Activity Report Saclay - Île-de-France 2016

List of Inria's Research Teams

Project-Team AMIB

Algorithms and Models for Integrative Biology

IN COLLABORATION WITH: Laboratoire d'informatique de l'école polytechnique (LIX)

IN PARTNERSHIP WITH:
CNRS

Ecole Polytechnique

Université Paris-Sud (Paris 11)

RESEARCH CENTER
Saclay - Île-de-France

THEME
Computational Biology

Table of contents

Project-Team AMIB

Creation of the Team: 2009 May 01, updated into Project-Team: 2011 January 01, end of the Project-Team: 2016 December 31

Keywords:

Computer Science and Digital Science:

> 3.4. - Machine learning and statistics
> 3.4.5. - Bayesian methods
> 7. - Fundamental Algorithmics
> 7.2. - Discrete mathematics, combinatorics

Other Research Topics and Application Domains:

> 1. - Life sciences
> 1.1. - Biology
> 1.1.1. - Structural biology
> 1.1.9. - Bioinformatics
> 1.1.10. - Mathematical biology
> 9.6. - Reproducibility

1. Members

Research Scientists
> Mireille Régnier [Team leader, Ecole Polytechnique, Senior Researcher, HDR]
> Yann Ponty [CNRS, Researcher]

Faculty Members
> Philippe Chassignet [Ecole Polytechnique, Associate Professor]
> Laurent Mouchard [Université de Rouen, Associate Professor, HDR]
> Jean-Marc Steyaert [Ecole Polytechnique, Professor, HDR]

Technical Staff
> Pauline Pommeret [Inria]

PhD Students
> Alice Héliou [Ecole Polytechnique]
> Amélie Héliou [Ecole Polytechnique]
> Juraj Michalik [Inria]
> Jorgelindo Moreira Da Veiga [CIFRE Soredab]
> Afaf Saaidi [CNRS]
> Antoine Soulé [Ecole Polytechnique, until Oct 2016]
> Wei Wang [Université Paris-Sud]

Administrative Assistant
> Evelyne Rayssac [Ecole Polytechnique]

2. Overall Objectives

2.1. Overall Objectives

Our project addresses a central question in bioninformatics, namely the molecular levels of organization in the cells. The biological function of macromolecules such as proteins and nucleic acids relies on their dynamic structural nature and their ability to interact with many different partners. Therefore, folding and docking are still major issues in modern structural biology and we currently concentrate our efforts on structure and interactions and aim at a contribution to RNA design. With the recent development of computational methods aiming to integrate different levels of information, protein and nucleic acid assemblies studies should provide a better understanding on the molecular processes and machinery occurring in the cell and our research extends to several related issues incomparative genomics.

On the one hand, we study and develop methodological approaches for dealing with macromolecular structures and annotation: the challenge is to develop abstract models that are computationally tractable and biologically relevant. Our approach puts a strong emphasis on the modeling of biological objects using classic formalisms in computer science (languages, trees, graphs...), occasionally decorated and/or weighted to capture features of interest. To that purpose, we rely on the wide array of skills present in our team in the fields of combinatorics, formal languages and discrete mathematics. The resulting models are usually designed to be amenable to a probabilistic analysis, which can be used to assess the relevance of models, or test general hypotheses.

On the other hand, once suitable models are established we apply these computational approaches to several particular problems arising in fundamental molecular biology. One typically aims at designing new specialized algorithms and methods to efficiently compute properties of real biological objects. Tools of choice include exact optimization, relying heavily on dynamic programming, simulations, machine learning and discrete mathematics. As a whole, a common toolkit of computational methods is developed within the group. The trade-off between the biological accuracy of the model and the computational tractability or efficiency is to be addressed in a close partnership with experimental biology groups. One outcome is to provide software or platform elements to predict structural models and functional hypotheses.

3. Research Program

3.1. RNA and protein structures

At the secondary structure level, we contributed novel generic techniques applicable to dynamic programming and statistical sampling, and applied them to design novel efficient algorithms for probing the conformational space. Another originality of our approach is that we cover a wide range of scales for RNA structure representation. For each scale (atomic, sequence, secondary and tertiary structure...) cutting-edge algorithmic strategies and accurate and efficient tools have been developed or are under development. This offers a new view on the complexity of RNA structure and function that will certainly provide valuable insights for biological studies.

3.1.1. Dynamic programming and complexity

Participants: Yann Ponty, Wei Wang, Antoine Soulé, Juraj Michalik.

Common activity with J. Waldispühl (McGill) and A. Denise (LRI).

Ever since the seminal work of Zuker and Stiegler, the field of RNA bioinformatics has been characterized by a strong emphasis on the secondary structure. This discrete abstraction of the 3D conformation of RNA has paved the way for a development of quantitative approaches in RNA computational biology, revealing unexpected connections between combinatorics and molecular biology. Using our strong background in enumerative combinatorics, we propose generic and efficient algorithms, both for sampling and counting structures using dynamic programming. These general techniques have been applied to study the sequence-structure relationship [44], the correction of pyrosequencing errors [37], and the efficient detection of multi-stable RNAs (riboswitches) [40], [41].
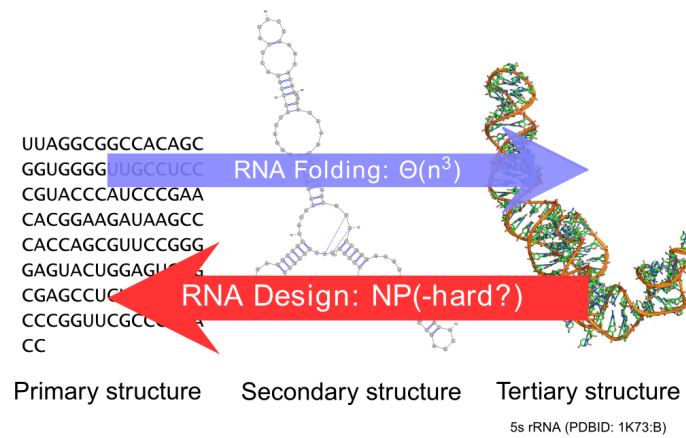
Figure 1. The goal of RNA design, aka RNA inverse folding, is to find a sequence that folds back into a given (secondary) structure.

3.1.2. *RNA design.*

Participants: Alice Héliou, Yann Ponty.

Joint project with A. Denise (sc Lri), J. Waldispühl (McGill), D. Barash (Univ. Ben-Gurion), and C. Chauve (Simon Fraser University).

It is a natural pursue to build on our understanding of the secondary structure to construct artificial RNAs performing predetermined functions, ultimately targeting therapeutic and synthetic biology applications. Towards this goal, a key element is the design of RNA sequences that fold into a predetermined secondary structure, according to established energy models (inverse-folding problem). Quite surprisingly, and despite two decades of studies of the problem, the computational complexity of the inverse-folding problem is currently unknown.

Within our group, we offer a new methodology, based on weighted random generation [24] and multidimensional Boltzmann sampling, for this problem. Initially lifting the constraint of folding back into the target structure, we explored the random generation of sequences that are compatible with the target, using a probability distribution which favors exponentially sequences of high affinity towards the target. A simple posterior rejection step selects sequences that effectively fold back into the latter, resulting in a *global sampling* pipeline that showed comparable performances to its competitors based on local search [31].

3.1.3. *Towards 3D modeling of large molecules*

Participants: Yann Ponty, Afaf Saaidi, Mireille Régnier, Amélie Héliou.

Joint projects with A. Denise (LRI), D. Barth (Versailles), J. Cohen (Paris-Sud), B. Sargueil (Paris V) and Jérome Waldispühl (McGill).

The modeling of large RNA 3D structures, that is predicting the three-dimensional structure of a given RNA sequence, relies on two complementary approaches. The approach by homology is used when the structure of a sequence homologous to the sequence of interest has already been resolved experimentally. The main problem then is to calculate an alignment between the known structure and the sequence. The ab initio approach is required when no homologous structure is known for the sequence of interest (or for some parts of it). We contribute methods inspired by both of these settings directions.

Modeling tasks can also be greatly helped by the availability of experimental data. However, high-resolution techniques such as crystallography or RMN, are notoriously costly in term of time and ressources, leading to the current gap between the amount of available sequences and structural data. As part of a colloboration with B. Sargueil's lab (Faculté de pharmacie, Paris V) funded by the Fondation pour la Recherche medical, we strive to propose a new paradigm for the analysis data produced using a new experimental technique, called SHAPE analysis (Selective 2'-Hydroxyl Acylation analyzed by Primer Extension). This experimental setup produces an accessibility profile associated with the different positions of an RNA, the *shadow* of an RNA. As part of A. Saaidi's PhD, we currently design new algorithmic strategies to infer the secondary structure of RNA from multiple SHAPE experiments performed by experimentalists at Paris V. Those are obtained on mutants, and will be coupled with a fragment-based 3D modeling strategy developed by our partners at McGill.

3.2. Séquences

Participants: Mireille Régnier, Philippe Chassignet, Yann Ponty, Jean-Marc Steyaert, Alice Héliou, Antoine Soulé.

String searching and pattern matching is a classical area in computer science, enhanced by potential applications to genomic sequences. In CPM/SPIRE community, a focus is given to general string algorithms and associated data structures with their theoretical complexity. Our group specialized in a formalization based on languages, weighted by a probabilistic model. Team members have a common expertise in enumeration and random generation of combinatorial sequences or structures, that are *admissible* according to some given constraints. A special attention is paid to the actual computability of formula or the efficiency of structures design, possibly to be reused in external software.

As a whole, motif detection in genomic sequences is a hot subject in computational biology that allows to address some key questions such as chromosome dynamics or annotation. Among specific motifs involved in molecular interactions, one may cite protein-DNA (cis-regulation), protein-protein (docking), RNA-RNA (miRNA, frameshift, circularisation). This area is being renewed by high throughput data and assembly issues. New constraints, such as energy conditions, or sequencing errors and amplification bias that are technology dependent, must be introduced in the models. A collaboration has beenestablished with LOB, at Ecole Polytechnique, who bought a sequencing machine, through the co-advised thesis of Alice Héliou. An other aim is to combine statistical sampling with a fragment based approach for decomposing structures, such as the cycle decomposition used within F. Major's group [33]. In general, in the future, our methods for sampling and sequence data analysis should be extended to take into account such constraints, that are continuously evolving.

3.2.1. *Combinatorial Algorithms and motifs*

Participants: Mireille Régnier, Philippe Chassignet, Alice Héliou.

Besides applications [39] of analytic combinatorics to computational biology problems, the team addressed general combinatorial problems on words and fundamental issues on languages and data structures.

