in order to take into account plant ontogeny and in particular the successive developmental stages before the floral transition for the wild type and selected mutants of *arabidopsis*.

5.3.5. Analyzing perturbations in Arabidopsis thaliana phyllotaxis

Participants: Christophe Godin, Yann Guédon, Yassin Refahi, Etienne Farcot, Teva Vernoux, Fabrice Besnard [RDP, ENS].

This research theme is supported by iSAM.

The geometric arrangement of lateral organs along plant stems, named phyllotaxis, shows a variety of striking patterns with remarkable regularities and symmetries. This has interested biologists, physicists, mathematicians and computer scientists for decades. These studies have lead to a commonly accepted standard interpretation of phyllotaxis that postulates that organs inhibit the formation of new organs in their vicinity. At a molecular scale, these inhibitory fields have been shown to result from the spatio-temporal distribution of the plant hormone auxin. This model theoretically explains a large part of the diversity of phyllotactic patterns observed in plants.

The cytokinin hormones are known to play a significant role in the regulation of phyllotaxis. Fabrice Besnard and Teva Vernoux realized that *Arabidopsis thalianaahp6* mutants, which are perturbed in the cytokinin signaling pathway, showed unusual chaotic perturbations of the phyllotaxis at macroscopic level.

In order to characterize these perturbations, we designed a pipeline of models and methods which relies of combinatorial and statistical techniques. Using this pipeline of methods, we have shown that the perturbation patterns in both wild-type and mutant plants can be explained by permutations in the order of insertion along the stem of 2 or 3 consecutive organs. The number of successive synchronized organs between two permutations reveals unexpected patterns that depend on the nature of the preceding permutation (2- or 3-permutation). We identified significant individual deviations of the level of baseline segments with reference to 137.5°, which confirms theoretical model predictions. Finally, we highlighted a marked relationship between permutation of organs and defects in the elongation of the internodes in between these organs.

We then looked at the origin of these permutations using confocal microscopy and realized that organs were in fact frequently co-initiated in the mutant, leading after development randomly in half of the cases to permutations. We concluded that the mutant is actually perturbed in the time between consecutive organ initiation (i.e. the plastochrone), while relative angular positions are not affected. After closer inspection, we realized that the mutated gene encode a protein diffusing from the organs and creating a field around the organs that regulates the plastochrone. We could demonstrate that in the mutant, the absence of this field lead to co-initiations and subsequently to the observed permutations.

Altogether, this study sheds a new light on our interpretation of phyllotaxis, revisiting the standard model and suggesting that several fields based on auxin and cytokinin with different properties are required to provide robustness to phyllotaxis. An overview of this work has been published in the journal Nature in December online [13]. Methodological developments were published more extensively in [20])

VISAGES Project-Team

6. New Results

6.1. Image Computing: Detection, Segmentation, Registration and Analysis

6.1.1. A Mathematical Framework for the Registration and Analysis of Multi-Fascicle Models for Population Studies of the Brain Microstructure

Participant: Olivier Commowick.

Diffusion tensor imaging (DTI) is unable to represent the diffusion signal arising from multiple crossing fascicles and freely diffusing water molecules. Generative models of the diffusion signal, such as multi-fascicle models, overcome this limitation by providing a parametric representation for the signal contribution of each population of water molecules. These models are of great interest in population studies to characterize and compare the brain microstructural properties. Central to population studies is the construction of an atlas and the registration of all subjects to it. However, the appropriate definition of registration and atlasing methods for multi-fascicle models have proven challenging. This paper proposes [32] a mathematical framework to register and analyze multi-fascicle models. Specifically, we define novel operators to achieve interpolation, smoothing and averaging of multi-fascicle models. We also define a novel similarity metric to spatially align multi-fascicle models. Our framework enables simultaneous comparisons of different microstructural properties that are confounded in conventional DTI. The framework is validated on multi-fascicle models from 24 healthy subjects and 38 patients with tuberous sclerosis complex, 10 of whom have autism. We demonstrate the use of the multi-fascicle models registration and analysis framework in a population study of autism spectrum disorder.

6.1.2. Multimodal rigid-body registration of 3D brain images using bilateral symmetry

Participants: Sylvain Prima, Olivier Commowick.

In this paper we show how to use the approximate bilateral symmetry of the brain with respect to its interhemispheric fissure for intra-subject (rigid-body) mono- and multimodal 3D image registration. We propose to define and compute an approximate symmetry plane in the two images to register and to use these two planes as constraints in the registration problem. This 6-parameter problem is thus turned into three successive 3-parameter problems. Our hope is that the lower dimension of the parameter space makes these three subproblems easier and faster to solve than the initial one. We implement two algorithms to solve these three subproblems in the exact same way, within a common intensity-based framework using mutual information as the similarity measure. We compare this symmetry-based strategy with the standard approach (i.e. direct estimation of a 6-parameter rigid-body transformation), also implemented within the same framework, using synthetic and real datasets. We show in [44] our symmetry-based method to achieve subvoxel accuracy with better robustness and larger capture range than the standard approach, while being slightly less accurate and slower. Our method also succeeds in registering clinical MR and PET images with a much better accuracy than the standard approach. Finally, we propose a third strategy to decrease the run time of the symmetry-based approach and we give some ideas, to be tested in future works, on how to improve its accuracy.

6.1.3. Distortion Correction in EPI Diffusion Weighted Images

Participants: Renaud Hedouin, Olivier Commowick.

We have compared and developed several methods which correct distortion of EPI images. The most popular method field map do not give optimal results. We have implemented and improved a method based on reversed phase encoding gradient which give good results. To correct diffusion weighted images this method only need one reversed phase encoding gradient $B0$ image which not need substantial additional acquisition time.

6.1.4. Using bilateral symmetry to improve non-local means denoising of MR brain images

Participants: Sylvain Prima, Olivier Commowick.

The popular NL-means denoising algorithm proposes to modify the intensity of each voxel of an image by a weighted sum of the intensities of similar voxels. The success of the NL-means rests on the fact that there are typically enough such similar voxels in natural, and even medical images; in other words, that there is some self-similarity/redundancy in such images. However, similarity between voxels (or rather, between patches around them) is usually only assessed in a spatial neighbourhood of the voxel under study. As the human brain exhibits approximate bilateral symmetry, one could wonder whether a voxel in a brain image could be more accurately denoised using information from both ipsi- and contralateral hemispheres. This is the idea we have investigated in this paper [45]. We define and compute a mid-sagittal plane which best superposes the brain with itself when mirrored about the plane. Then we use this plane to double the size of the neighbourhoods and hopefully find additional interesting voxels to be included in the weighted sum. We evaluate this strategy using an extensive set of experiments on both simulated and real datasets.

6.1.5. Detection of Multiple Sclerosis Lesions using Dictionary Learning

Participants: Hrishikesh Deshpande, Pierre Maurel, Christian Barillot.

Multiple sclerosis (MS) is a chronic, autoimmune, inflammatory disease of the central nervous system, in which certain areas of brain develop MS lesions, which are characterized by demyelination. Over the last years, various models combined with supervised and unsupervised classification methods have been proposed for detection of MS lesions using magnetic resonance images. Recently, signal modeling using sparse representations (SR) has gained tremendous attention and is an area of active research. SR allows coding data as sparse linear combinations of the elements of over-complete dictionary and has led to interesting image recognition results. The dictionary used for sparse coding plays a key role in the classification process. In this work, we have proposed to learn class specific dictionaries and develop new classification scheme, to automatically detect MS lesions in 3-D multi-channel magnetic resonance images.

6.1.6. Multiple Sclerosis Lesion Detection in Clinically Isolated Syndromes

Participants: Yogesh Karpate, Olivier Commowick, Christian Barillot.

Quantitative assessment of Multiple Sclerosis Lesions (MSL) in Clinically Isolated Syndromes (CIS) is important, as they are a precursor to subsequent stages of the disease. We address the problem of lesion patch detection with respect to Normally Appearing Brain Tissues (NABT). Our approach consists in learning rotationally invariant MSL and NABT multimodal intensity signatures based on 3D spherical gabor descriptors. This learning step, done once and for all, is followed by a testing step for the patient patches with an exemplar SVM. First, we develop a framework for selecting focused region of interest (fROI) using linear SVM for scoring. This allows an excellent trade-off between speed and accuracy. Second, building rotational invariant and scale independent features for accurate representation of image signatures. The extracted features are sensitive to the orientation of the analyzed image. This is a drawback in classification and retrieval applications. We handle this problem by using shperical Gabor descriptors. And last, we apply max pooling for down sampling of feature vectors. For the classification purpose we use a standard linear Support Vector Machine(SVM). The main contribution of the work is to build binary classifier to discriminate NABTs and MSLs based upon robust image representation. We have validated our approach on synthetic and real patient data. The synthetic lesion data is generated with noise, without noise and with bias field. Further, validation is carried out in three different scenarios. First, we evaluate our classifier using K-fold started with cross validation using NABT from healthy volunteers and MSL from CIS patients, then the detection of NABT and MSL from CIS patients on known patches is performed. The last evaluation concerned the full search algorithm.

6.1.7. Intensity Normalization in Longitudinal MS Patients

Participants: Yogesh Karpate, Olivier Commowick, Christian Barillot.

This work proposes a longitudinal intensity normalization algorithm for multi-channel MRI of brain of MS patient in the presence of lesions, aiming towards stable and consistent longitudinal segmentation. This approach is parametric and developed using two different forms of Robust Expectation Maximization (EM). The first is Spatio-Temporal Robust Expected Maximization (STREM) and other being EM with beta divergence. We validated our method on real longitudinal multiple sclerosis subjects.

6.2. Image processing on Diffusion Weighted Magnetic Resonance Imaging

6.2.1. Statistical Analysis of White Matter Integrity for the Clinical Study of Specific Language Impairment in Children

Participants: Olivier Commowick, Camille Maumet, Aymeric Stamm, Jean-Christophe Ferré, Christian Barillot.

Children affected by Specific Language Impairment (SLI) fail to develop a normal language capability. To date, the etiology of SLI remains largely unknown. It induces difficulties with oral language which cannot be directly attributed to intellectual deficit or other developmental delay. Whereas previous studies on SLI focused on the psychological and genetic aspects of the pathology, few imaging studies investigated defaults in neuroanatomy or brain function. We have proposed [53] to investigate the integrity of white matter in SLI thanks to diffusion Magnetic Res- onance Imaging. An exploratory analysis was performed without a priori on the impaired regions. A region of interest statistical analysis was performed based, first, on regions defined from Catani's atlas and, then, on tractography-based regions. Both the mean fractional anisotropy and mean apparent diffusion coefficient were compared across groups. To the best of our knowledge, this is the first study focusing on white matter integrity in specific language impairment. 22 children with SLI and 19 typically developing children were involved in this study. Overall, the tractography-based approach to group comparison was more sensitive than the classical ROI-based approach. Group differences between controls and SLI patients included decreases in FA in both the perisylvian and ventral pathways of language, comforting findings from previous functional studies.

6.2.2. Adaptive Multi-modal Particle Filtering for Probabilistic White Matter Tractography

Participants: Aymeric Stamm, Olivier Commowick, Christian Barillot.

Particle filtering has recently been introduced to perform probabilistic tractography in conjunction with DTI and Q-Ball models to estimate the diffusion information. Particle filters are particularly well adapted to the tractography problem as they offer a way to approximate a probability distribution over all paths originated from a specified voxel, given the diffusion information. In practice however, they often fail at consistently capturing the multi-modality of the target distribution. For brain white matter tractography, this means that multiple fiber pathways are unlikely to be tracked over extended volumes. We have proposed [51] to remedy this issue by formulating the filtering distribution as an adaptive M-component non-parametric mixture model. Such a formulation preserves all the properties of a classical particle filter while improving multi-modality capture. We apply this multi-modal particle filter to both DTI and Q-Ball models and propose to estimate dynamically the number of modes of the filtering distribution. We show on synthetic and real data how this algorithm outperforms the previous versions proposed in the literature.

6.2.3. Tracking the Cortico-Spinal Tract from Low Spatial and Angular Resolution Diffusion MRI

Participants: Aymeric Stamm, Olivier Commowick, Christian Barillot.

We have participated to the annual MICCAI workshop on DTI tractography [52]. We presented a pipeline to reconstruct the corticospinal tract (CST) that connects the spinal cord to the motor cortex. The proposed method combines a new geometry-based multi-compartment diffusion model coined Diffusion Directions Imaging and a new adaptive multi-modal particle filter for tractography. The DTI Tractography challenge proposes to test our methods in the context of neurosurgical planning of tumor removal, where very low spatial and angular resolution diffusion data is available due to severe acquisition time constraints. We took up the challenge and present our reconstructed CSTs derived from a single-shell acquisition scheme at b = 1000 s/mm2 with only 20 or 30 diffusion gradients (low angular resolution) and with images of 5 mm slice thickness (low spatial resolution).

6.3. Medical Image Computing in Brain Pathologies

6.3.1. Semi-Automatic Classification of Lesion Patterns in Patients with Clinically Isolated Syndrome

Participants: Olivier Commowick, Jean-Christophe Ferré, Gilles Edan, Christian Barillot.

Multiple sclerosis (MS) is neuro-degenerative disease of the Central Nervous System characterized by the loss of myelin. A Clinically Isolated Syndrome (CIS) is a first neurological episode caused by inflammation/demyelination in the central nervous system which may lead to MS. Better understanding of the disease at its onset will lead to a better discovery of pathogenic mechanisms, allowing suitable therapies at an early stage. We have proposed [37] an automatic segmentation algorithm for two different contrast agents, used within a framework for early characterization of CIS patients according to lesion patterns, and more specifically according to the nature of the inflammatory patterns of these lesions. We expect that the proposed framework can infer new prospective figures from the earliest imaging signs of MS since it can provide a classification of different types of lesions across patients. The lesion detection algorithm based on intensity normalization and subtraction of the used MRI data is a pivotal step, since it avoids the time-demanding task of manual delineation.

6.3.2. Multiple Sclerosis Lesions Evolution in Patients with Clinically Isolated Syndrome

Participants: Olivier Commowick, Jean-Christophe Ferré, Gilles Edan, Christian Barillot.

Multiple sclerosis (MS) is a disease with heterogeneous evolution among the patients. Some classifications have been carried out according to either the clinical course or the immunopathological profiles. Epidemiological data and imaging are showing that MS is a two-phase neurodegenerative inflammatory disease. At the early stage it is dominated by focal inflammation of the white matter (WM), and at a latter stage it is dominated by diffuse lesions of the grey matter and spinal cord. A Clinically Isolated Syndrome (CIS) is a first neurological episode caused by inflammation/demyelination in the central nervous system which may lead to MS. Few studies have been carried out so far about this initial stage. Better understanding of the disease at its onset will lead to a better discovery of pathogenic mechanisms, allowing suitable therapies at an early stage. We have proposed [36] a new data processing framework able to provide an early characterization of CIS patients according to lesion patterns, and more specifically according to the nature of the inflammatory patterns of these lesions. The method is based on a two layers classification. Initially, the spatio-temporal lesion patterns are classified using a tensor-like representation. The discovered lesion patterns are then used to identify group of patients and their correlation to 15 months follow-up total lesion loads (TLL), which is so far the only image-based figure that can potentially infer future evolution of the pathology. We expect that the proposed framework can infer new prospective figures from the earliest imaging sign of MS since it can provide a classification of different types of lesion across patients.

6.3.3. Arterial Spin Labeling at 3T in semantic dementia: perfusion abnormalities detection and comparison with FDG-PET

Participants: Isabelle Corouge, Jean-Christophe Ferré, Elise Bannier, Aymeric Stamm, Christian Barillot, Jean-Yves Gauvrit.

Arterial Spin Labeling (ASL) is a non invasive perfusion imaging technique which has shown great diagnosis potential in dementia. However, it has never been applied to semantic dementia (SD), a rare subtype of frontotemporal lobar degeneration characterized by the gradual loss of conceptual knowledge, which is actually explored by a now well established marker of SD: ^{18}fluorodeoxyglucose-positron emission tomography (FDG-PET) imaging. Although ASL and FDG-PET respectively measure perfusion and metabolism, they have been shown to be strongly correlated. In this work, we explore the ability of ASL to detect perfusion abnormalities in SD in comparison with FDG-PET. Using patients and healthy subjects data from an ongoing clinical study, we apply our analysis framework starting with visual comparison of ASL and FDG-PET, and focusing on ASL data preprocessing and statistical analysis at the individual and group level. Preliminary results yield concordant observations between ASL and FDG-PET as well as expected hypoperfusions in SD, namely in the left temporal lobe, thus suggesting the potential of ASL to assess perfusion impairments in SD.

6.4. Vascular Imaging and Arterial Spin Labelling

6.4.1. Patient-specific detection of perfusion abnormalities combining within-subject and between-subject variances in Arterial Spin Labeling.

Participants: Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot.

In this paper, patient-specific perfusion abnormalities in Arterial Spin Labeling (ASL) were identified by comparing a single patient to a group of healthy controls using a mixed-effect hierarchical General Linear Model (GLM). Two approaches are currently in use to solve hierarchical GLMs: (1) the homoscedastic approach assumes homogeneous variances across subjects and (2) the heteroscedastic approach is theoretically more efficient in the presence of heterogeneous variances but algorithmically more demanding. In practice, in functional magnetic resonance imaging studies, the superiority of the heteroscedastic approach is still under debate. Due to the low signal-to-noise ratio of ASL sequences, within-subject variances have a significant impact on the estimated perfusion maps and the heteroscedastic model might be better suited in this context. In this paper we studied how the homoscedastic and heteroscedastic approaches behave in terms of specificity and sensitivity in the detection of patient-specific ASL perfusion abnormalities. Validation was undertaken on a dataset of 25 patients diagnosed with brain tumors and 36 healthy volunteers. We showed evidence of heterogeneous within-subject variances in ASL and pointed out an increased false positive rate of the homoscedastic model. In the detection of patient-specific brain perfusion abnormalities with ASL, modeling heterogeneous variances increases the sensitivity at the same specificity level [24].

6.4.2. An a contrario approach for the detection of activated brain areas in fMRI

Participants: Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot.

BOLD functional MRI (fMRI) is now a widespread imaging technique to study task-related activity in the brain. However, getting the areas of activation at the individual subject level is still an open issue. The standard massively univariate statistical analysis is usually performed after smoothing the data and makes use of a single p-value for final thresholding of the results [1]. In group fMRI studies, the need for compensation of cross-subjects misregistrations clearly justifies the smoothing. However, at the individual level, where neat delineations of the activated areas are of interest, the use of Gaussian smoothing as a pre-processing step is more questionable. In this paper, we propose to study the ability of an a contrario approach, recently adapted for basal perfusion abnormalities detection [2], to correctly detect areas of functional activity [42].

6.4.3. Robust perfusion maps in Arterial Spin Labeling by means of M-estimators

Participants: Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot.

Non-invasive measurement of Cerebral Blood Flow (CBF) is now feasible thanks to the introduction of Arterial Spin Labeling (ASL) Magnetic Resonance Imaging (MRI) techniques. To date, due to the low signal-to-noise ratio of ASL, a single acquisition (pair of control/label scans) is not sufficient to estimate perfusion reliably. Instead, the acquisition is usually repeated several times and the perfusion information is calculated by averaging across the repetitions. However, due to its zero breakdown point, the sample mean is very sensitive to outliers. In this paper, we propose to compute ASL CBF maps using Huber's M-estimator, a robust statistical function that is not overly impacted by outliers. This method is compared to an empirical approach, introduced in [1], based on z-score thresholding [43].

6.4.4. Quantifying CBF from Aterial Spin Labeling via Diverse-TI: sampling diversity or repetitions ?

Participants: Lei Yu, Pierre Maurel, Christian Barillot.

Arterial Spin Labeling (ASL) is a noninvasive perfusion technique which allows the absolute quantification of Cerebral Blood Flow (CBF). The perfusion is obtained from the difference between images with and without magnetic spin labeling of the arterial blood and the captured signal is around 0.5-2% of the magnitude of the labeling images, so the noise is one of the main problems for further data analysis. Classical method, *Mono-TI*, for CBF quantification is averaging repetitions with only one Inversion Time (TI) - the time delay between labeling and acquisition to allow the labeled blood to arrive the imaging slice. It improves the robustness to noise, however, cannot compensate the variety of Arterial Arrival Time (AAT). In this paper, *Diverse-TI* is proposed to exploit different TI sampling instants (sampling diversity) to improve the robustness to variety of AAT and simultaneously average repetitions with each TI (sampling repetitions) to improve the robustness to noise. Generally, the sampling diversity is relatively small and can be considered as compressed measurements, thus the Compressive Matched Filter (CMF) enlightened from sparsity is exploited to directly reconstruct CBF and AAT directly from compressed measurements. Meanwhile, regarding the CBF quantification performance, the compromise between the sampling repetition and sampling diversity is discussed and the empirical protocol to determine the sampling diversity is proposed. Simulations are carried out to highlight our discussions. This is a joint work with Remi Gribonval (Panama Team) [56].

6.4.5. Peripheral angiography and neurovascular imaging

Participants: Hélène Raoult, Jean-Yves Gauvrit, Elise Bannier, Pierre Maurel, Clement Neyton, Christian Barillot, Jean-Christophe Ferré.

Vascular imaging contributions were performed on two different regions during the evaluation period: first on peripheral angiography, then on neurovascular imaging. Arteriography and MR angiography are routinely performed in patients presenting vascular pathologies. Yet, contrast agent injection is contraindicated in patients with renal insufficiency and the underlying risk of developing nephrogenic systemic fibrosis further encourages research on non-contrast enhanced MR angiography techniques (NCE MRA). In this context, we have been working on new MR sequences to reliably detect vascular abnormalities.