Motif detection combines an algorithmic search of potential sites and a significance assessment. Assessment significance requires a quantitative criterion such as the p-value.

In the recent years, a general scheme of derivation of analytic formula for the pvalue under different constraints (k-occurrence, first occurrence, overrepresentation in large sequences,...) has been provided. It relies on a representation of continuous sequences of overlapping words, currently named *clumps* or *clusters* in a graph [35]. Recursive equations to compute p-values may be reduced to a traversal of that graph, leading to a linear algorithm. This improves over the space and time complexity of the generating function approach or previous probabilistic weighted automata.

This research area is widened by new problems arising from *de novo* genome assembly or re-assembly.

In [43], it is claimed that half of the genome consists of different types of repeats. One may cite microsatellites, DNA transposons, transposons, long terminal repeats (LTR), long interspersed nuclear elements (LINE), ribosomal DNA, short interspersed nuclear elements (SINE). Therefore, knowledge about the length of repeats is a key issue in several genomic problems, notably assembly or re-sequencing. Preliminary theoretical results are given in [28], and, recently, heuristics have been proposed and implemented [25], [38], [22]. A dual problem is the length of minimal absent words. Minimal absent words are words that do not occur but whose proper factors all occur in the sequence. Their computation is extremly related to finding maximal repeats (repeat that can not be extended on the right nor on the left). The comparison of the sets of minimal absent words provides a fast alternative for measuring approximation in sequence comparison [21], [23].

Recently, it was shown that considering the words which occur in one sequence but do no in another can be used to detect biologically significant events [42]. We have studied the computation of minimal absent words and we have provided new linear implementations [18],[16]. We are now working on a dynamic approach to compute minimal absent words for a sliding window. For a sequence of size n, we expect a complexity of O(n) in time and space, independent of the size of the window. This approach could be use to align a sequence on a larger sequence using minimal absent words for comparison.

According to the current knowledge, cancer develops as a result of the mutational process of the genomic DNA. In addition to point mutations, cancer genomes often accumulate a significant number of chromosomal rearrangements also called structural variants (SVs). Identifying exact positions and types of these variants may lead to track cancer development or select the most appropriate treatment for the patient. Next Generation Sequencing opens the way to the study of structural variants in the genome, as recently described in [20]. This is the subject of an international collaboration with V. Makeev's lab (IOGENE, Moscow), MAGNOME project-team and V. Boeva (Curie Institute). One goal is to combine two detection techniquesbased either on paired-end mapping abnormalities or on variation of the depth of coverage. A second goal is to develop a model of errors, including a statistical model, that takes into account the quality of data from the different sequencing technologies, their volume and their specificities such as the GC-content or the mappability.

3.2.2. *Random generation*

Participants: Yann Ponty, Juraj Michalik.

Analytical methods may fail when both sequential and structural constraints of sequences are to be modelled or, more generally, when molecular *structures* such as RNA structures have to be handled. The random generation of combinatorial objects is a natural, alternative, framework to assess the significance of observed phenomena. General and efficient techniques have been developed over the last decades to draw objects uniformly at random from an abstract specification. However, in the context of biological sequences and structures, the uniformity assumption becomes unrealistic, and one has to consider non-uniform distributions in order to derive relevant estimates. Typically, context-free grammars can handle certain kinds of long-range interactions such as base pairings in secondary RNA structures.

In 2005, a new paradigm appeared in the *ab initio* secondary structure prediction [26]: instead of formulating the problem as a classic optimization, this new approach uses statistical sampling within the space of solutions. Besides giving better, more robust, results, it allows for a fruitful adaptation of tools and algorithms derived in a purely combinatorial setting. Indeed, in a joint work with A. Denise (LRI), we have done significant and original progress in this area recently [34], [39], including combinatorial models for structures with pseudoknots. Our aim is to combine this paradigm with a fragment based approach for decomposing structures, such as the cycle decomposition used within F. Major's group [33].

3.3. 3D interaction and structure prediction

Participant: Amélie Héliou.

The biological function of macromolecules such as proteins and nucleic acids relies on their dynamic structural nature and their ability to interact with many different partners. This is specially challenging as structure flexibility is key and multi-scale modelling [19], [27] and efficient code are essential [32].

Our project covers various aspects of biological macromolecule structure and interaction modelling and analysis. First protein structure prediction is addressed through combinatorics. The dynamics of these types of structures is also studied using statistical and robotics inspired strategies. Both provide a good starting point to perform 3D interaction modelling, accurate structure and dynamics being essential.

Our group benefits from a good collaboration network, mainly at Stanford University (USA), HKUST (Hong-Kong) and McGill (Canada). The computational expertise in this field of computational structural biology is represented in a few large groups in the world (e.g. Pande lab at Stanford, Baker lab at U.Washington) that have both dry and wet labs. At Inria, our interest for structural biology is shared by the ABS and ORPAILLEUR project-teams. Our activities are however now more centered around protein-nucleic acid interactions, multi-scale analysis, robotics inspired strategies and machine learning than protein-protein interactions, algorithms and geometry. We also shared a common interest for large biomolecules and their dynamics with the NANO-D project team and their adaptative sampling strategy. As a whole, we contribute to the development of geometric and machine learning strategies for macromolecular docking.

Game theory was used by M. Boudard in her PhD thesis, defended in 2015, to predict the 3d structure of RNA. In her PhD thesis, co-advised by J. Cohen (LRI), A. Héliou is extending the approach to predict protein structures.

3.3.1. *Robotics-inspired structure and dynamics*
Participant: Amélie Héliou.

We recently work one a robotics approach to sample the conformational space of macromolecules like RNAs [1]. The robotics approach allows maintaining the secondary structure of the RNA fixed, as an unfolding is very unlikely and energetically demanding. By this approach we also dramatically reduce the number of degrees of freedom in the molecule. The conformational space becomes possible to be sampled. This reduction does not reduce the quality of the sampling.

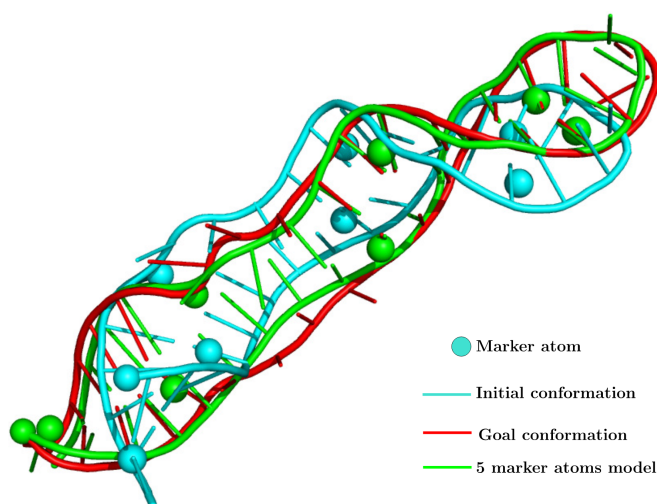

Figure 2. The cyan structure is the initial conformation, the red structure is the goal conformation. The full-atom initial conformation was driven toward the goal conformation using only the position of the goal sphere atoms. The green conformation is the result obtained; spheres perfectly overlap with the goal position and the overall conformation is really close to the goal conformation.

Our current work consists in applying the same approach to a targeted move. The motion is then driven either by the position of a few atoms or the distances between couple of atoms. Theses two aspects are under development and will increase the analysis possibility of experimental data. Our method can drive a RNA conformation toward another conformation of the same RNA given only the position of a few atoms (marker atoms).

For instance double electron-electron resonance (DEER) experimental results are distributions of distances. Probes are attached to the molecules and the distances between to probes is measured and outputted as a distribution. Our method is able to sample an ensemble of all-atom conformations that can explain the distance distribution.

3.3.2. *Game theory and protein folding*

Participant: Amélie Héliou.

M.Boudard used game theory to sample folded conformations of RNA. We work in apply game theory to sample folded conformations of proteins. This is challenging as a protein is generally less flexible than a RNA and thus accept less conformations.

Our work is first to find an algorithm that can guarantee the convergence to an Nash equilibrium (a state were no player would increase his payoff by playing something different alone) and prove their convergence. At the same time, we are looking for efficient and biologically relevant ways of defining the game settings so that Nash equilibria correspond to folded states. One direction would be to draw a parallel between Nash equilibria and local minima of the kinetic landscape.

4. New Software and Platforms

4.1. GenRGenS

GENeration of Random GENomic Sequences
KEYWORDS: Bioinformatics - Genomic sequence
FUNCTIONAL DESCRIPTION

A software dedicated to the random generation of sequences. Supports different lasses of models, including weighted context-free grammars, Markov models, ProSITE patterns..

- Participants: Yann Ponty and Alain Denise
- Contact: Yann Ponty
- URL: https://www.lri.fr/~genrgens/

4.2. VARNA

Interactive drawing and editing of the RNA secondary structure
KEYWORDS: Bioinformatics - Structural Biology
SCIENTIFIC DESCRIPTION

VARNA is Java lightweight Applet dedicated to drawing the secondary structure of RNA. It is also a Swing component that can be very easily included in an existing Java code working with RNA secondary structure to provide a fast and interactive visualization.

Being free of fancy external library dependency and/or network access, the VARNA Applet can be used as a base for a standalone applet. It looks reasonably good and scales up or down nicely to adapt to the space available on a web page, thanks to the anti-aliasing drawing primitives of Swing.
FUNCTIONAL DESCRIPTION

Varna is a new tool for the automated drawing, visualization and annotation of the secondary structure of RNA, designed as a companion software for web servers and databases.

Varna implements four drawing algorithms, supports input/output using the classic formats dbn, ct, bpseq and RNAML and exports the drawing as five picture formats, either pixel-based (JPEG, PNG) or vector-based (SVG, EPS and XFIG).

It also allows manual modification and structural annotation of the resulting drawing using either an interactive point and click approach, within a web server or through command-line arguments.

- Participants: Alain Denise and Yann Ponty
- Partners: CNRS - Ecole Polytechnique - Université Paris-Sud
- Contact: Yann Ponty
- URL: http://varna.lri.fr/

5. New Results

5.1. RNA design

We extended our previous results on RNA design [29], obtained in collaboration with J. Hales, J. Manuch and L. Stacho (Simon Fraser University/Univ. British Columbia, Canada).