A first study [29] was published, where we assessed the feasibility and image quality of an improved non-gated carotid NATIVE TrueFISP NCE MRA sequence providing an extended field of view and a shorter acquisition time as compared to Time-of-Flight (TOF) imaging. A second study [48] was recently accepted for publication in Radiology on intracranial NCE MRA for arteriovenous malformation imaging with a high temporal resolution over 2 cardiac cycles. Combined with image post-processing, it allows improved depiction of venous drainage necessary to evaluate hemorrhagic risk and quantification. This ongoing work was just submitted.

WILLOW Project-Team

6. New Results

6.1. 3D object and scene modeling, analysis, and retrieval

../../../../projets/willow/IMG/Aubry13.jpg

Figure 1. Our system automatically aligns and recovers the viewpoint of paintings, drawings, and historical photographs to a 3D model of an architectural site.

6.1.1. Painting-to-3D Model Alignment Via Discriminative Visual Elements

Participants: Mathieu Aubry, Bryan Russell [Intel Labs], Josef Sivic.

In this work we describe a technique that can reliably align arbitrary 2D depictions of an architectural site, including drawings, paintings and historical photographs, with a 3D model of the site. This is a tremendously difficult task as the appearance and scene structure in the 2D depictions can be very different from the appearance and geometry of the 3D model, e.g., due to the specific rendering style, drawing error, age, lighting or change of seasons. In addition, we face a hard search problem: the number of possible alignments of the painting to a large 3D model, such as a partial reconstruction of a city, is huge. To address these issues, we develop a new compact representation of complex 3D scenes. The 3D model of the scene is represented by a small set of discriminative visual elements that are automatically learnt from rendered views. Similar to object detection, the set of visual elements, as well as the weights of individual features for each element, are learnt in a discriminative fashion. We show that the learnt visual elements are reliably matched in 2D depictions of the scene despite large variations in rendering style (e.g. watercolor, sketch, historical photograph) and structural changes (e.g. missing scene parts, large occluders) of the scene. We demonstrate an application of the proposed approach to automatic re-photography to find an approximate viewpoint of historical paintings and photographs with respect to a 3D model of the site. The proposed alignment procedure is validated via a human user study on a new database of paintings and sketches spanning several sites. The results demonstrate that our algorithm produces significantly better alignments than several baseline methods. This work has been accepted for publication to the ACM Transactions on Graphics (ACM ToG). The problem addressed in this work is illustrated in Figure 1 and example results are shown in figure 2 . The pre-print is available online at [10].

6.1.2. Learning and Calibrating Per-Location Classifiers for Visual Place Recognition

Participants: Petr Gronat, Josef Sivic, Guillaume Obozinski [ENPC / 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. This work has been published at CVPR 2013 [6].

6.1.3. Visual Place Recognition with Repetitive Structures

Participants: Akihiko Torii [Tokyo Institute of Technology], Josef Sivic, Tomáš Pajdla [CTU in Prague], Masatoshi Okutomi [Tokyo Institute of Technology].

Repeated structures such as building facades, fences or road markings often represent a significant challenge for place recognition. Repeated structures are notoriously hard for establishing correspondences using multi-view geometry. Even more importantly, they violate the feature independence assumed in the bag-of-visual-words representation which often leads to over-counting evidence and significant degradation of retrieval performance. In this work we show that repeated structures are not a nuisance but, when appropriately represented, they form an important distinguishing feature for many places. We describe a representation of repeated structures suitable for scalable retrieval. It is based on robust detection of repeated image structures and a simple modification of weights in the bag-of-visual-word model. Place recognition results are shown on datasets of street-level imagery from Pittsburgh and San Francisco demonstrating significant gains in recognition performance compared to the standard bag-of-visual-words baseline and more recently proposed burstiness weighting. This work has been published at CVPR 2013 [9].

../../../../projets/willow/IMG/Aubry13im3.jpg

Figure 2. Example alignments of non-photographic depictions to 3D models. Notice that we are able to align depictions rendered in different styles and having a variety of viewpoints with respect to the 3D models.

6.1.4. Trinocular Geometry Revisited

Participants: Jean Ponce, Martial Hebert [CMU].

When do the visual rays associated with triplets of point correspondences converge, that is, intersect in a common point? Classical models of trinocular geometry based on the fundamental matrices and trifocal tensor associated with the corresponding cameras only provide partial answers to this fundamental question, in large part because of underlying, but seldom explicit, general configuration assumptions. In this project, we use elementary tools from projective line geometry to provide necessary and sufficient geometric and analytical conditions for convergence in terms of transversals to triplets of visual rays, without any such assumptions. In turn, this yields a novel and simple minimal parameterization of trinocular geometry for cameras with non-collinear or collinear pinholes. This work has been submitted to CVPR 2014.

6.2. Category-level object and scene recognition

6.2.1. Learning Graphs to Match

Participants: Minsu Cho, Karteek Alahari, Jean Ponce.

Many tasks in computer vision are formulated as graph matching problems. Despite the NP-hard nature of the problem, fast and accurate approximations have led to significant progress in a wide range of applications. Learning graph models from observed data, however, still remains a challenging issue. This work presents an effective scheme to parameterize a graph model, and learn its structural attributes for visual object matching. For this, we propose a graph representation with histogram-based attributes, and optimize them to increase the matching accuracy. Experimental evaluations on synthetic and real image datasets demonstrate the effectiveness of our approach, and show significant improvement in matching accuracy over graphs with pre-defined structures. The work is illustrated in Figure 3 . This work has been published ICCV 2013 [3].

6.2.2. Finding Matches in a Haystack: A Max-Pooling Strategy for Graph Matching in the Presence of Outliers

Participants: Minsu Cho, Olivier Duchenne [Intel], Jian Sun, Jean Ponce.

A major challenge in real-world matching problems is to tolerate the numerous outliers arising in typical visual tasks. Variations in object appearance, shape, and structure within the same object class make it hard to distinguish inliers from outliers due to clutters. In this project, we propose a novel approach to graph matching, which is not only resilient to deformations but also remarkably tolerant to outliers. By adopting a max-pooling strategy within the graph matching framework, the proposed algorithm evaluates each candidate match using its most promising neighbors, and gradually propagates the corresponding scores to update the neighbors. As final output, it assigns a reliable score to each match together with its supporting neighbors, thus providing contextual information for further verification. We demonstrate the robustness and utility of our method with synthetic and real image experiments. This work has been submitted to CVPR 2014.

6.2.3. Decomposing Bag of Words Histograms

Participants: Ankit Gandhi [IIIT India], Karteek Alahari, C.v. Jawahar [IIIT India].

We aim to decompose a global histogram representation of an image into histograms of its associated objects and regions. This task is formulated as an optimization problem, given a set of linear classifiers, which can effectively discriminate the object categories present in the image. Our decomposition bypasses harder problems associated with accurately localizing and segmenting objects. We evaluate our method on a wide variety of composite histograms, and also compare it with MRF-based solutions. In addition to merely measuring the accuracy of decomposition, we also show the utility of the estimated object and background histograms for the task of image classification on the PASCAL VOC 2007 dataset. This work has been published at ICCV 2013 [5].

6.2.4. Image Retrieval using Textual Cues

Participants: Anand Mishra [IIIT India], Karteek Alahari, C.v. Jawahar [IIIT India].

../../../../projets/willow/IMG/cho13.jpg

Figure 3. Graph learning for matching. Our approach learns a graph model from labeled data to provide the best match to instances of a target class. It shows significant improvement over previous approaches for matching. (Best viewed in color.)

We present an approach for the text-to-image retrieval problem based on textual content present in images. Given the recent developments in understanding text in images, an appealing approach to address this problem is to localize and recognize the text, and then query the database, as in a text retrieval problem. We show that such an approach, despite being based on state-of-the-art methods, is insufficient, and propose a method, where we do not rely on an exact localization and recognition pipeline. We take a query-driven search approach, where we find approximate locations of characters in the text query, and then impose spatial constraints to generate a ranked list of images in the database. The retrieval performance is evaluated on public scene text datasets as well as three large datasets, namely IIIT scene text retrieval, Sports-10K and TV series-1M, we introduce. This work has been published at ICCV 2013 [7].

6.2.5. Learning Discriminative Part Detectors for Image Classification and Cosegmentation

Participants: Jian Sun, Jean Ponce.

In this work, we address the problem of learning discriminative part detectors from image sets with category labels. We propose a novel latent SVM model regularized by group sparsity to learn these part detectors. Starting from a large set of initial parts, the group sparsity regularizer forces the model to jointly select and optimize a set of discriminative part detectors in a max-margin framework. We propose a stochastic version of a proximal algorithm to solve the corresponding optimization problem. We apply the proposed method to image classification and cosegmentation, and quantitative experiments with standard bench- marks show that it matches or improves upon the state of the art. This work has been published at CVPR 2013 [8].

6.2.6. Learning and Transferring Mid-Level Image Representations using Convolutional Neural Networks

Participants: Maxime Oquab, Leon Bottou [MSR New York], Ivan Laptev, Josef Sivic.

Convolutional neural networks (CNN) have recently shown outstanding image classification performance in the large-scale visual recognition challenge (ILSVRC2012). The success of CNNs is attributed to their ability to learn rich mid-level image representations as opposed to hand-designed low-level features used in other image classification methods. Learning CNNs, however, amounts to estimating millions of parameters and requires a very large number of annotated image samples. This property currently prevents application of CNNs to problems with limited training data. In this work we show how image representations learned with CNNs on large-scale annotated datasets can be efficiently transferred to other visual recognition tasks with limited amount of training data. We design a method to reuse layers trained on the ImageNet dataset to compute mid-level image representation for images in the PASCAL VOC dataset. We show that despite differences in image statistics and tasks in the two datasets, the transferred representation leads to significantly improved results for object and action classification, outperforming the current state of the art on Pascal VOC 2007 and 2012 datasets. We also show promising results for object and action localization. The pre-print of this work is available online [11]. Results are illustrated in Figure 4 .

6.2.7. Seeing 3D chairs: exemplar part-based 2D-3D alignment using a large dataset of CAD models

Participants: Mathieu Aubry, Bryan Russell [Intel labs], Alyosha Efros [UC Berkeley], Josef Sivic.

We present an approach for the text-to-image retrieval problem based on textual content present in images. Given the recent developments in understanding text in images, an appealing approach to address this problem is to localize and recognize the text, and then query the database, as in a text retrieval problem. We show that such an approach, despite being based on state-of-the-art methods, is insufficient, and propose a method, where we do not rely on an exact localization and recognition pipeline. We take a query-driven search approach, where we find approximate locations of characters in the text query, and then impose spatial constraints to generate a ranked list of images in the database. The retrieval performance is evaluated on public scene text datasets as well as three large datasets, namely IIIT scene text retrieval, Sports-10K and TV series-1M, we introduce. This work has been submitted to CVPR 2014.