Our results provided complete characterizations for the structures that can be designed using restricted alphabets. We provided a complete characterization of designable structures without unpaired bases. When unpaired bases are allowed, we provided partial characterizations for classes of designable/undesignable structures, and showed that the class of designable structures is closed under the stutter operation. Membership of a given structure to any of the classes can be tested in linear time and, for positive instances, a solution could be found in linear time. Finally, we considered a structure-approximating version of the problem that allows to extend helices and, assuming that the input structure avoids two motifs, we provided a linear-time algorithm that produces a designable structure with at most twice more base pairs than the input structure, as illustrated by Fig. 3.
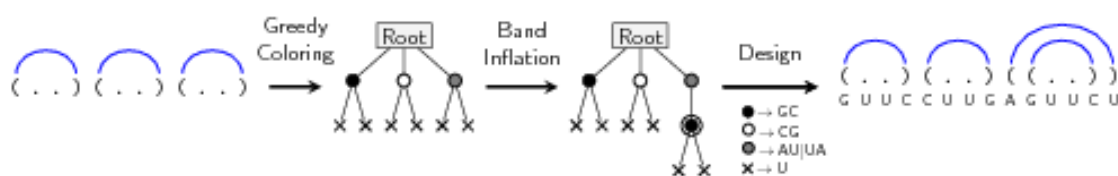
Figure 3. Principle of our structure-approximating version of RNA design: Starting from a potentially undesignable structure, a greedy coloring can be performed and corrected such that the final structure is provably designable in linear time.

In a manuscript accepted for publication in Algorithmica [4], we have shown that our previous results [29] hold for more sophisticated energy models where base-pairs are associated with arbitrary energy contributions. This result, which required a complete overhaul of our previous proofs (e.g. using arguments based on graph coloring), allows us to foresee an extension of (at least some of) our results to state-of-the-art models, such as the Turner energy model.

We also initiated a collaboration with Danny Barash's group at Ben-Gurion university (Israel). We contributed a review of existing tools and techniques for RNA design, to appear as an article within the *Briefings in Bioinformatics* series [2]. We also combined previously contributed methods for design into a new method and web-server for the design of RNAs [3]. This collaboration stemmed from the observation that `IncaRNAtion` [36], a random generation algorithm for RNA design recently developed in collaboration with Jérôme Waldispühl's group at McGill University (Montreal, Canada), produced excellent starting points (seed) for classic algorithms based on local-search. In particular, the combination of `IncaRNAtion` and `RNAfbInv` [45] was found to yield particularly promising candidates for design.

5.2. Algorithmics and combinatorics of motifs occurrences

We have developed a new algorithm to compute minimal absent words in external memory. Minimal absent words are used in sequence comparison [23] or to detect biologically significant events. For instance, it was shown that there exist three minimal words in Ebola virus genomes which are absent from human genome [42]. The identification of such specific-species sequences may prove to be useful for the development of diagnosis and therapeutics. We have already provide an $O(n)$-time and $O(n)$-space algorithm to compute minimal absent words, with an implementation that can be executed in parallel. However these implementations require a large amount of RAM, thus they cannot be used for the human genome on a desktop computer. In our new contribution we developed an external memory implementation, it can compute minimal absent words of length at most 11 for the human genome using only 1GB of RAM in less than 4 hours (manuscript submitted [16]).

Repetitive patterns in genomic sequences have a great biological significance. This is a key issue in several genomic problems as many repetitive structures can be found in genomes. One may cite microsatellites, retrotransposons, DNA transposons, long terminal repeats (LTR), long interspersed nuclear elements (LINE), ribosomal DNA, short interspersed nuclear elements (SINE). Knowledge about the length of a maximal repeat also has algorithmic implications, most notably the design of assembly algorithms that rely upon de Bruijn graphs.

Analytic combinatorics allowed us to derive formula for the expected length of repetitions in a random sequence [9]. The originality of the approach is the demonstration of a Large Deviation principle and the use of Lagrange multipliers. This allowed for a generalization of previous works on a binary alphabet. Simulations on random sequences confirmed the accuracy of our results. As an application, the sample case of Archaea genomes illustrated how biological sequences may differ from random sequences, and in turns provides tools to extract repetitive sequences.

5.3. Integrative RNA structural modeling

To circumvent expensive, low-throughput, 3D experimental techniques such as X-ray crystallography, a low resolution/high throughput technology called SHAPE is increasingly favored for structural modeling by structural biologists.

Within Afaf Saaidi's thesis, funded by the *Fondation pour la Recherche Médicale* and co-supervised by Bruno Sargueil at Faculté de Pharmacie of Université Paris V, we have developed integrative modeling strategies based on Boltzmann sampling. Preliminary results, obtained by applying these methods to model the structures of 3'UTR regions in Ebola, were presented at JOBIM 2016 [14].

Moreover, in collaboration with McGill University (Canada), we cross-examined mutate-and-map data (MaM [30]) in the light of evolutionary data. MaM data consist in the sequential SHAPE probing of a set of mutant RNAs, obtained by systematic point-wise mutations, to highlight structurally-dependent nucleotides, later to use dependent pairs as constraints in (an automated) structural modeling. We chose to adopt an alternative perspective on MaM data, and used the perturbation of the SHAPE profiles as a proxy for the structural disruption induced by a mutation. Disruptive mutations are rescued within homologs, *i.e.* compensated to re-establish the structure. However, our analysis also revealed the existence of non-structurally local (neither on the 2D or 3D levels) nucleotides which have significant mutual-information with highly disruptive positions, despite not being involved in any obvious compensatory relationship.
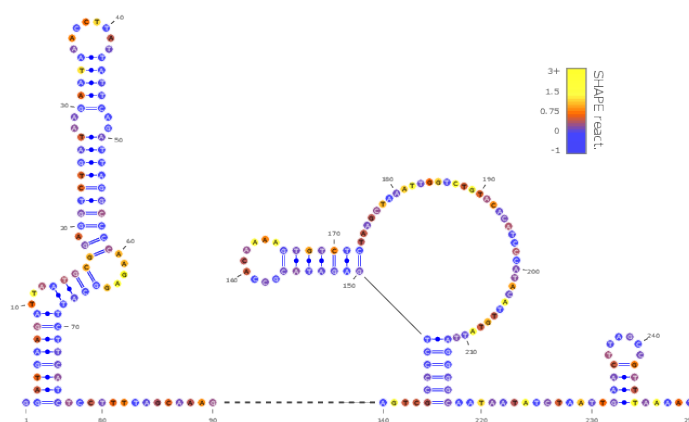
Figure 4. Conserved and thermodynamically-stable structure elements revealed by our analysis of an Ebola UTR region.

We hypothesized that such mutations are revealing of interactions involving RNA. In a manuscript published in *Nucleic Acids* journal, we tested and validated this hypothesis by showing its capacity to discriminate discriminative positions that are known to be in contact with specific ligands (proteins, DNA, small molecules...) [10].

a.   *b.*   *c.*
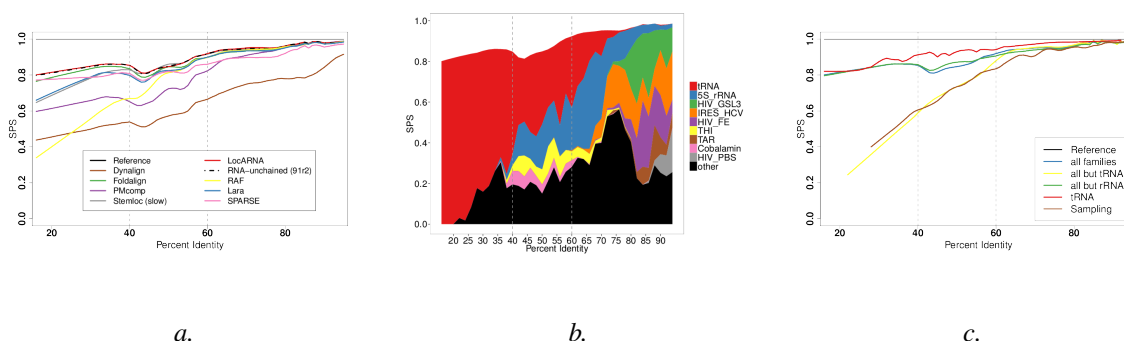
Figure 5. Typical software for comparative RNA structure prediction exhibit a dent in performance within the 40%-60% sequence identity range when benchmarked using the popular Bralibase data set (a.). However, this is due to the overrepresentation of a well-predicted type of RNA (tRNAs, red area) for low-identity ranges (b.). A re-evaluation of state-of-the-art software on an unbiased (c., brown line) reveals much more modest predictive capacities than initially believed in the community.

A fruitful line of research for RNA structure prediction is based on a comparative approach. Whenever homologous RNAs are identified, a classic strategy is to perform a simultaneous alignment and folding of several RNAs. Many software (30+) have been contributed over the past decades for this problem, leading to the introduction of benchmarks, one of the most prominent being the `Bralibase`, to position new developments and identify axes of progression. One such desired improvement, as illustrated in Figure 5, was the difficulties experienced by most software around the 40-60% sequence identity range, which was believed to arise from deep algorithmic reasons. In collaboration with Cedric Chauve (Simon Fraser University, Canada) Benedikt Löwes and Robert Giegerich (Bielefeld University, Germany), we showed that this perceived difficulty was simply

the consequence of a strong bias towards tRNAs in the 40-60% sequence identity region. Moreover, we argued that the overall performance of existing tools for low sequence identities were largely overestimated [8].

Finally, we presented at JOBIM 2016 an efficient implementation, called `LicoRNA`, of our parameterized complexity algorithm based on tree-decomposition for the sequence/structure alignment of RNA [15]. Specifically, we showed that our `LicoRNA`, by including an expressive scoring scheme and capturing pseudoknots of arbitrary complexity, generally outperforms previously contributions for the problem.

5.4. Combinatorial foundations

Pairwise ordered tree alignment are combinatorial objects that appear in RNA secondary structure comparison. However, the usual representation of tree alignments as supertrees is ambiguous, i.e. two distinct supertrees may induce identical sets of matches between identical pairs of trees. This ambiguity is uninformative, and detrimental to any probabilistic analysis. In a recent collaboration with Cédric Chauve (SFU Vancouver, Canada) and Julien Courtiel (LIPN, Paris XII) presented at the ALCOB'16 conference, we considered tree alignments up to equivalence [11]. Our first result was a precise asymptotic enumeration of tree alignments, obtained from a context-free grammar by means of basic analytic combinatorics. Our second result focused on alignments between two given ordered trees. By refining our grammar to align specific trees, we obtained a decomposition scheme for the space of alignments, and used it to design an efficient dynamic programming algorithm for sampling alignments under the Gibbs-Boltzmann probability distribution. This generalizes existing tree alignment algorithms, and opens the door for a probabilistic analysis of the space of suboptimal RNA secondary structures alignments.
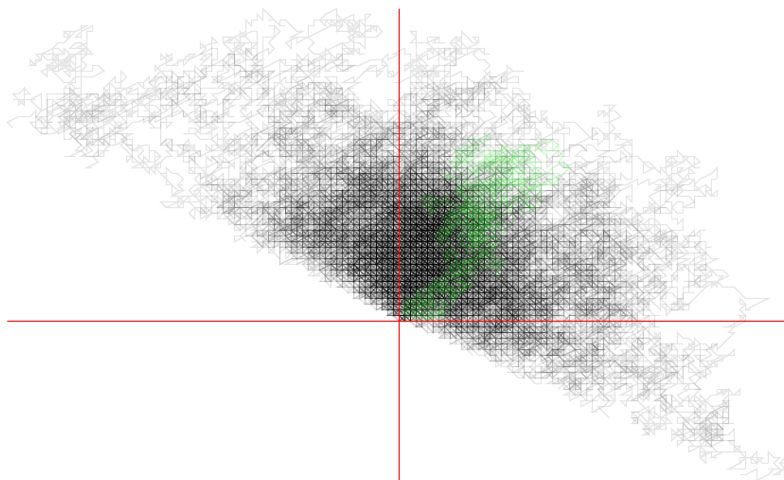
Figure 6. Random 2D walks (green walks) confined in the positive can be generated efficiently by performing rejection from a well-chosen 1D model (black walks) [12].

Finally, in collaboration with Marni Mishna (Simon Fraser University, Canada) and Jérémie Lumbroso (Princeton University, USA), we considered the uniform random generation of *difficult*, or *reluctant*, 2D discrete walks that remain confined in the positive quarter plane. We proposed a naive dynamic programming algorithm having complexity $O(n^4)$ for any step set. We also exploited the remark that any quarterplane walks can be transformed into a well-chosen 1D model having the same exponential growth factor. However, such a 1D model takes irrational steps, leading us to explore new avenues for the random generation. This work was presented at the GASCOM'16 conference [12].

5.5. Comparative genomics

D. Iakovishina defended in 2015 a PhD thesis co-advised by M. Régnier and V. Boeva (Curie Institute). She proposed a new computational method to detect structural variants using whole genome sequencing data. It combines two techniques that are based either on the detection of paired-end mapping abnormalities or on the detection of the depth of coverage. SV-Bay relies on a probabilistic Bayesian approach and includes a modelization of possible sequencing errors, read mappability profile along the genome and changes in the GC-content. Keeping only somatic SVs is an additional option when matched normal control data are provided. SV-Bay compares favorably with existing tools on simulated and experimental data sets [6] Software SV-Bay is freely available https://github.com/InstitutCurie/SV-Bay.

As a side product, a novel exhaustive catalogue of SV types -to date the most comprehensive SV classification-was built. On the grounds of previous publications and experimental data, seven new SV types, ignored by the existing SV calling algorithms, were exhibited.

We also contributed, in collaboration with Céline Scornavacca's group (ISEM, Montpellier) to the algorithmic foundations of the `EcceTERA` software [7] for the reconciliation of gene and species phylogenetic trees. This software adopts a maximum parsimony approach to predict in an evolutionary model that includes duplications, losses and transfers of genes.

6. Partnerships and Cooperations

6.1. National Initiatives

6.1.1. FRM

Yann Ponty is the Bioinformatics PI for a *Fondation de la Recherche Médicale*-funded project.

Fondation pour la Recherche Medicale – *Analyse Bio-informatique pour la recherche en Biologie* program

- Approche comparatives haut-débit pour la modelisation de l'architecture 3D des ARN à partir de données experimentales
- 2015–2018
- Yann Ponty, A. Denise, M. Regnier, A. Saaidi (PhD funded by FRM)
- B. Sargueil (Paris V – Experimental partner), J. Waldispuhl (Univ. McGill)

6.2. European Initiatives

6.2.1. *Collaborations in European Programs, Except FP7 & H2020*

Yann Ponty is the French PI for the French/Austrian RNALands project, jointly funded by the French ANR and the Austrian FWF, in partnership with the Theoretical Biochemistry Institute (University of Vienna, Austria), LRI (Univ. Paris-Sud) and EPI Bonsai (Inria Lille-Nord Europe).

French/Austrian International Program

RNALands (ANR-14-CE34-0011)

Fast and efficient sampling of structures in RNA folding landscapes

01/10/2014–30/09/2018

Coordinated by AMIB (Inria Saclay) and TBI Vienna (University of Vienna)

EPI Bonsai/Inria Lille - Nord Europe, Vienna University (Austria), LRI, Université Paris-Sud (France)

The main goal of the RNALands project is to provide efficient tools for studying the kinetics of RiboNucleic Acids, based on efficient sampling strategies.

Partenariat Henri Curien (CampusFrance, programme Staël)

Random constrained permutations: Algorithms and Analysis

01/01/2015–31/12/2016

Coordinated by CMAP (Ecole Polytechnique) and Maths Dept (Univ. Zürich, Switzerland)

LIX (Ecole Polytechnique), LIPN (Univ Paris XIII), LIGM (Univ Marne-la-Vallée)

The goals of this collaborative network is to initiate or push several collaborations related to the structure of random constrained permutations. AMIB bring their strength in random generation and discrete algorithms, and benefit from the considerable expertise accumulated at University of Zürich on permutation patterns.

6.3. International Initiatives

6.3.1. Inria International Labs

6.3.1.1. Declared Inria International Partners

Title: AMAVI - Combinatorics and Algorithms for the Genomic sequences

International Partner (Institution - Laboratory - Researcher):

Vavilov Institute of General Genetics (Russia (Russian Federation)) - Department of Computational Biology - Vsevolod Makeev

Duration: 2013 - 2017

Start year: 2013

VIGG and AMIB teams has a more than 12 years long collaboration on sequence analysis. The two groups aim at identifying DNA motifs for a functional annotation, with a special focus on conserved regulatory regions. In the current 3-years project CARNAGE, our collaboration, that includes Inria-team MAGNOME, is oriented towards new trends that arise from Next Generation Sequencing data. Combinatorial issues in genome assembly are addressed. RNA structure and interactions are also studied.

The toolkit is pattern matching algorithms and analytic combinatorics, leading to common software.

6.3.1.2. Regular International Partners

AMIB enjoys regular interactions with the following institutions:

- Simon Fraser University (Vancouver, Canada). The Mathematics department at SFU has ongoing projects on RNA design, comparative genomics and RNA structure comparison with our team. M. Mishna (SFU) will also visit Inria Saclay in January 2017 to push an oingoing collaboration on 2D walks;

- McGill University (Montréal, Canada). Following our productive collaboration with J. Waldispühl (Computer Science Dept, McGill), and the recent defense of V. Reinharz's PhD, whose thesis was co-supervised by AMIB members, we plan to increase our interactions on SHAPE data analysis by applying for an Inria associate team;

- King's college (London, UK). Our collaboration with L. Mouchard (AMIB associate) and S. Pissis on string processing and data structures is at the core of Alice Héliou's PhD. To finalize the implementation of her algorithms and apply them on real data, Alice has spent a two month period during the summer of 2016 at the EBI.

6.3.2. Participation in Other International Programs

France-Stanford exchange program

Duration: 2014 - 2016

Start year: 2014

See also: http://francestanford.stanford.edu/collaborative_projects

Amélie Héliou is co-supervised by H. Van Den Bedem in Stanford. Her tow-months visit to Stanford during the Fall of 2016 was funded by France-Stanford.

6.4. International Research Visitors

6.4.1. Visits of International Scientists

6.4.1.1. Internships

Frédéric Lavner

Date: 01/07/2016- 31/08/2016

Institution: ENSEA (France)

Supervisor: Mireille Régnier

Maria Waldl

Date: 01/08/2016 - 31/09/2016

Institution: TBI, University of Vienna (Austria)

Supervisor: Yann Ponty

6.4.2. Visits to International Teams

- Yann Ponty has visited M. Mishna and C. Chauve at the Simon Fraser University for two weeks in July 2016;
- Juraj Michalik has visited A. Tanzer and I. Hofacker at the university of Vienna (Austria) for one week in November 2016;
- Mireille Régnier has visited MIPT (Moscow) and Novossibirsk University to enhance student exchanges between these universities and Ecole polytechnique.

6.4.2.1. Research Stays Abroad

- Alice Héliou has visited the EBI (UK) for two months during the Fall of 2016;
- Amélie Héliou has visited Stanford University (USA) for two months during the Summer of 2016;

7. Dissemination

7.1. Promoting Scientific Activities

7.1.1. Scientific Events Selection

7.1.1.1. Member of the Conference Program Committees

- Yann Ponty: RECOMB'17, BICOB'17, SeqBio'16, ECCB'16, BioVis'16, ISMB'16, and BI-COB'16

7.1.1.2. Reviewer

For international conferences:

- Yann Ponty: MFCS'16

7.1.2. Journal

7.1.2.1. Member of the Editorial Boards

M. Régnier is an editor of PeerJ Computer Science.

7.1.2.2. Reviewer - Reviewing Activities

M. Régnier and Y. Ponty reviewed manuscripts for a large selection of journals in Mathematics, Computer Science and Bioinformatics: Discrete Mathematics and Theoretical Computer Science, Theoretical Computer Science, Bioinformatics, BMC Bioinformatics, Journal of Mathematical Biology, IEEE/ACM Transactions on Computational Biology and Bioinformatics, Journal of Discrete Algorithms, Algorithms for Molecular Biology, PLOS One, Journal of Theoretical Biology, RNA, Nucleic Acids Research...

7.1.3. Invited Talks

Mireille Regnier was an invited speaker of the *Advanced Algorithms on Strings* workshop in the honor of Alberto Apostolico.

Yann Ponty gave invited talks at the *Journées Combinatoires Franco-Vancouvéroises* (Vancouver, Canada), the college of life sciences (Wuhan University, China), and multiple seminars (C3BI@Pasteur, LIGM, LAMSADE, 2x Bioinfo@LIX/LRI...)

7.1.4. Leadership within the Scientific Community

Yann Ponty is the scientific animator of the *macromolecular structure and interactions axis* of the CNRS *Molecular Bioinformatics* workgroup (GdR BIM).

M. Régnier is a member of DIGITEO progam Committee and SDV working group in Saclay area.

7.1.5. Scientific Expertise

Yann Ponty acted as an external expert for the French *Agence Nationale de la Recherche* (ANR, JCJC program), and for the Canadian *Sciences and Engineering Research Council* (NSERC/CRSNG, Discovery grant program);

Yann Ponty acted as an external reviewer for the evaluation of the assistant professor position of Jing Qin at University of Southern Denmark, towards her promotion as an associate professor;

M. Régnier is a member of DIGITEO progam Committee and SDV working group in Saclay area.

7.1.6. Research Administration

Since 2016, M. Regnier acts as the head of LIX (CNRS/Ecole Polytechnique);

Until Sept. 2016, Yann Ponty was an elected member of the *comité national du CNRS*, and took part in the evaluation of CNRS research scientists and structures at a national level in Computer Science (Section 6) and Life Science interfaces (CID 51);

Yann Ponty is an elected member of the *conseil de laboratoire* of LIX.