../../../../projets/willow/IMG/oquab13.jpg

Figure 4. Recognition and localization results of our method for a Pascal VOC test image. Output maps are shown for six object categories with the highest responses.

6.3. Image restoration, manipulation and enhancement

6.3.1. 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 published at CVPR 2013 [4] and example results are shown in figure 5 .

6.3.2. Efficient, Blind, Spatially-Variant Deblurring for Shaken Images

Participants: Oliver Whyte [Microsoft Redmond], Josef Sivic, Andrew Zisserman, Jean Ponce.

In this chapter we discuss modeling and removing spatially-variant blur from photographs. We describe a compact global parameterization of camera shake blur, based on the 3D rotation of the camera during the exposure. Our model uses three-parameter homographies to connect camera motion to image motion and, by assigning weights to a set of these homographies, can be seen as a generalization of the standard, spatially-invariant convolutional model of image blur. As such we show how existing algorithms, designed for spatially-invariant deblurring, can be "upgraded" in a straightforward manner to handle spatially-variant blur instead. We demonstrate this with algorithms working on real images, showing results for blind estimation of blur parameters from single images, followed by non-blind image restoration using these parameters. Finally, we introduce an efficient approximation to the global model, which significantly reduces the computational cost of modeling the spatially-variant blur. By approximating the blur as locally-uniform, we can take advantage of fast Fourier-domain convolution and deconvolution, reducing the time required for blind deblurring by an order of magnitude.

This work has been accepted for publication as a book chapter in the upcoming book "Motion Deblurring: Algorithms and Systems" to be published by Cambridge University Press in May 2014. [2] The demo implementing deblurring of images degraded by camera shake is available online at: http://www.di.ens.fr/willow/research/saturation/.

6.4. Human activity capture and classification

6.4.1. Layered Segmentation of People in Stereoscopic Movies

Participants: Karteek Alahari, Guillaume Seguin, Josef Sivic, Ivan Laptev.

In this work we seek to obtain a pixel-wise segmentation and pose estimation of multiple people in a stereoscopic video. This involves challenges such as dealing with unconstrained stereoscopic video, non-stationary cameras, and complex indoor and outdoor dynamic scenes. The contributions of our work are two-fold: First, we develop a segmentation model incorporating person detection, pose estimation, as well as colour, motion, and disparity cues. Our new model explicitly represents depth ordering and occlusion. Second, we introduce a stereoscopic dataset with frames extracted from feature-length movies "StreetDance 3D" and "Pina". The dataset contains 2727 realistic stereo pairs and includes annotation of human poses, person bounding boxes, and pixel-wise segmentations for hundreds of people. The dataset is composed of indoor and outdoor scenes depicting multiple people with frequent occlusions. We demonstrate results on our new challenging dataset, as well as on the H2view dataset from (Sheasby et al. ACCV 2012). This work has been published at ICCV 2013 [1].

[2] http://www.cambridge.org/fr/academic/subjects/engineering/image-processing-and-machine-vision/motion-deblurring-algorithms-and-systems

../../../../projets/willow/IMG/devy13.jpg

Figure 5. Sample deblurring results on real images. From left to right: blurry image, deblurred image, close-up corresponding to the boxes shown in red. Note that our estimated deblurred image has more detail.

6.4.2. Finding Actors and Actions in Movies

Participants: Piotr Bojanowski, Francis Bach [Inria Sierra], Ivan Laptev, Jean Ponce, Cordelia Schmid [Inria Lear], Josef Sivic.

We address the problem of learning a joint model of actors and actions in movies using weak supervision provided by scripts. Specifically, we extract actor/action pairs from the script and use them as constraints in a discriminative clustering framework. The corresponding optimization problem is formulated as a quadratic program under linear constraints. People in video are represented by automatically extracted and tracked faces together with corresponding motion features. First, we apply the proposed framework to the task of learning names of characters in the movie and demonstrate significant improvements over previous methods used for this task. Second, we explore the joint actor/action constraint and show its advantage for weakly supervised action learning. We validate our method in the challenging setting of localizing and recognizing characters and their actions in feature length movies Casablanca and American Beauty. This work has been published at ICCV 2013 [2] and example results are shown in figure 6 . The corresponding software has been also made publicly available (see the software section of this report).

6.4.3. Highly-Efficient Video Features for Action Recognition and Counting

Participants: Vadim Kantorov, Ivan Laptev.

Local video features provide state-of-the-art performance for action recognition. While the accuracy of action recognition has been 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 2014.

../../../../projets/willow/IMG/bojanowski13im3.jpg

Figure 6. Examples of automatically assigned names and actions in the movie Casablanca. Top row: Correct name and action assignments for tracks that have an actor/action constraint in the script. Bottom row: Correct name and action assignments for tracks that do not have a corresponding constraint in the script, but are still correctly classified. Note that even very infrequent characters are correctly classified (Annina and Yvonne). See more examples on the project web-page: http://www.di.ens.fr/willow/research/actoraction/

WIMMICS Project-Team

6. New Results

6.1. QAKiS: Question Answering wiKiframework-based System

Participants: Elena Cabrio, Julien Cojan, Amine Hallili, Serena Villata.

We worked on an extension of QAKiS, the system for open domain Question Answering over Linked Data, that allows to query DBpedia multilingual chapters. Such chapters can contain different information with respect to the English version, e.g. they provide more specificity on certain topics, or fill information gaps. QAKiS exploits the alignment between properties carried out by DBpedia contributors as a mapping from Wikipedia terms to a common ontology, to exploit information coming from DBpedia multilingual chapters (English, French and German), broadening therefore its coverage. We also worked on proposing an argumentation theory model to reason over the inconsistent information sets obtained from DBpedia multilingual chapters, and provide nevertheless a unique and motivated answer to the user.

A demo of the system is available online [6]. The results of this research have been published in [26], [27], [54], [34].

6.2. Combining Argumentation Theory and Natural Language Processing

Participants: Elena Cabrio, Serena Villata.

With the growing use of the Social Web, an increasing number of applications for exchanging opinions with other people are becoming available online. To cut in on a debate, the participants need first to evaluate the opinions of the other users to detect whether they are in favor or against the debated issue. An automated framework to detect the relations among the arguments represented by the natural language formulation of the users opinions is therefore needed. The work in this area proposes the use of natural language techniques to identify the arguments and their relations. In particular, the textual entailment approach is adopted, i.e. a generic framework for applied semantics, where linguistic objects are mapped by means of semantic inferences at a textual level. Textual entailment is then coupled with an abstract bipolar argumentation system which allows to identify the arguments that are accepted in the considered online debate.

The same framework is also experimented to support the management of argumentative discussions in wiki-like platforms. The results of this research have been published in [16], [28], [29].

6.3. Understanding Query Behavior and Explaining Linked Data

Participants: Fabien Gandon, Rakebul Hasan.

Our main research is to understand how to assist users in querying [63] and consuming [64] Linked Data. In querying Linked Data, we help users by providing information on how a query may behave. In addition, we provide information about the behavior of similar queries executed in the past. Users can use these information for query construction and refinement. Accurately predicting query behavior is also important for workload management, query scheduling, query optimization. In consuming Linked Data, we explain why a given piece of data exists and how the data was derived. Users can use these explanations to understand and debug Linked Data. Overall, we address the followings research questions:

i. *How to predict query behavior prior to executing the query?*

ii. *How to explain Linked Data?*

[6] http://qakis.org/qakis2/

6.3.1. Predicting query behavior

To predict query behavior prior to query execution, we apply machine learning techniques on the logs of executed queries. We work with SPARQL queries and predict how long a query would take to execute. We use the frequencies and the cardinalities of SPARQL algebra operators of a query as its features. We also extract a compact set of features from the basic graph patterns belonging to the query. We achieve high accuracy ($R^2 = 0.837$) using the k-nearest neighbors regression. We also suggest similar queries from the query log using an efficient neighbors search. Users can use these suggestions to understand behaviors of similar past queries, and construct and refine their queries accordingly.

6.3.2. Explaining Linked Data

The diverse and distributed nature of Linked Data presents opportunities for large-scale data integration and reasoning over cross-domain data. In this scenario, consumers of Linked Data may need explanations for debugging or understanding ontologies. A consumer may also want a short explanation to have an overview of the reasoning. We propose to publish the explanation related metadata as Linked Data. This enables us to explain derived data in the distributed setting of Linked Data. We present the *Ratio4TA*[7] vocabulary to describe explanation metadata and guidelines to publish these metadata as Linked Data. In addition, we summarize explanations using four measures: centrality, coherence, abstractness, and similarity. Users can specify their explanation filtering criteria - types of information they are interested in. We evaluate our summarization approach by comparing the summarized explanations generated by our approach and ground truth summarized explanations generated by humans. Our explanation summarization approach performs roughly with 60% to 70% accuracy for small summaries.

6.4. Linguistic Knowledge Representation: the Unit Graphs Formalism

Participants: Fabien Gandon, Maxime Lefrançois.

As any community of interest, linguists produce knowledge. Generic needs arise with such produced knowledge: how to represent it, how to manipulate it, how to share it, how to query it, and how to reason with it. To answer these needs is the goal of the knowledge representation (KR) domain. Existing KR formalisms such as the Semantic Web formalisms are standard solutions, and their specialization to the linguistic domain is under active development. Yet, the description logic behind the OWL formalism fails to represent how the meaning of words combine to build up the meaning of sentences. To tackle this specific problem, we introduced the new so-called Unit Graphs KR framework that is portable to existing KR standards but that introduces its own formal logic. UGs are defined over a UG-support that contains: i) a hierarchy of unit types which is strongly driven by the actantial (from action) structure of unit types, ii) a hierarchy of circumstantial symbols, and iii) a set of unit identifiers. On these foundational concepts, we defined UGs, justified the introduction of a deep-semantic representation level for the Meaning–Text Theory, we represented lexicographic definitions of lexical units, and we introduced two formal semantics: one based on UGs closure and homomorphism, and one based on model theoretic semantics. The UGs formalism has been the object of 6 publications in [42], [43], [44], [45], [46], [47].

6.4.1. Editor of Formal Lexicographic Definitions

Participants: Fabien Gandon, Alain Giboin, Romain Gugert, Maxime Lefrançois.

A prototype of a GUI of an editor of formal dictionary definitions aimed at lexicographers was developed based on the formalism of Units Graphs and on Meaning-Text Theory. The development of the GUI was preceded by the elaboration of scenarios of how users would interact with Units Graphs objects. It was followed by user tests of the GUI with actual lexicographers unfolding the scenarios. This work is reported in [46].

6.5. Access Control and Presentation for Linked Data

Participants: Luca Costabello, Fabien Gandon, Serena Villata.

[7] http://ns.inria.fr/ratio4ta/

PRISSMA is an *adaptive* rendering engine for Linked Data resources. PRISSMA tweaks RDF visualization to the mobile context in which the resource consumption is performed. The work in 2013 has been focused on designing the algorithm that selects the best RDF visualization according to the real, sensed context. Such *selection algorithm* finds optimal error-tolerant subgraph isomorphisms between RDF graphs using the notion of graph edit distance and is sublinear in the number of context declarations in the system. The PRISSMA selection algorithm has been implemented as an Android library, and a test campaign assessed response time and memory consumption. A proof-of-concept, PRISSMA-equipped, mobile RDF browser has been developed to test PRISSMA in a real-world application.

We proposed an extension of Shi3ld, a context-aware access control framework for the Web of Data, developed last year. In particular, we showed how the Shi3ld attribute-based authorization framework for SPARQL endpoints has been progressively converted to protect HTTP operations on RDF. We started by supporting the SPARQL 1.1 Graph Store Protocol and shifted towards a solution without SPARQL for the Linked Data Platform. The resulting authorization framework provides the same functionalities of its SPARQL-based counterpart, including the adoption of Semantic Web languages only. Moreover, a user-friendly interface allowing non expert users to create Shi3ld access policies through a GUI has been designed and developed. These results have been published in [22], [36], [35].

Luca Costabello co-supervised a six-month master student internship related to the Shil3d project. The student, Iacopo Vagliano, from Politecnico di Torino (Italy) developed a Web application to manage Shi3ld access control policies.

6.6. Reasoning about Data Licensing in the Web of Data

Participants: Fabien Gandon, Serena Villata.

In the domain of Linked Open Data a need is emerging for developing automated frameworks able to generate the licensing terms associated to data coming from heterogeneous distributed sources. Together with Guido Governatori (NICTA, Australia) and Antonino Rotolo (University of Bologna, Italy), we proposed and evaluated a deontic logic semantics which allows to define the deontic components of the licenses, i.e., permissions, obligations, and prohibitions, and generate a composite license compliant with the licensing items of the composed different licenses. The AND-composition and OR-composition heuristics have been proposed to support the data publisher in choosing the licenses composition strategy which better suits her needs w.r.t. the data she is publishing. The approach has been evaluated using the SPINdle defeasible reasoner, where the proposed heuristics have been hard coded in the reasoner. The results of this research line have been published in [50], [38].

6.7. Semantic and Temporal Analysis of Online Communities

Participants: Catherine Faron Zucker, Fabien Gandon, Zide Meng.

This work is done in the PhD of Zide Meng in the OCKTOPUS ANR project.

Data Formalization: We use *FOAF* and *SIOC* schema to formalize a dataset from the popular question-answer site *StackOverflow* into RDF format. For some mis-matched vocabulary, we introduce *ugc* schema, which refer to user generated content. Moreover, in order to enrich the dataset, we link tag entity of our dataset to the corresponding entity in DBpedia by using cosine distance of two entities description to solve the disambiguation problem.

Analysis: After formalizing the dataset, we begin to exploit some graph mining algorithms, such as community detection algorithm, to analyse the dataset. We extract different kinds of graph from the RDF dataset, such as question-answer graph, co-answer graph, tag co-occurrence graph etc. We aim at finding useful information such as interest groups, experts and tag groups from this kind of question-answer site. By studying the state of the art of community detection algorithm, we analyse the advantage and disadvantage of different approaches, then try to introduce a better algorithm which could outperform others in this scenario.

Plan: During our analysis, we find out some difficult problems which haven't been well solved, such as question intent understanding and community evolution. We will use semantic technology, combining with social network analysis to solve this problem. In the future, we would develop an information management system for such dataset by using analysis algorithms we introduced to improve the performance of information retrieval on user generated content sites.

6.8. RDF Mining

Participant: Andrea Tettamanzi.

We started investigating an approach to RDF mining based on grammatical evolution and possibility theory: the aim is to mine large RDF graphs by automatically generating and testing OWL 2 axioms based on the known facts. This research effort brings together expertise on metaheuristics for machine learning and data mining, fuzzy logic and possibility theory for representing and handling uncertainty, and the core interests of the Wimmics team, namely, knowledge graphs and the Semantic Web.

Finally, an article describing work on the automatic design of multilayer feedforward neural networks with evolutionary algorithms carried out while still at the University of Milan, got published in [75].

6.9. Combination of Evolutionary and Semantic Web Techniques for Protein Design

Participants: David Simoncini, Andrea Tettamanzi.

Proteins are fundamental components of all living cells and are among the most studied biological molecules. They are involved in numerous diseases and being able to determine their 3D structures and interactions is essential to understand the mechanisms of cell functions. *De novo* computational protein design refers to the problem of finding a sequence of amino acids corresponding to a protein with the desired three-dimensional structure, or the desired biological function. It is a longstanding goal in computational structural biology and only a few examples of successful *de novo* computational protein designs can be found in the literature. Computational protein design has many industrial applications, such as biofuels, drug synthesis and food processing (through computational design of enzymes) or targetted drug delivery systems (through bio-nanotechnologies).

In this context, our research focuses on knowledge extraction from protein structure databases for the development of new computational protein design frameworks. Whereas most of the current methods ignore available structural information, our algorithm takes into account known profitable interactions between amino acids and uses this information to guide the energy minimization process and propose more realistic sequences of proteins.

6.10. Logical Foundations of Cognitive Agents

Participants: Andrea Tettamanzi, Serena Villata.

We carried on work on the logical foundations of cognitive agents in collaboration with Salem Benferhat of CRIL and Célia da Costa Pereira of I3S [25] and on the application of such theoretical framework to the problem of exploiting untrustworthy communication in vehicular ad-hoc networks, in collaboration with Ana L. Bazzan and Andrew Koster of the Federal University of Rio Grande do Sul in Brasil and Célia da Costa Pereira of I3S [41]; still related to the issue of trust in multi-agent systems, we took part, with Serena Villata and Célia da Costa Pereira of I3S in a joint investigation with a research team, led by Cristiano Castelfranchi, of the CNR-ISTC in Rome [19].

6.11. Requirement Engineering

Participants: Isabelle Mirbel, Zeina Azmeh.

The participation of stakeholders (and especially end-users) in requirement engineering is recognized as a key element in the development of useful and usable systems. But in practice, the involvement of end-users is often difficult to implement. Today's Web has given rise to several platforms serving the purpose of collaborative software development. Thanks to these environments, it is possible, among others, for anyone to suggest new requirements for a software under development. A lot of requirements are thus proposed by users and it becomes difficult, after a while, for the persons in charge of the software which development is hosted by the platform to understand this large set of new requirements in its entirety. An important limitation of these new approaches resides in the information overload, lacking structure and semantics.

In this context, we proposed an approach based on Semantic Web languages as well as concept lattices to identify relevant groups of stakeholders depending on their past participation. We also developped a tool supporting this approach. This work relies on Semantic Web languages and formal concept analysis. Semantic Web languages are used to annotate the data extracted from the plateform and to reason about it. Formal Concept Analysis is a theory of data analysis which identifies conceptual structures among data sets. We use it to classify users as well as requirements into lattices which can then be exploited as road maps to examine new requirements. The results of this research have been published in [24].

6.12. Management of Technical and Regulatory Knowledge

Participants: Khalil Bouzidi, Michel Buffa, Catherine Faron Zucker, Nhan Le Than.

In the framework of a long-term collaboration with CSTB (Centre Scientifique et Technique du Bâtiment) on the management of technical and regulatory knowledge based on Semantic Web models and techniques, Catherine Faron Zucker and Nhan Le Than co-supervised the PhD thesis of Khalil Riad Bouzidi which has been defended on September 2013.

In the continuation of this work, Catherine Faron Zucker and Michel Buffa got involved with CSTB and three other partners on a project proposal submitted to ANR on the recommendation of technical documents in a social network of building professionals, based on the capitalization and sharing of best search practices.

6.13. Co-Construction of Community Ontologies and Corpus in a Limited Technological Environment

Participants: Olivier Corby, Papa Fary Diallo, Isabelle Mirbel.

In this thesis, we study the implementation of an online platform to build and share the collective memory of citizens in Senegal and revive stories by using a semantic layer. During the first year of this thesis, the first step has been to describe some Use Cases about the platform we would like to develop. We started to define what community means in our work which are group of people with a shared history, culture, ethnicity or interest and want to exchange or collaborate via the Web to share their knowledge of this area.

Our communities are characterized by three components: 1) a common socio-cultural interest, 2) exchange, collaboration and sharing among members and 3) use of the Internet to interact. Thanks to the use cases, we define two main types of users. A community member who is an user who participates in the construction of information and who has interactions with other users. The second type is a simple user who visits the platform for having information, he can be a tourist who want to have information about Senegalese communities activities.

With these use cases, we determine some features that the platform should have. Community members should have, among other thing, a place where they interact to collaborate. To have a "living" community, the system must notify the members of the community about new entries on their focus. Also, to have a catchy presentation, we plan to use maps with different kinds of information.

The second step has been to do a state of the art of online communities. This review allows us to find different definitions and typologies which differ from the study domain – anthropologist, sociologist, psychologist – or the objective – demographic, technological environment, members characteristic. The broadest definition takes into account our concept of community in the context of a knowledge-sharing platform is that of Porter [80]. Despite the fact that numerous typologies are proposed, none is completely consistent with our vision of community. However, the "Toronto School" proposes a category in the classification based on the knowledge transmission called ”knowledge-building community” applied in the education area. We think that this type of community could be generalized in the field of socio-constructivisme development, which our communities belong to, for sharing socio-cultural knowledge.

Then, the second phase of this review has been to present the WestAfricapedia project which takes place in this thesis. The main objective is to enhance and sustain the socio-cultural heritage of Senegalese communities through a framework of sharing and co-construction of sociocultural knowledge. Thus we distinguish two main types of communities: knowledge-building community extended in the culture area and exchange information community that has sub-categories such as sports community, commercial community, etc.

6.14. Semantic Wiki

Participants: Pavel Arapov, Michel Buffa.

We worked on Semantic Web tools, more particularly on WikiNEXT, a semantic application wiki. WikiNEXT lies on the border between application wikis and modern Web based IDEs like jsbin [8], JSFIDDLE[9], cloud9 IDE[10], etc. It has been initially created for writing documents that integrate data from external data sources of the Web of Data, such as DBPedia.org or FreeBase.com, or for writing interactive tutorials (e.g. an HTML5 tutorial, a semantic Web programming tutorial) that mix text and interactive examples in the same page. The system combines some powerful aspects from (i) wikis, such as ease of use, collaboration and openness, (ii) semantic Web/wikis such as making information processable by machines and (iii) Web-based IDEs such as instant development and code testing in a Web browser.

WikiNEXT can be used for writing documents/pages as well as for writing Web applications that manipulate semantic data, either locally or coming from the Web of Data. These applications can be created, edited or cloned in the browser and can be used for integrating data visualizations in wiki pages, for annotating content with metadata, or for any kind of processing. WikiNEXT is particularly suited for teaching Web technologies or for writing documents that integrate data from the Web of data.

6.15. Semantic Aggregation

Participant: Christophe Desclaux.