7.2. Teaching - Supervision - Juries

7.2.1. Teaching

We have and we will go on having trained a group of good multi-disciplinary students both at the Master and PhD level. Being part of this community as a serious training group is obviously an asset. Our project is also very much involved in two major student programs in France: the Master AMI2B at Paris-Saclay (previously BIBS (Bioinformatique et Biostatistique) at Université Paris-Sud/École Polytechnique) and the parcours d'Approfondissement en Bioinformatique at École Polytechnique.

At Ecole Polytechnique, M. Régnier is in charge of M1 and M2. Most team members are teaching in this master program.

Beyond the plateau de Saclay, Yann Ponty taught 12h at the M2 level for University Pierre et Marie Curie in the BIM Master program.

7.2.2. *Supervision*

HDR

L. Mouchard *Contributions algorithmiques à l'analyse des séquences génomiques* : J.-M. Steyaert

PhD

Vladimir Reinharz, *Algorithmic properties of evolved structured RNAs*, McGill University (Montréal, Canada), July 2016, supervised by J. Waldispühl (CS Dept, McGill Univ.) and Yann Ponty

PhD in progress

Alice Héliou, *Identification et caractérisation d'ARN circulaires dans des séquences NGS*, Ecole Polytechnique, Encadrant(e|s): Mireille Régnier and Hubert Becker

Wei Wang, *Homology based approaches for predicting 3D structure of RNA molecules*, Univ. Paris XI, Encadrant(e|s): Alain Denise and Yann Ponty;

Amélie Héliou, *Game theory and conformation sampling for multi-scale and multi-body macromolecule docking*, Ecole Polytechnique, Encadrant(e|s): Johanne Cohen;

Afaf Saaidi, *Differential analysis of RNA SHAPE probing data*, Ecole Polytechnique, Encadrants: Yann Ponty and Mireille Régnier.

Antoine Soulé, *Evolutionary study of RNA-RNA interactions in yeast* , Ecole Polytechnique, Encadrants: Jean-Marc Steyaert and J. Waldispuhl (U. McGill, Canada);

Jorgelindo Moreira da Veiga, *Caractérisation dynamique et optimisation des flux métaboliques* , Ecole Polytechnique, Encadrants: L. Schwartz (AP-HP)Sabine Peres (U. Paris-Sud)

7.2.3. *Juries*

HDR

S. Bérard *Histoires évolutives et autres comptes* : M. Régnier

M. Magnin *Contributions à l'élaboration de connaissances qualitatives en bio-informatique* : M. Régnier

PhD

Manuel Lafond, Comparative Genomics, Université de Montréal, Canada, August 2016

Karen Druart, Computational Structural Biology, Ecole Polytechnique, December 2016

7.3. Popularization

Afaf Saaidi participated to the *Ma thèse en trois minute* contest, and won the best poster award at the *Journées de l'école doctoral Interfaces* of Paris Saclay University.

8. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] E. BIGAN, L. PAULEVÉ, J.-M. STEYAERT, S. S. DOUADY.*Necessary and sufficient conditions for protocell growth*, in "Journal of Mathematical Biology", April 2016 [*DOI :* 10.1007/s00285-016-0998-0], https://hal.archives-ouvertes.fr/hal-01338156.

[2] A. CHURKIN, M. DRORY RETWITZER, V. REINHARZ, Y. PONTY, J. WALDISPÜHL, D. BARASH.*Design of RNAs: Comparing Programs for inverse RNA folding*, in "Briefings in Bioinformatics", 2017, forthcoming, https://hal.inria.fr/hal-01392958.

[3] M. DRORY RETWITZER, V. REINHARZ, Y. PONTY, J. WALDISPÜHL, D. BARASH.*incaRNAfbinv : a web server for the fragment-based design of RNA sequences*, in "Nucleic Acids Research", 2016, vol. 44, n⁰ W1, p. W308 - W314 [*DOI :* 10.1093/NAR/GKW440], https://hal.inria.fr/hal-01319682.

[4] J. HALEŠ, A. HÉLIOU, J. MAŇUCH, Y. PONTY, L. STACHO.*Combinatorial RNA Design: Designability and Structure-Approximating Algorithm in Watson-Crick and Nussinov-Jacobson Energy Models*, in "Algorithmica", 2016, forthcoming, https://hal.inria.fr/hal-01285499.

[5] A. HELIOU, M. LÉONARD, L. MOUCHARD, M. SALSON.*Efficient dynamic range minimum query*, in "Theoretical Computer Science", 2017 [*DOI :* 10.1016/J.TCS.2016.07.002], https://hal.archives-ouvertes.fr/hal-01255499.

[6] D. IAKOVISHINA, I. JANOUEIX-LEROSEY, E. BARILLOT, M. REGNIER, V. BOEVA.*SV-Bay: structural variant detection in cancer genomes using a Bayesian approach with correction for GC-content and read map-pability*, in "Bioinformatics (Oxford, England)", January 2016, https://hal.inria.fr/hal-01253126.

[7] E. JACOX, C. CHAUVE, G. J. SZÖLLŐSI, Y. PONTY, C. SCORNAVACCA.*ecceTERA: comprehensive gene tree-species tree reconciliation using parsimony*, in "Bioinformatics (Oxford, England)", July 2016, p. 2056-8, Accepted, forthcoming, https://hal.inria.fr/hal-01276903.

[8] B. LÖWES, C. CHAUVE, Y. PONTY, R. GIEGERICH.*The BRaliBase dent-a tale of benchmark design and interpretation.* , in "Briefings in Bioinformatics", March 2016 [*DOI :* 10.1093/BIB/BBW022], https://hal.inria.fr/hal-01273406.

[9] M. REGNIER, P. CHASSIGNET. *Accurate prediction of the statistics of repetitions in random sequences: a case study in Archaea genomes*, in "Frontiers in Bioengineering and Biotechnology", May 2016, https://hal.inria.fr/hal-01304366.

[10] V. REINHARZ, Y. PONTY, J. WALDISPÜHL.*Combining structure probing data on RNA mutants with evolutionary information reveals RNA-binding interfaces*, in "Nucleic Acids Research", 2016, vol. 44, n⁰ 11, e104 - e104 [*DOI :* 10.1093/NAR/GKW217], https://hal.inria.fr/hal-01291754.

International Conferences with Proceedings

[11] C. CHAUVE, J. COURTIEL, Y. PONTY.*Counting, generating and sampling tree alignments*, in "ALCOB - 3rd International Conference on Algorithms for Computational Biology - 2016", Trujillo, Spain, June 2016, https://hal.inria.fr/hal-01154030.

[12] J. LUMBROSO, M. MISHNA, Y. PONTY.*Taming reluctant random walks in the positive quadrant*, in "GASCOM - 10th conference on random generation of combinatorial structures - 2016", Bastia, France, Electronic Notes in Discrete Mathematics, June 2016, forthcoming, https://hal.inria.fr/hal-01291164.

National Conferences with Proceeding

[13] V. J. HENRY, A. FERRÉ, C. FROIDEVAUX, A. A. GOELZER, V. V. FROMION, S. COHEN-BOULAKIA, S. S. DEROZIER, M. DINH, G. FIÉVET, S. FISCHER, J.-F. J.-F. GIBRAT, V. L. LOUX, S. PÉRÈS.*Représentation*

systémique multi-échelle des processus biologiques de la bactérie, in "IC2016: Ingénierie des Connaissances", Montpellier, France, June 2016, https://hal.archives-ouvertes.fr/hal-01442727.

[14] A. SAAIDI, D. ALLOUCHE, B. SARGUEIL, Y. PONTY.*Towards structural models for the Ebola UTR regions using experimental SHAPE probing data*, in "JOBIM - Journées Ouvertes en Biologie, Informatique et Mathématiques - 2016", Lyon, France, June 2016, https://hal.inria.fr/hal-01332469.

[15] W. WANG, M. BARBA, P. RINAUDO, A. DENISE, Y. PONTY.*Homology modeling of complex structural RNAs*, in "JOBIM - Journées Ouvertes en Biologie, Informatique et Mathématiques - 2016", Lyon, France, June 2016, https://hal.inria.fr/hal-01332642.

Other Publications

[16] C. BARTON, A. HELIOU, L. MOUCHARD, S. P. PISSIS.*Parallelising the Computation of Minimal Absent Words*, January 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01255489.

[17] D. SURUJON, Y. PONTY, P. CLOTE.*Small-world networks and RNA secondary structures*, January 2017, working paper or preprint, https://hal.inria.fr/hal-01424452.

References in notes

[18] C. BARTON, A. HELIOU, L. MOUCHARD, S. PISSIS.*Linear-time computation of minimal absent words using suffix array*, in "BMC Bioinformatics", 2014, vol. 15, 11 [*DOI :* 10.1186/S12859-014-0388-9], https://hal.inria.fr/hal-01110274.

[19] J. BERNAUER, S. C. FLORES, X. HUANG, S. SHIN, R. ZHOU.*Multi-Scale Modelling of Biosystems: from Molecular to Mesocale - Session Introduction*, in "Pacific Symposium on Biocomputing", 2011, p. 177-80 [*DOI :* 10.1142/9789814335058_0019], http://hal.inria.fr/inria-00542791.

[20] V. BOEVA, T. POPOVA, K. BLEAKLEY, P. CHICHE, J. CAPPO, G. SCHLEIERMACHER, I. JANOUEIX-LEROSEY, O. DELATTRE, E. BARILLOT.*Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data*, in "Bioinformatics", 2012, vol. 28, n° 3, p. 423-425, http://dx.doi.org/10.1093/bioinformatics/btr670.

[21] S. CHAIRUNGSEE, M. CROCHEMORE.*Using minimal absent words to build phylogeny*, in "Theoretical Computer Science", 2012, vol. 450, n° 0, p. 109-116.

[22] R. CHIKHI, P. MEDVEDEV.*Informed and automated k-mer size selection for genome assembly*, in "Bioinformatics", Jan 2014, vol. 30, n° 1, p. 31–37, http://dx.doi.org/10.1093/bioinformatics/btt310.

[23] M. CROCHEMORE, G. FICI, R. MERCAS, S. PISSIS.*Linear-Time Sequence Comparison Using Minimal Absent Words*, in "LATIN 2016: Theoretical Informatics - 12th Latin American Symposium", Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2016, http://arxiv.org/abs/1506.04917.

[24] A. DENISE, Y. PONTY, M. TERMIER.*Controlled non uniform random generation of decomposable structures*, in "Theoretical Computer Science", 2010, vol. 411, n° 40-42, p. 3527-3552 [*DOI :* 10.1016/J.TCS.2010.05.010], http://hal.inria.fr/hal-00483581.

[25] H. DEVILLERS, S. SCHBATH.*Separating significant matches from spurious matches in DNA sequences*, in "Journal of Computational Biology", 2012, vol. 19, n⁰ 1, p. 1–12, http://dx.doi.org/10.1089/cmb.2011.0070.

[26] Y. DING, C. CHAN, C. LAWRENCE.*RNA secondary structure prediction by centroids in a Boltzmann weighted ensemble*, in "RNA", 2005, vol. 11, p. 1157–1166.

[27] S. C. FLORES, J. BERNAUER, S. SHIN, R. ZHOU, X. HUANG.*Multiscale modeling of macromolecular biosystems*, in "Briefings in Bioinformatics", July 2012, vol. 13, n⁰ 4, p. 395-405 [*DOI :* 10.1093/BIB/BBR077], http://hal.inria.fr/hal-00684530.

[28] Z. GU, H. WANG, A. NEKRUTENKO, W. H. LI.*Densities, length proportions, and other distributional features of repetitive sequences in the human genome estimated from 430 megabases of genomic sequence*, in "Gene", Dec 2000, vol. 259, n⁰ 1-2, p. 81–88.

[29] J. HALEŠ, J. MAŇUCH, Y. PONTY, L. STACHO.*Combinatorial RNA Design: Designability and Structure-Approximating Algorithm*, in "Annual Symposium on Combinatorial Pattern Matching", Springer, 2015, p. 231–246.

[30] W. KLADWANG, R. DAS.*A mutate-and-map strategy for inferring base pairs in structured nucleic acids: proof of concept on a DNA/RNA helix*, in "Biochemistry", 2010, vol. 49, n⁰ 35, p. 7414–7416.

[31] A. LEVIN, M. LIS, Y. PONTY, C. W. O'DONNELL, S. DEVADAS, B. BERGER, J. WALDISPÜHL.*A global sampling approach to designing and reengineering RNA secondary structures*, in "Nucleic Acids Research", November 2012, vol. 40, n⁰ 20, p. 10041-52 [*DOI :* 10.1093/NAR/GKS768], http://hal.inria.fr/hal-00733924.

[32] S. LORIOT, F. CAZALS, J. BERNAUER.*ESBTL: efficient PDB parser and data structure for the structural and geometric analysis of biological macromolecules*, in "Bioinformatics", April 2010, vol. 26, n⁰ 8, p. 1127-8 [*DOI :* 10.1093/BIOINFORMATICS/BTQ083], http://hal.inria.fr/inria-00536404.

[33] M. PARISIEN, F. MAJOR.*The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data*, in "Nature", 2008, vol. 452, n⁰ 7183, p. 51–55.

[34] Y. PONTY.*Efficient sampling of RNA secondary structures from the Boltzmann ensemble of low-energy: The boustrophedon method*, in "Journal of Mathematical Biology", Jan 2008, vol. 56, n⁰ 1-2, p. 107–127, http://www.lri.fr/~ponty/docs/Ponty-07-JMB-Boustrophedon.pdf.

[35] M. REGNIER, E. FURLETOVA, M. ROYTBERG, V. YAKOVLEV.*Pattern occurrences Pvalues, Hidden Markov Models and Overlap Graphs*, 2013, submitted, http://hal.inria.fr/hal-00858701.

[36] V. REINHARZ, Y. PONTY, J. WALDISPÜHL.*A weighted sampling algorithm for the design of RNA sequences with targeted secondary structure and nucleotides distribution*, in "ISMB/ECCB - 21st Annual international conference on Intelligent Systems for Molecular Biology/12th European Conference on Computational Biology - 2013", Berlin, Allemagne, 2013, http://hal.inria.fr/hal-00811607.

[37] V. REINHARZ, Y. PONTY, J. WALDISPÜHL.*Using Structural and Evolutionary Information to Detect and Correct Pyrosequencing Errors in Noncoding RNAs*, in "Journal of Computational Biology", November 2013, vol. 20, n⁰ 11, p. 905-19, Extended version of RECOMB'13 [*DOI :* 10.1089/CMB.2013.0085], http://hal.inria.fr/hal-00828062.

[38] G. RIZK, D. LAVENIER, R. CHIKHI.*DSK: k-mer counting with very low memory usage*, in "Bioinformatics", Mar 2013, vol. 29, n⁰ 5, p. 652–653 [*DOI : 10.1093/BIOINFORMATICS/BTT020*], http://bioinformatics.oxfordjournals.org/content/early/2013/02/01/bioinformatics.btt020.full.

[39] C. SAULE, M. REGNIER, J.-M. STEYAERT, A. DENISE.*Counting RNA pseudoknotted structures*, in "Journal of Computational Biology", October 2011, vol. 18, n⁰ 10, p. 1339-1351 [*DOI : 10.1089/CMB.2010.0086*], http://hal.inria.fr/inria-00537117.

[40] E. SENTER, S. SHEIKH, I. DOTU, Y. PONTY, P. CLOTE.*Using the Fast Fourier Transform to Accelerate the Computational Search for RNA Conformational Switches*, in "PLoS ONE", December 2012, vol. 7, n⁰ 12 [*DOI : 10.1371/JOURNAL.PONE.0050506*], http://hal.inria.fr/hal-00769740.

[41] E. SENTER, S. SHEIKH, I. DOTU, Y. PONTY, P. CLOTE.*Using the Fast Fourier Transform to accelerate the computational search for RNA conformational switches (extended abstract)*, in "RECOMB - 17th Annual International Conference on Research in Computational Molecular Biology - 2013", Beijing, Chine, 2013, http://hal.inria.fr/hal-00766780.

[42] R. M. SILVA, D. PRATAS, L. CASTRO, A. J. PINHO, P. J. S. G. FERREIRA.*Three minimal sequences found in Ebola virus genomes and absent from human DNA*, in "Bioinformatics", 2015 [*DOI : 10.1093/BIOINFORMATICS/BTV189*].

[43] T. J. TREANGEN, S. L. SALZBERG.*Repetitive DNA and next-generation sequencing: computational challenges and solutions*, in "Nat Rev Genet", Jan 2012, vol. 13, n⁰ 1, p. 36–46, http://dx.doi.org/10.1038/nrg3117.

[44] J. WALDISPÜHL, Y. PONTY.*An unbiased adaptive sampling algorithm for the exploration of RNA mutational landscapes under evolutionary pressure*, in "Journal of Computational Biology", November 2011, vol. 18, n⁰ 11, p. 1465-79 [*DOI : 10.1089/CMB.2011.0181*], http://hal.inria.fr/hal-00681928.

[45] L. WEINBRAND, A. AVIHOO, D. BARASH.*RNAfbinv: an interactive Java application for fragment-based design of RNA sequences*, in "Bioinformatics", 2013, btt494.

Project-Team AVIZ

Analysis and Visualization

RESEARCH CENTER
Saclay - Île-de-France

THEME
Interaction and visualization

Table of contents

Project-Team AVIZ

Creation of the Team: 2007 February 08, updated into Project-Team: 2008 January 01

Keywords:

Computer Science and Digital Science:

3.1.4. - Uncertain data
3.1.7. - Open data
3.3. - Data and knowledge analysis
3.3.1. - On-line analytical processing
3.5.1. - Analysis of large graphs
5.1. - Human-Computer Interaction
5.2. - Data visualization

Other Research Topics and Application Domains:

1.1. - Biology
1.3. - Neuroscience and cognitive science
9.4.5. - Data science
9.5. - Humanities
9.5.1. - Psychology
9.5.3. - Economy, Finance
9.5.6. - Archeology, History
9.5.10. - Digital humanities

1. Members

Research Scientists
Jean-Daniel Fekete [Team leader, Inria, Senior Researcher, HDR]
Petra Isenberg [Inria, Senior Researcher]
Pierre Dragicevic [Inria, Researcher]
Tobias Isenberg [Inria, Researcher]

Technical Staff
Romain Di Vozzo [Inria]
Christoph Kinkeldey [Inria, Post-Doc, from May 2016]

PhD Students
Sriram Karthik Badam [Inria, from Jun 2016 until Sep 2016]
Lonni Besancon [Univ. Paris XI]
Evanthia Dimara [Inria, granted by Fondation Cooper. Scient. Campus Paris Saclay-DIGITEO]
Pascal Goffin [Inria, until Sep 2016, granted by ANR FITOC project]
Mathieu Le Goc [Inria]
Paola Tatiana Llerena Valdivia [Univ. of São Paulo, from Jul 2016]
Bart Postma [Inria, until Jun 2016]
Yanhong Wu [Hong-Kong Univ. of Science and Technology, from May 2016]

Administrative Assistant
Katia Evrat [Inria]

Others

Frédéric Vernier [Univ. Paris-Sud, LIMSI, Associate Professor]
Evelyne Lutton [INRA, Research Scientist]
Andre Spritzer [INRA, Post-doc, until Oct 2016]

2. Overall Objectives

2.1. Objectives

Aviz (Analysis and VIsualiZation) is a multidisciplinary project that seeks to improve visual exploration and analysis of large, complex datasets by tightly integrating analysis methods with interactive visualization.

Our work has the potential to affect practically all human activities for and during which data is collected and managed and subsequently needs to be understood. Often data-related activities are characterized by access to new data for which we have little or no prior knowledge of its inner structure and content. In these cases, we need to interactively *explore* the data first to gain insights and eventually be able to act upon the data contents. Interactive visual analysis is particularly useful in these cases where automatic analysis approaches fail and human capabilities need to be exploited and augmented.

Within this research scope Aviz focuses on five research themes:

- Methods to visualize and smoothly navigate through large datasets;
- Efficient analysis methods to reduce huge datasets to visualizable size;
- Visualization interaction using novel capabilities and modalities;
- Evaluation methods to assess the effectiveness of visualization and analysis methods and their usability;
- Engineering tools for building visual analytics systems that can access, search, visualize and analyze large datasets with smooth, interactive response.

2.2. Research Themes

Aviz's research on Visual Analytics is organized around five main Research Themes:

Methods to visualize and smoothly navigate through large data sets: Large data sets challenge current visualization and analysis methods. Understanding the structure of a graph with one million vertices is not just a matter of displaying the vertices on a screen and connecting them with lines. Current screens only have around two million pixels. Understanding a large graph requires both data reduction to visualize the whole and navigation techniques coupled with suitable representations to see the details. These representations, aggregation functions, navigation and interaction techniques must be chosen as a coordinated whole to be effective and fit the user's mental map.