Christophe spent one year in the Wimmics team (October 2012 to October 2013) as an invited engineer funded by the BoostYourCode contest he won in 2012. The aim of the BoostYourCode contest (organized by Inria) is to offer to a junior engineer a one year full time contract to work on an innovating OpenSource project.

We worked on an RSS feed aggregation tool using Named Entities Recognition. Reador.NET [11] provides a specialized tool for monitoring news from various sources like RSS, twitter or facebook feeds. Reador.NET semantically increases news for a better classification for the user. We worked on document clustering, natural language processing, RDF datastores and building efficient SPARQL queries.

6.16. Semantic Mappings

Participants: Thi Hoa Hue Nguyen, Nhan Le Thanh.

[8] http://www.jsbin.com
[9] http://www.jsfiddle.net
[10] http://www.cloud9ide.com
[11] http://www.reador.net

This PhD thesis is about semantic mappings with a control flow-based business workflow: an approach to develop control flow applications using knowledge-based systems.

Although software systems employed to create and execute automatically business processes have been becoming more and more available and advanced, each system is built to deal with a particular workflow type. In addition, these systems require a great deal of time and effort of expert programmers as well as the knowledge of domain experts to set up. Therefore, it is desirable to develop an alternative approach.

Our objective is to represent control flow-based business workflow patterns (CBWPs) in knowledge base by a declarative approach. We first propose an ontological model to represent Coloured Petri Nets (CPNs) with OWL DL. On this basis, we define a meta-knowledge base for CBWPs management. We then develop a graphical interface to design and simulate CBWPs. Our ongoing work is to develop a middleware prototype for mapping and using a CBWP with a user's knowledge base in order to illustrate the feasibility of our approach [49].

6.17. Emotional and Social Web

6.17.1. Modeling, Detection and Annotation of Emotional States using an Algebraic Multidimensional Vector Space

Participants: Nhan Le Thanh, Imen Tayari.

In this research work, we presents a generic solution of emotional data exchange between heterogeneous multi-modal applications. This proposal is based on a new algebraic representation of emotions and it is composed of three distinct layers: the psychological layer, the formal computational layer and the language layer. Moreover, our proposal provides powerful mathematical tools for the analysis and the processing of these emotions and it enables the exchange of the emotional states regardless to the modalities and sensors used in the detection step. The validation of the proposed solution is done with K-nearest neighbor classification algorithm for detecting and evaluating emotion from Eight-Emotion Sentics Data.

6.17.2. Social radio: a Case Studies of Social Network Services

Participants: Amosse Edouard, Nhan Le Thanh.

In this project, we carry out some case studies of social radio that is an information service on social networking. Two case studies are conducted on the topics of traffic incidents and geo-epidemiologies. These case studies allow us to study a formal model of spatiotemporal annotations on social network.

6.17.3. Participatory Mapping and Social Bookmarking

Participants: Michel Buffa, Alain Giboin.

In continuation of ISICIL, collaboration began this year between the ITCS-HSS research teams Wimmics and Tech-CICO (UTT), in association with Mnemotix and Wannago startups, in order to design a platform of participatory mapping in the field of sustainable tourism. This platform will enable the various actors in the field (tourists, tourism service providers, scientific experts of fauna, flora and geology, associations, and so on) exchange knowledge about the site and thus enhance the site attraction. This platform is called "socio-semantic" because it offers a unique combination of Semantic Web and Social Web technologies.

The article [51] details one of the planned scenarios of use of the platform and illustrates some proposed functionalities such as Webmarks (Wimmics and Mnemotix) and multiple viewpoints (Tech-CICO). The article also shows how ICT and HSS researchers will collaborate to analyze the innovative uses of the platform on the first fields of application (in Provence- Alpes-Côte d'Azur Region).

6.17.4. Modeling Team Processes

Participants: Pierre Robillard, Isabelle Mirbel, Zeina Azmeh, Alain Giboin, Mathieu Lavallée.

Recent studies outline the importance of software development teams' interactions, suggesting that poor team dynamics can lead to poor software. The relationship between "soft" issues like team dynamics and "hard" issues like software quality is difficult to observe, however. To bridge the gap between these two kinds of issues, and to help development teams prevent quality issues through the planning of relevant team activities, we worked on an assessment method of the quality of team dynamics based on a taxonomy of episodes of interactions encountered in software development teams [83] – the CoDyMA (Collaborative Dynamics Measurement and Analysis) method. We proposed an analysis procedure of episodes based on the Formal Concept Analysis (FCA) approach. This procedure uses as input the data (namely the accounts of face-to-face interactions) reported by the developers in their activity diary. The entries are coded in terms of interaction episodes and artifact types to produce a FCA lattice. The observed lattice is compared to a prescribed lattice, and adjustments can be proposed to the team if necessary. The procedure was applied to data from a case study. This work is described in a paper submitted for publication.

6.17.5. Modeling Users and Groups of Users

Participants: Isabelle Mirbel, Zeina Azmeh, Alain Giboin.

6.17.5.1. Emphasizing Dysfunctional Group Dynamics in Collaboration Personas: Specification of an Approach

Comparing Collaboration Personas and Individual Personas for the design and evaluation of collaboration software, Judge, Matthews, and Whittaker [79] found that practitioners preferred collaboration personas, but required that the method put more emphasis on problematic or dysfunctional group dynamics. Because Judge et al. only outlined a possible approach to meet this requirement, we decided to contribute to the specification of the approach. In [37] we report the first steps of this specification work.

6.17.5.2. Using Formal Concept Analysis to elicit Personas

Personas are built from a clustering of behavioral variables common to a set of users. Behavioral variables are ways in which users behavior differ (e.g., goals and attitudes); it is important to elicit them because they have an impact on the system to be designed. Today, the clustering is mainly performed manually. To automate it, we started this year to explore the use of Formal Concept Analysis tools.

6.17.6. Modeling Multimodal Grounding Processes in Design Teams

Participants: Aurore Defays, Alain Giboin.

Grounding is the process used by participants to a collective activity to coordinate both the content and process of their communication to be successful [71]. Grounding is also defined as the process of elaborating and maintaining the Common Ground (i.e., mutual knowledge, mutual beliefs, and mutual assumptions) necessary to participants' mutual understanding [72]. So far, grounding has been studied mainly from a unimodal point of view, i.e., from the point of view of the verbal modality (oral or written). Some authors have begun to study grounding from a bimodal viewpoint; for example [78] have studied the use of verbal and gestural modalities necessary to ensure mutual understanding in interactions between Japanese airline pilots and an American flight instructor. In the context of her PhD thesis in Ergonomics applied to architectural design digital tools, Aurore Defays extended the study of grounding to actual multimodality (with n modality: oral, written, gestures, gazes, etc.).

With Aurore Defays, we focused this year on improving the methodology of analysis of multimodal grounding proposed in [74]. To do this, we relied on the data of an existing study by Defays on remote collaboration between dyads and triads of architects interacting through a collaborative digital studio (the Distributed Collaborative Digital Studio, DSDC). Our initial research question was: Is a multimodal shared representation preferable to a unimodal representation to collaborate effectively? Analyzing the data, this question was gradually transformed into: Which modalities are relevant to build the common ground necessary for a particular type of collaboration to succeed?

6.17.7. The "Design Thinking" Toolset: Application to Discovery Hub

Participants: Gessica Puri, Alain Giboin, Nicolas Marie, Damien Legrand.

Last year was developed a ”design thinking” toolset (including a framework) for helping developers think in terms of a user’s point of view when they design and evaluate link visualization and manipulation applications such as graph visualization applications [81]. The toolset was used this year to perform a qualitative evaluation of Discovery Hub, so contributing to the development of a new version of the discovery engine.

The ShowCaseMachine project led by Damien Legrand won the 11th Challenge *Jeune Pousses* at Telecom Valley in Sophia Antipolis.

6.18. Graph-based Knowledge Representation

6.18.1. SPARQL Based Pretty Printing Language

Participant: Olivier Corby.

We have designed SPARQL Template, a pretty-printing rule language for RDF graphs. It enables to pretty print RDF graphs representing Abstact Syntax Trees of languages such as SPIN or OWL RDF syntax. We have implemented a pretty printing engine that interprets SPARQL Template.

An example of template for a OWL ”someValuesFrom” statement is shown below. The SPARQL 1.1 ”where” part specifies the conditions to apply the rule on a focus node ”?in”. The template part specifies the result of the pretty print of the focus node. Variables in the template part are recursively replaced by the result of their pretty print.

```
template {
  "someValuesFrom(" ?p " " ?c ")"
}
where {
  ?in a owl:Restriction ;
    owl:onProperty ?p ;
    owl:someValuesFrom ?c
}
```

We have introduced named templates that are called explicitely using a ”kg:template” extension function.

The pretty printing language and engine have been validated on five RDF AST [12]: SPIN, OWL 2, SQL, Turtle and a mockup of mathematical expressions pretty printed into Latex. The SPIN pretty printer is used in the PhD Thesis of Oumy Seye on "Rules for the Web of Data" and the SQL pretty printer is used in the PhD Thesis of Corentin Follenfant on "Usage semantics of analytics and Business Intelligence tools".

6.18.2. Federated Semantic Data Query

Participants: Olivier Corby, Alban Gaignard.

Another activity of the team addresses the data explosion challenges faced in e-Science. Semantic Web technologies are well adopted to represent the knowledge associated to both e-Science data and processing tools. A PhD thesis [76], addressing the distributed knowledge production and sharing in collaborative e-Science platforms, has successfully been defended this year. Moreover, we have been participating in the organization of the second edition of the CrEDIBLE workshop [13], gathering international experts to discuss the challenges of federating distributed biomedical imaging data and knowledge.

In this area, the main scientific results are (i) a software architecture for transparently querying multiple data sources through the SPARQL language [73], (ii) a set of querying strategies and optimizations dedicated to limit the cost of distributed query processing, while still considering enough expressivity (full SPARQL 1.1 support, including named graphs, property path expressions, optional, aggregates, etc.).

[12] ftp://ftp-sop.inria.fr/wimmics/soft/pprint
[13] http://credible.i3s.unice.fr

Performance-oriented experiments have been conducted on the Grid'5000 distributed computing infrastructure to compare our approach with state-of-the-art engines such as FedX [85], Splendid [77], or DARQ [82]. Experiments, based on the FedBench benchmark [84] show performances between DARQ, Splendid, and FedX, while still high expressivity.

Since distributed query processing lead to complex and costly processes, we started to collect provenance information which opens interesting perspectives towards enhanced trust and reproducibility in Linked Data querying and reasoning.

These distributed query processing strategies have been implemented and integrated into Corese through two main components, namely a data source federator, and a data source endpoint. A prototype Web application has also been developed to demonstrate our approach. End-users can configure and launch distributed SPARQL querying and finally visualize SPARQL results and their associated provenance.

6.18.3. Rules for the Web of Data

Participants: Olivier Corby, Catherine Faron Zucker, Oumy Seye.

This work takes place in the PhD Thesis of Oumy Seye.

The objective of this year is to foster knowledge reuse on the Web based on the principles of Linked Data. Our approach is to consider rule bases like data sources that can be published, shared and queried as Linked Data, thus enabling the selection and reuse of relevant and useful shared rules in any particular context or application. We propose to select rules by querying either metadata annotating rules, rules content or both. To make rules content queryable, we use RDF representations of SPARQL rules with the SPIN format [14].

This idea joins the principles of the Semantic Web that encourages the sharing and reuse of knowledge. We used the SPIN syntax (which allows the representation of a SPARQL query in RDF) obtained with the SPIN pretty printer of Corese. We have subsequently been able to select rules of interest with Corese. The proposal enables to search rules based on their content. This allows us to help users extract relevant set of rules for their data, and thus leverage more easily shared rules. This idea can be used to build a search engine for rules on the Web or a tool for automatically connect rules with semantic data.

In the remainder of this work, we will focus on updating harvested rules. A poster on this work was presented for the GLC pole day July 8, and at the summer school ESWC September 2.

6.18.4. Semantic Web and Business Intelligence

Participants: Corentin Follenfant, Olivier Corby, Fabien Gandon.

This PhD Thesis is done with a CIFRE industrial grant from SAP Research.

The bilateral contract with SAP aims at converging Semantic Web and Business Intelligence through a framework applying the read/write Web principles to the business knowledge carried within Business Intelligence reports. These reports often provide a dynamic view upon numerical data from various enterprise sources, mainly relational databases. Reports are authored with a complex process that can be reduced to writing, directly or through different layers of user interfaces, SQL queries that will query the sources and feed the dynamic reports. In order to simplify the query authoring process, complementary approaches are envisioned.

Our approach proposes to model the queries as knowledge through their abstract syntax trees (ASTs) with Semantic Web tools, query and manipulate them through appropriate standards, respectively RDF/S and SPARQL. Indeed RDF enables us to model the actual structure of the ASTs by integrating the knowledge related to syntax and semantics of the SQL queries: types can be captured with XML Schema Datatypes, while more specific business knowledge can also be designed according to the source business models and annotate various entities referenced within the SQL queries. Regarding the query and manipulation part, a library of SPARQL queries was designed to perform generic AST manipulation (generic from a DSL perspective), and is usable to search, extract, edit, prune or graft parts of RDF-modelled ASTs.

[14] http://www.w3.org/Submission/spin-overview/

While this year was mostly dedicated to manuscript writing, additional experiments were run to demonstrate the validity of our model: a large set of ANSI SQL queries generated with a TPC-DS benchmark was converted to its RDF representation. Inversely, a generic pretty printer system developed into the Corese engine was validated by the internship of Abdoul Macina who developed a set of rules to have the pretty printer turn RDF-modelled SQL queries back to their concrete syntactic form. This enables iterative query design by leveraging AST patterns rather than manually editing brute syntax.

ZENITH Project-Team

6. New Results

6.1. Big Data Integration

6.1.1. Probabilistic Data Integration

Participants: Reza Akbarinia, Naser Ayat, Patrick Valduriez.

Data uncertainty in scientific applications can be due to many different reasons: incomplete knowledge of the underlying system, inexact model parameters, inaccurate representation of initial boundary conditions, inaccuracy in equipments, error in data entry, etc.

An important problem that arises in big data integration is that of Entity Resolution (ER). ER is the process of identifying tuples that represent the same real-world entity. The problem of *entity resolution over probabilistic data* (which we call ERPD) arises in many distributed application domains that have to deal with probabilistic data, ranging from sensor databases to scientific data management. The ERPD problem can be formally defined as follows. Let e be an uncertain entity represented by multiple possible alternatives, i.e. tuples, each with a membership probability. Let D be an uncertain database composed of a set of tuples each associated with a membership probability. Then, given e, D, and a similarity function F, the problem is to find the entity-tuple pair (t, t_i) (where $t \in e, t_i \in D$) such that (t, t_i) has the highest cumulative probability to be the most similar in all possible worlds. This entity-tuple pair is called the *most probable match pair* of e and D, denoted by MPMP(e, D).

Many real-life applications produce uncertain data distributed among a number of databases. Dealing with the ERPD problem for distributed data is quite important for such applications. A straightforward approach for answering distributed ERPD queries is to ask all distributed nodes to send their databases to a central node that deals with the problem of ER by using one of the existing centralized solutions. However, this approach is very expensive and does not scale well neither in the size of databases, nor in the number of nodes.

In [20], we proposed FD (Fully Distributed), a decentralized algorithm for dealing with the ERPD problem over distributed data, with the goal of minimizing bandwidth usage and reducing processing time. It has the following salient features. First, it uses the novel concepts of *Potential* and *essential-set* to prune data at local nodes. This leads to a significant reduction of bandwidth usage compared to the baseline approaches. Second, its execution is completely distributed and does not depend on the existence of certain nodes. We validated FD through implementation over a 75-node cluster and simulation using both synthetic and real-world data. The results show very good performance, in terms of bandwidth usage and response time.

6.1.2. Open Data Integration

Participants: Emmanuel Castanier, Patrick Valduriez.

Working with open data sources can yield high value information but raises major problems in terms of metadata extraction, data source integration and visualization. For instance, Data Publica provides more than 12 000 files of public data. However, even though data formats become richer and richer in terms of semantics and expressivity (e.g. RDF), most data producers do not use them much in practice, because they require too much upfront work, and keep using simpler tools like Excel. Unfortunately, no integration tool is able to deal in an effective way with spreadsheets. Only few initiatives (OpenII and Google Refine) deal with Excel files. However, their importers are very simple and make some strict restrictions over the input spread-sheets.

In [31], we describe a demonstration of WebSmatch, a flexible environment for Web data integration. WebSmatch supports the full process of importing, refining and integrating data sources and uses third party tools for high quality visualization. We use a typical scenario of public data integration which involves problems not solved by currents tools: poorly structured input data sources (XLS files) and rich visualization of integrated data.

6.1.3. Pricing Integrated Data

Participant: Patrick Valduriez.

Data is a modern commodity, being bought and sold. Electronic data market places and independent vendors integrate data and organize their online distribution. Yet the pricing models in use either focus on the usage of computing resources, or are proprietary, opaque, most likely ad hoc, and not conducive of a healthy commodity market dynamics. In [39], we propose a generic data pricing model that is based on minimal provenance, i.e. minimal sets of tuples contributing to the result of a query. We show that the proposed model fulfills desirable properties such as contribution monotonicity, bounded-price and contribution arbitrage-freedom. We present a baseline algorithm to compute the exact price of a query based on our pricing model. We show that the problem is NP-hard. We therefore devise, present and compare several heuristics. We conduct a comprehensive experimental study to show their effectiveness and effciency.

In most data markets, prices are prescribed and accuracy is determined by the data. Instead, we consider a model in which accuracy can be traded for discounted prices: "what you pay for is what you get". The data market model consists of data consumers, data providers and data market owners. The data market owners are brokers between the data providers and data consumers. A data consumer proposes a price for the data that she requests. If the price is less than the price set by the data provider, then she gets an approximate value. The data market owners negotiate the pricing schemes with the data providers. They implement these schemes for the computation of the discounted approximate values. In [38], we propose a theoretical and practical pricing framework with its algorithms for the above mechanism. In this framework, the value published is randomly determined from a probability distribution. The distribution is computed such that its distance to the actual value is commensurate to the discount. The published value comes with a guarantee on the probability to be the exact value. The probability is also commensurate to the discount. We present and formalize the principles that a healthy data market should meet for such a transaction. We define two ancillary functions and describe the algorithms that compute the approximate value from the proposed price using these functions. We prove that the functions and the algorithm meet the required principles.

6.2. Distributed Indexing and Searching

6.2.1. P2P Search and Recommendation

Participants: Esther Pacitti, Maximilien Servajean.

In crossdiscipline domains, users belonging to different communities produce various scientific material that they own, share, or endorse. In that context, we are interested in querying and recommending scientific material in the form of documents. Such documents cover various topics such as models for plant phenotyping, statistics on specific kinds of plants, or biological experiments.

In [40], we investigate profile diversity, a novel idea in searching scientific documents. Combining keyword relevance with popularity in a scoring function has been the subject of different forms of social relevance. On the other hand, content diversity has been thoroughly studied in search and advertising, database queries, and recommendations.

We introduce profile diversity for scientific document search as a complement to traditional content diversity. Profile diversity combines the discipline and communities to which a user belongs. We propose an adaptation of Fagin's threshold-based algorithms to return the most relevant and most popular documents that satisfy content and profile diversities. To validate our scoring function, DivRSci, we ran experiments that use two benchmarks: a realistic benchmark with scientists and TREC'09. We show that DivRSci presents the best compromise between all requirements we have identified. DivRSci also shows to be the best generating list of inter-disciplinary and inter-community documents. Finally, it yields very good gains (by a factor of 6), suited for profile diversification

6.2.2. Spatial Queries in Wireless Data Broadcasting

Participant: Patrick Valduriez.

The main requirements for spatial query processing via mobile terminals include rapid and accurate searching and low energy consumption. Most location-based services (LBSs) are provided using an on-demand method, which is suitable for light-loaded systems where contention for wireless channels and server processing is not severe. However, as the number of users of LBSs increases, performance deteriorates rapidly since the servers' capability to process queries is limited. Furthermore, the response time of a query may significantly increase with the concentration of users' queries in a server at the same time. That is because the server has to check the locations of users and potential objects for the final result and then individually send answers to clients via a point-to-point channel. At this time, an inefficient structure of spatial index and searching algorithm may incur an extremely large access latency.

To address this problem, we propose in [27] the Hierarchical Grid Index (HGI), which provides a light-weight sequential location-based index structure for efficient LBSs. We minimize the index size through the use of hierarchical location-based identifications. And we support efficient query processing in broadcasting environments through sequential data transfer and search based on the object locations. We also propose Top-Down Search and Reduction- Counter Search algorithms for efficient searching and query processing. HGI has a simple structure through elimination of replication pointers and is therefore suitable for broadcasting environments with one-dimensional characteristics, thus enabling rapid and accurate spatial search by reducing redundant data. Our performance evaluation shows that our proposed index and algorithms are accurate and fast and support efficient spatial query processing.

6.3. Big Data Analysis

6.3.1. Big Data Analysis using Algebraic Workflows

Participants: Jonas Dias, Patrick Valduriez.

Analyzing big data requires the support of dataflows with many activities to extract and explore relevant information from the data. Recent approaches such as Pig Latin propose a high-level language to model such dataflows. However, the dataflow execution is typically delegated to a MapReduce implementation such as Hadoop, which does not follow an algebraic approach, thus it cannot take advantage of the optimization opportunities of PigLatin algebra.