Aviz designs new visualization representations and interactions to efficiently navigate and manipulate large data sets.

Efficient analysis methods to reduce huge data sets to visualizable size: Designing analysis components with interaction in mind has strong implications for both the algorithms and the processes they use. Some data reduction algorithms are suited to the principle of sampling, then extrapolating, assessing the quality and incrementally enhancing the computation: for example, all the linear reductions such as PCA, Factorial Analysis, and SVM, as well as general MDS and Self Organizing Maps. Aviz investigates the possible analysis processes according to the analyzed data types.

Visualization interaction using novel capabilities and modalities: The importance of interaction to Visualization and, in particular, to the interplay between interactivity and cognition is widely recognized. However, information visualization interactions have yet to take full advantage of these new possibilities in interaction technologies, as they largely still employ the traditional desktop, mouse, and keyboard setup of WIMP (Windows, Icons, Menus, and a Pointer) interfaces. At Aviz we investigate in particular interaction through tangible and touch-based interfaces to data.

Evaluation methods to assess their effectiveness and usability: For several reasons appropriate evaluation of visual analytics solutions is not trivial. First, visual analytics tools are often designed to be applicable to a variety of disciplines, for various different data sources, and data characteristics, and because of this variety it is hard to make general statements. Second, in visual analytics the specificity of humans, their work environment, and the data analysis tasks, form a multi-faceted evaluation context which is difficult to control and generalize. This means that recommendations for visual analytics solutions are never absolute, but depend on their context.

In our work we systematically connect evaluation approaches to visual analytics research—we strive to develop and use both novel as well as establish mixed-methods evaluation approaches to derive recommendations on the use of visual analytics tools and techniques. Aviz regularly published user studies of visual analytics and interaction techniques and takes part in dedicated workshops on evaluation.

Engineering tools: for building visual analytics systems that can access, search, visualize and analyze large data sets with smooth, interactive response.

Currently, databases, data analysis and visualization all use the concept of data tables made of tuples and linked by relations. However, databases are storage-oriented and do not describe the data types precisely. Analytical systems describe the data types precisely, but their data storage and computation model are not suited to interactive visualization. Visualization systems use in-memory data tables tailored for fast display and filtering, but their interactions with external analysis programs and databases are often slow.

Aviz seeks to merge three fields: databases, data analysis and visualization. Part of this merging involves using common abstractions and interoperable components. This is a long-term challenge, but it is a necessity because generic, loosely-coupled combinations will not achieve interactive performance.

Aviz's approach is holistic: these five themes are facets of building an analysis process optimized for discovery. All the systems and techniques Aviz designs support the process of understanding data and forming insights while minimizing disruptions during navigation and interaction.

3. Research Program

3.1. Scientific Foundations

The scientific foundations of Visual Analytics lie primarily in the domains of Visualization and Data Mining. Indirectly, it inherits from other established domains such as graphic design, Exploratory Data Analysis (EDA), statistics, Artificial Intelligence (AI), Human-Computer Interaction (HCI), and Psychology.

The use of graphic representation to understand abstract data is a goal Visual Analytics shares with Tukey's Exploratory Data Analysis (EDA) [58], graphic designers such as Bertin [46] and Tufte [57], and HCI researchers in the field of Information Visualization [44].

EDA is complementary to classical statistical analysis. Classical statistics starts from a *problem*, gathers *data*, designs a *model* and performs an *analysis* to reach a *conclusion* about whether the data follows the model. While EDA also starts with a problem and data, it is most useful *before* we have a model; rather, we perform visual analysis to discover what kind of model might apply to it. However, statistical validation is not always required with EDA; since often the results of visual analysis are sufficiently clear-cut that statistics are unnecessary.

Visual Analytics relies on a process similar to EDA, but expands its scope to include more sophisticated graphics and areas where considerable automated analysis is required before the visual analysis takes place. This richer data analysis has its roots in the domain of Data Mining, while the advanced graphics and interactive exploration techniques come from the scientific fields of Data Visualization and HCI, as well as the expertise of professions such as cartography and graphic designers who have long worked to create effective methods for graphically conveying information.

The books of the cartographer Bertin and the graphic designer Tufte are full of rules drawn from their experience about how the meaning of data can be best conveyed visually. Their purpose is to find effective visual representation that describe a data set but also (mainly for Bertin) to discover structure in the data by using the right mappings from abstract dimensions in the data to visual ones.

For the last 25 years, the field of Human-Computer Interaction (HCI) has also shown that interacting with visual representations of data in a tight perception-action loop improves the time and level of understanding of data sets. Information Visualization is the branch of HCI that has studied visual representations suitable to understanding and interaction methods suitable to navigating and drilling down on data. The scientific foundations of Information Visualization come from theories about perception, action and interaction.

Several theories of perception are related to information visualization such as the "Gestalt" principles, Gibson's theory of visual perception [51] and Triesman's "preattentive processing" theory [56]. We use them extensively but they only have a limited accuracy for predicting the effectiveness of novel visual representations in interactive settings.

Information Visualization emerged from HCI when researchers realized that interaction greatly enhanced the perception of visual representations.

To be effective, interaction should take place in an interactive loop faster than 100ms. For small data sets, it is not difficult to guarantee that analysis, visualization and interaction steps occur in this time, permitting smooth data analysis and navigation. For larger data sets, more computation should be performed to reduce the data size to a size that may be visualized effectively.

In 2002, we showed that the practical limit of InfoVis was on the order of 1 million items displayed on a screen [49]. Although screen technologies have improved rapidly since then, eventually we will be limited by the physiology of our vision system: about 20 millions receptor cells (rods and cones) on the retina. Another problem will be the limits of human visual attention, as suggested by our 2006 study on change blindness in large and multiple displays [47]. Therefore, visualization alone cannot let us understand very large data sets. Other techniques such as aggregation or sampling must be used to reduce the visual complexity of the data to the scale of human perception.

Abstracting data to reduce its size to what humans can understand is the goal of Data Mining research. It uses data analysis and machine learning techniques. The scientific foundations of these techniques revolve around the idea of finding a good model for the data. Unfortunately, the more sophisticated techniques for finding models are complex, and the algorithms can take a long time to run, making them unsuitable for an interactive environment. Furthermore, some models are too complex for humans to understand; so the results of data mining can be difficult or impossible to understand directly.

Unlike pure Data Mining systems, a Visual Analytics system provides analysis algorithms and processes compatible with human perception and understandable to human cognition. The analysis should provide understandable results quickly, even if they are not ideal. Instead of running to a predefined threshold, algorithms and programs should be designed to allow trading speed for quality and show the tradeoffs interactively. This is not a temporary requirement: it will be with us even when computers are much faster, because good quality algorithms are at least quadratic in time (e.g. hierarchical clustering methods). Visual Analytics systems need different algorithms for different phases of the work that can trade speed for quality in an understandable way.

Designing novel interaction and visualization techniques to explore huge data sets is an important goal and requires solving hard problems, but how can we assess whether or not our techniques and systems provide real improvements? Without this answer, we cannot know if we are heading in the right direction. This is why we have been actively involved in the design of evaluation methods for information visualization [55], [54], [52], [53], [50]. For more complex systems, other methods are required. For these we want to focus on longitudinal evaluation methods while still trying to improve controlled experiments.
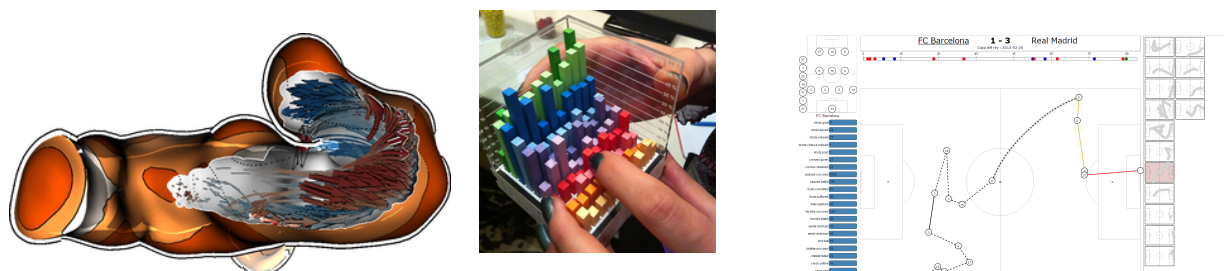
3.2. Innovation

Figure 1. Example novel visualization techniques and tools developed by the team. Left: a non-photorealistic rendering technique that visualizes blood flow and vessel thickness. Middle:a physical visualization showing economic indicators for several countries, right: SoccerStories a tool for visualizing soccer games.

We design novel visualization and interaction techniques (see, for example, Figure 1). Many of these techniques are also evaluated throughout the course of their respective research projects. We cover application domains such as sports analysis, digital humanities, fluid simulations, and biology. A focus of Aviz' work is the improvement of graph visualization and interaction with graphs. We further develop individual techniques for the design of tabular visualizations and different types of data charts. Another focus is the use of animation as a transition aid between different views of the data. We are also interested in applying techniques from illustrative visualization to visual representations and applications in information visualization as well as scientific visualization.

3.3. Evaluation Methods

Evaluation methods are required to assess the effectiveness and usability of visualization and analysis methods. Aviz typically uses traditional HCI evaluation methods, either quantitative (measuring speed and errors) or qualitative (understanding users tasks and activities). Moreover, Aviz is also contributing to the improvement of evaluation methods by reporting on the best practices in the field, by co-organizing workshops (BELIV 2010, 2012, 2014, 2016) to exchange on novel evaluation methods, by improving our ways of reporting, interpreting and communicating statistical results, and by applying novel methodologies, for example to assess visualization literacy.

3.4. Software Infrastructures

We want to understand the requirements that software and hardware architectures should provide to support exploratory analysis of large amounts of data. So far, "big data" has been focusing on issues related to storage management and predictive analysis: applying a well-known set of operations on large amounts of data. Visual Analytics is about exploration of data, with sometimes little knowledge of its structure or properties. Therefore, interactive exploration and analysis is needed to build knowledge and apply appropriate analyses; this knowledge and appropriateness is supported by visualizations. However, applying analytical operations on large data implies long-lasting computations, incompatible with interactions, and generates large amounts of results, impossible to visualize directly without aggregation or sampling. Visual Analytics has started to tackle these problems for specific applications but not in a general manner, leading to fragmentation of results and difficulties to reuse techniques from one application to the other. We are interested in abstracting-out the issues and finding general architectural models, patterns, and frameworks to address the Visual Analytics challenge in more generic ways.

3.5. Emerging Technologies

Figure 2. Example emerging technology solutions developed by the team for multi-display environments, wall displays, and token-based visualization.