In [35], we propose an approach for big data analysis based on algebraic workflows, which yields optimization and parallel execution of activities and supports user steering using provenance queries. We illustrate how a big data processing dataflow can be modeled using the algebra. Through an experimental evaluation using real datasets and the execution of the dataflow with Chiron, an engine that supports our algebra, we show that our approach yields performance gains of up to 19.6% using algebraic optimizations in the dataflow and up to 39% of time saved on a user steering scenario.

This work was done in the context of the CNPq-Inria Hoscar project and FAPERJ-Inria P2Pcloud project .

6.3.2. Big Data Partitioning

Participants: Reza Akbarinia, Miguel Liroz, Esther Pacitti, Patrick Valduriez.

The amount of data that is captured or generated by modern computing devices has augmented exponentially over the last years. For processing this *big data*, parallel computing has been a major solution in both industry and research. This is why, the MapReduce framework, which provides automatic distribution parallelization and fault-tolerance in a transparent way over lowcost machines, has become one of the standards in big data analysis.

For processing a big dataset over a cluster of nodes, one main step is data partitioning (or fragmentation) to divide the dataset to the nodes. In our team, we study the problem of data partitioning in two different contexts: (1) in scientific databases that are continuously growing and (2) in the MapReduce framework. In both cases, we propose automatic approaches, which are performed transparently to the users, in order to free them from the burden of complex partitioning.

In [25], we consider applications with very large databases, where data items are continuously appended. Thus, the development of efficient data partitioning is one of the main requirements to yield good performance. In particular, this problem is harder in the case of some scientific databases, such as astronomical catalogs. The complexity of the schema limits the applicability of traditional automatic approaches based on the basic partitioning techniques. The high dynamicity makes the usage of graph-based approaches impractical, as they require to consider the whole dataset in order to come up with a good partitioning scheme. In our work, we propose *DynPart* and *DynPartGroup*, two dynamic partitioning algorithms for continuously growing databases [25]. These algorithms efficiently adapt the data partitioning to the arrival of new data elements by taking into account the affinity of new data with queries and fragments. In contrast to existing static approaches, our approach offers constant execution time, no matter the size of the database, while obtaining very good partitioning efficiency. We validate our solution through experimentation over real-world data; the results show its effectiveness.

In [37] and [43], we address the problem of high data transfers in MapReduce, and propose a technique that repartitions tuples of the input datasets. Our technique optimizes the distribution of key-values over mappers, and increases the data locality in reduce tasks. It captures the relationships between input tuples and intermediate keys by monitoring the execution of a set of MapReduce jobs which are representative of the workload. Then, based on those relationships, it assigns input tuples to the appropriate chunks. With this data repartitioning and a smart scheduling of reducer tasks, our approach significantly contributes to the reduction of transferred data between mappers and reducers in job executions. We evaluate our approach through experimentation in a Hadoop deployment on top of Grid5000 using standard benchmarks. The results show high reduction in data transfer during the shuffle phase compared to Native Hadoop.

6.4. Data Stream Mining

6.4.1. Mining Uncertain Data Streams

Participants: Reza Akbarinia, Florent Masseglia.

Discovering Probabilistic Frequent Itemsets (PFI) is very challenging since algorithms designed for deterministic data are not applicable in probabilistic data. The problem is even more difficult for probabilistic data streams where massive frequent updates need to be taken into account while respecting data stream constraints. In [28], we propose FEMP (Fast and Exact Mining of Probabilistic data streams), the first solution for exact PFI mining in data streams with sliding windows. FEMP allows updating the frequentness probability of an itemset whenever a transaction is added or removed from the observation window. Using these update operations, we are able to extract PFI in sliding windows with very low response times. Furthermore, our method is exact, meaning that we are able to discover the exact probabilistic frequentness distribution function for any monitored itemset, at any time. We implemented FEMP and conducted an extensive experimental evaluation over synthetic and real-world data sets; the results illustrate its very good performance.

6.4.2. Itemset Mining over Tuple-Evolving Data Streams

Participant: Florent Masseglia.

In many data streaming applications today, tuples inside the streams may get revised over time. This type of data stream brings new issues and challenges to the data mining tasks. In [42] we present a theoretical analysis for mining frequent itemsets from sliding windows over such data. We define conditions that determine whether an infrequent itemset will become frequent when some existing tuples inside the streams have been updated. We design simple but effective structures for managing both the evolving tuples and the candidate frequent itemsets. Moreover, we provide a novel verification method that efficiently computes the counts of candidate itemsets. Experiments on real-world datasets show the efficiency and effectiveness of our proposed method.

6.5. Scalable Data Analysis

6.5.1. Scalable Mining of Small Visual Objects

Participants: Pierre Letessier, Julien Champ, Alexis Joly.

Automatically linking multimedia documents that contain one or several instances of the same visual object has many applications including: salient events detection, relevant patterns discovery in scientific data or simply web browsing through hyper-visual links. Whereas efficient methods now exist for searching rigid objects in large collections, discovering them from scratch is still challenging in terms of scalability, particularly when the targeted objects are small compared to the whole image. In a previous work, we revisited formally the problem of mining or discovering such objects, and then generalized two kinds of existing methods for probing candidate object seeds: weighted adaptive sampling and hashing based methods. This year, we continued working on the subject by improving our high-dimensional data hashing strategy, that works first at the visual level, and then at the geometric level. We conducted new experiments on a dedicated evaluation dataset [5] and we did show that our the recall or our approach definitely outperforms the reference method [46].
Based on this contribution, we then address the problem of suggesting object-based visual queries in a multimedia search engine [22], [36]. State-of-the-art visual search systems are usually based on the query-by-window paradigm: a user selects any image region containing an object of interest and the system returns a ranked list of images that are likely to contain other instances of the query object. User's perception of these tools is however affected by the fact that many submitted queries actually return nothing or only junk results (complex non-rigid objects, higher-level visual concepts, etc.). In [22], we addressed the problem of suggesting only the object's queries that actually contain relevant matches in the dataset. This requires to first discover accurate object's clusters in the dataset (as an offline process); and then to select the most relevant objects according to user's intent (as an on-line process). We therefore introduce a new object's instances clustering framework based on a bipartite shared-neighbours clustering algorithm that is used to gather object's seeds discovered by our visual mining method. Shared nearest neighbours methods were not studied beforehand in the case of bipartite graphs and never used in the context of object discovery. Experiments show that this new method outperforms state-of-the-art object mining and retrieval results on the Oxford Building dataset. We finally describe two real-word object-based visual query suggestion scenarios using the proposed framework and show examples of suggested object queries. A demo was presented at ACM Multimedia 2013 [36].
This method was finally integrated within a visual-based media event detection system in the scope of a French project called the Transmedia Observatory [33]. It allows the automatic discovery of the most circulated images across the main news media (news websites, press agencies, TV news and newspapers). The main originality of the detection is to rely on the transmedia contextual information to denoise the raw visual detections and consequently focus on the most salient trans-media events.

6.5.2. Rare Events Identification for Large-Scale Applications

Participant: Florent Masseglia.

While significant work in data mining has been dedicated to the detection of single outliers in the data, less research has approached the problem of isolating a group of outliers, i.e. rare events representing micro-clusters of less – or significantly less – than 1% of the whole dataset. This research issue is critical for example in medical applications. The problem is difficult to handle as it lies at the frontier between outlier detection and clustering and distinguishes by a clear challenge to avoid missing true positives. In [41], we address this challenge and propose a novel two-stage framework, based on a backward approach, to isolate abnormal groups of events in large datasets. The key of our backward approach is to first identify the core of the dense regions and then gradually augment them based on a density-driven condition. The framework outputs a small subset of the dataset containing both rare events and outliers. We tested our framework on a biomedical application to find micro-clusters of pathological cells. The comparison against two common clustering (DBSCAN) and outlier detection (LOF) algorithms show that our approach is a very efficient alternative to the detection of rare events – generally a recall of 100% and a higher precision, positively correlated wih the size of the rare event – while also providing a $\mathcal{O}(N)$ solution to the existing algorithms dominated by a $\mathcal{O}(N^2)$ complexity.

6.5.3. Large-scale content-based plants identification from social image data

Participants: Hervé Goëau, Alexis Joly, Julien Champ, Saloua Litayem.

[5] http://www-sop.inria.fr/members/Alexis.Joly/BelgaLogos/FlickrBelgaLogos.html

Speeding up the collection and integration of raw botanical observation data is a crucial step towards a sustainable development of agriculture and the conservation of biodiversity. Initiated in the context of a citizen sciences project in collaboration with the botanists of the AMAP UMR team and Tela Botanica social network, the overall contribution of this work [23] is an innovative collaborative workflow focused on image-based plant identification as a mean to enlist new contributors and facilitate access to botanical data. Since 2010, hundreds of thousands of geo-tagged and dated plant photographs were collected and revised by hundreds of novice, amateur and expert botanists of a specialized social network. An image-based identification tool - available as both a web and a mobile application - is synchronized with that growing data and allows any user to query or enrich the system with new observations. Extensive experiments of the visual search engine as well as system-oriented and user-oriented evaluations of the application did show that it is very helpful to determine a plant among hundreds or thousands of species [23]. As a concrete result, more than 80K people in about 150 countries did download the iPhone end point of the application [32].

From a data management and data analysis perspective, our main contribution concerns the scalability of the system. At the time of writing, the content-based search engine actually works on 120K images covering more than 5000 species (which already makes it the largest identification tool built anytime). The resulting training dataset contains several hundreds of millions feature vectors, each with several hundreds of float attributes (i.e. high-dimensional feature vectors describing the visual content). At query time, thousands of such feature vectors are extracted from the query pictures and have to be searched online in the training set to find the most similar pictures. The underlying search of approximate nearest neighbors is speed-up thanks to a data-dependent high-dimensional hashing framework based on Random Maximum Margin Hashing (RMMH), a new hash function family that we introduced in 2011. RMMH is used for both compressing the original feature vectors into compact binary hash codes and for partitioning the data into a well balanced hash table. Search is then performed through adaptive multi-probe accesses in the hash table and a top-k search refinement step on the full binary hash codes. Last improvements brought in 2013 include a multi-threaded version of the search, the use of a probabilistic asymmetric distance instead of the Hamming distance and the integration of a query optimization training stage in the compressed feature space instead of the original space. A beta version of Pl@ntNet visual search engine based on these new contributions is currently being tested and is about 8 times faster than the one used in production.

Besides scalability and efficiency, we also did work on improving the identification performances of the system [29]. We notably improved the quality of the top-K returned images by weighting each match according to its Hamming distance to the query rather than using a simple vote. We then improved the multi-cue fusion strategy by indexing separately each type of visual features rather than concatenating them in an early phase. We finally did train the optimal selection of features for each of the considered plant organ (flower, leaf, bark, fruit). Beyond the use of the visual content itself, we explored the usefulness of associated metadata and we did prove that some of them like the date can improve the identification performances (contrary to the geo-coordinates that surprisingly degraded the results). Overall, as a result of our participation to ImageCLEF plant identification benchmark [34], we obtained the second best run among 12 international groups and a total of 33 submitted runs.