We want to empower humans to make use of data using different types of display media and to enhance how they can understand and visually and interactively explore information. This includes novel display equipment and accompanying input techniques. The Aviz team specifically focuses on the exploration of the use of large displays in visualization contexts as well as emerging physical and tangible visualizations. In terms of interaction modalities our work focuses on using touch and tangible interaction. Aviz participates to the Digiscope project that funds 11 wall-size displays at multiple places in the Paris area (see http://www.digiscope.fr), connected by telepresence equipment and a Fablab for creating devices. Aviz is in charge of creating and managing the Fablab, uses it to create physical visualizations, and is also using the local wall-size display (called WILD) to explore visualization on large screens. The team also investigates the perceptual, motor and cognitive implications of using such technologies for visualization.

3.6. Psychology

More cross-fertilization is needed between psychology and information visualization. The only key difference lies in their ultimate objective: understanding the human mind vs. helping to develop better tools. We focus on understanding and using findings from psychology to inform new tools for information visualization. In many cases, our work also extends previous work in psychology. Our approach to the psychology of information visualization is largely holistic and helps bridge gaps between perception, action and cognition in the context of information visualization. Our focus includes the perception of charts in general, perception in large display environments, collaboration, perception of animations, how action can support perception and cognition, and judgment under uncertainty.

4. Application Domains

4.1. Domains

Research in visual analytics can profit from the challenges and requirements of real-world datasets. Aviz develops active collaboration with users from a range of application domains, making sure it can support their specific needs. By studying similar problems in different domains, we can begin to generalize our results and have confidence that our solutions will work for a variety of applications.

We apply our techniques to important medical applications domains such as bioinformatics and brain studies. In particular, we are interested in helping neuroscientists make sense of evolving functional networks, in the form of weighted and/or dynamic graphs.

Other application domains include:

- Digital Humanities in general, with the Cendari European project with historians from most European countries, the project "Interactive Network Visualization" with Microsoft Research-Inria Joint Centre on Graph Visualization, and with our work on Word-Scale Visualizations;
- Many traditional scientific research fields such as astronomy, fluid dynamics, structural biology, and neurosciences;
- Scientific illustration that can benefit from illustrative visualization techniques for scientific data;
- Personal visualization and visual analytics in which we develop solutions for the general audience.

5. Highlights of the Year

5.1. Highlights of the Year

We had a number of highlights this year:

- Aviz researchers contributed 35 publications this year. Amongst these 6 papers were presented at IEEE VIS, the largest international Visualizations and Visual Analytics conference. One full paper was presented at UIST, one the most prestigious international conference on human computer interaction;
- Aviz researchers organized two workshops at international conferences (IEEE VIS);
- Three awards were won by Aviz researchers for papers (see below);
- We welcomed four international students to our lab for research visits;
- Aviz researchers taught four lectures at various French and international universities.

5.1.1. Awards

[8]

E. DIMARA, P. DRAGICEVIC, A. BEZERIANOS.*The Attraction Effect in Information Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", 2017, vol. 23, no 1 [DOI : 10.1109/TVCG.2016.2598594], https://hal.inria.fr/hal-01355750

BEST PAPERS AWARDS :

[24] **Proceedings of the Symposium on User Interface Software and Technology (UIST)**. M. LE GOC, L. H. KIM, A. PARSAEI, J.-D. FEKETE, P. DRAGICEVIC, S. FOLLMER.

[25] **IEEE VIS 2016**. S. K. BADAM, C. KINKELDEY, P. ISENBERG.

6. New Software and Platforms

6.1. Zooids

Participants: Mathieu Le Goc [correspondant], Lawrence Kim, Ali Parsaei, Jean-Daniel Fekete, Pierre Dragicevic, Sean Follmer.

Zooids are autonomous robots that handle both display and interaction. We distribute them as an open-source open-hardware platform for developing tabletop swarm interfaces [24]. All information, related content and material can be found at http://www.aviz.fr/swarmui.

6.2. Reorder.js

Participant: Jean-Daniel Fekete [correspondant].

Visualizing data tables and graph/network can be done using a matrix visualization. Jacques Bertin, the French cartographer and visualization pioneer explained in his book "Semiology of Graphics" that, to make sense of a matrix, it should first be correctly ordered. This is what the `Reorder.js` library is doing.

Figure 3. Zooids can be held as tokens, manipulated collectively or individually, behave as physical pixels, act as handles and controllers, and can move dynamically under machine control. They are building blocks for a new class of user interface we call swarm user interfaces.

Ordering is also useful for other purposes. For example, if you want to visualize with Parallel Coordinates, you should provide an order for the dimension axes. Reorder.js can be used to find a suitable order. See also the poster paper [48].

The library also provide examples of visualizations using reordering; they are based on the d3.js library. For more information, see our survey of methods for matrix reordering [3].
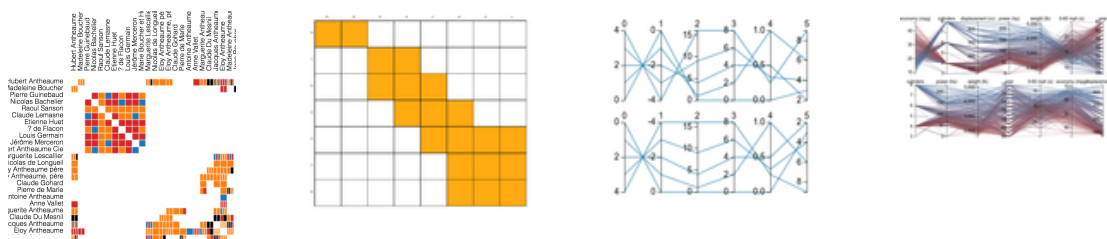
Figure 4. Correctly ordered matrices of a network (left), and parallel coordinate plots with dimensions ordered according to their correlation (right)

6.3. NetworkCube

Participants: Jean-Daniel Fekete [correspondant], Nathalie Henry-Riche, Benjamin Bach.

Network visualizations support research in a range of scientific domains from biology to humanities. NetworkCube is a platform to bridge the gap between domain scientists and visualisation researchers; NetworkCube aims in being a fast way to deploy experimental visualizations from research to domain experts analyzing dynamic networks. In turn, InfoVis researchers benefit from studying how their visualizations are used in the wild [45].

NetworkCube is made of three parts: a core, the Vistorian which is specialized for visualizations for historians, and the Connectoscope which is specialized for Brain Researchers. NetworkCube provides multiple representations for dynamic networks, allowing complex explorations from Web clients.

6.4. Vispubdata

Participants: Petra Isenberg [correspondant], Florian Heimerl, Steffen Koch, Tobias Isenberg, Panpan Xu, Charles Stolper, Michael Sedlmair, Torsten Möller, John Stasko.
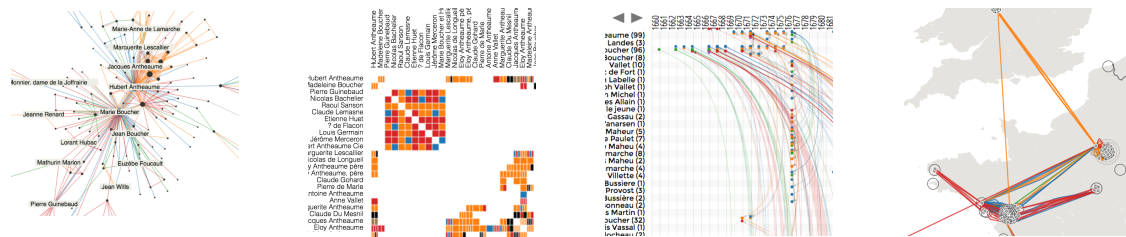
Figure 5. Four different visualization techniques to explore dynamic networks provided by NetworkCube
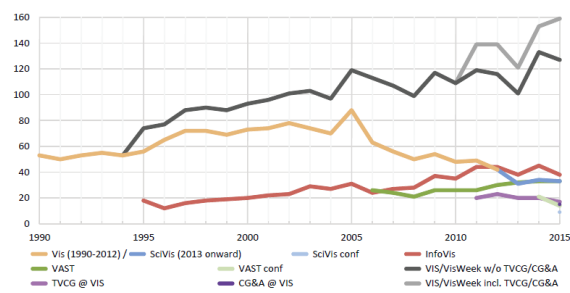
Figure 6. Overview of the files included in the dataset.

We have created and made available to all a dataset with information about every paper that has appeared at the IEEE Visualization (VIS) set of conferences: InfoVis, SciVis, VAST, and Vis. The information about each paper includes its title, abstract, authors, and citations to other papers in the conference series, among many other attributes. This data is meant to be useful to the broad data visualization community to help understand the evolution of the field and as an example document collection for text data visualization research.

6.5. Time Curves

Participants: Benjamin Bach, Pierre Dragicevic [correspondant], Conglei Shi, Nicolas Heulot.

Figure 7. Wikipedia article on abortion illustrated with time curves.

We introduced *time curves* as a general approach for visualizing patterns of evolution in temporal data [2]. Examples of such patterns include slow and regular progressions, large sudden changes, and reversals to previous states. These patterns can be of interest in a range of domains, such as collaborative document editing, dynamic network analysis, and video analysis. Time curves employ the metaphor of folding a timeline visualization into itself so as to bring similar time points close to each other. This metaphor can be applied to any dataset where a similarity metric between temporal snapshots can be defined, thus it is largely datatype-agnostic. In our paper and on the online Website, we illustrate how time curves can visually reveal informative patterns in a range of different datasets.

More on the project Web page: www.aviz.fr/ bbach/timecurves.

6.6. CENDARI Note-Taking-Environment

SCIENTIFIC DESCRIPTION

CENDARI (http://www.aviz.fr/Research/CENDARI) Is a European Infrastructure project funded by the EU for 4 years: 2012-2016. Aviz is in charge of the Human-Computer Interface for the project, and develops a tool to allow historians and archivists to take notes, enter them online, manage their images in relations with the notes and documents, and visualize the entities they find in the documents and notes. This system is an extension of the original EditorsNotes project, integrating several innovative components asked by the historians: visualizations, relations with the Semantic Web, and a management of access rights respecting the researchers' desire of privacy for their notes, as well as desire of sharing entities and relations gathered throught the notes and documents.

FUNCTIONAL DESCRIPTION

Editors' Notes is an open-source, web-based tool for recording, organizing, preserving, and opening access to research notes, built with the needs of documentary editing projects, archives, and library special collections in mind.

- Participants: Evanthia Dimara, Nadia Boukhelifa Sari Ali and Jean-Daniel Fekete
- Contact: Jean-Daniel Fekete
- URL: https://github.com/CENDARI/editorsnotes

6.7. Hybrid Image Visualisation
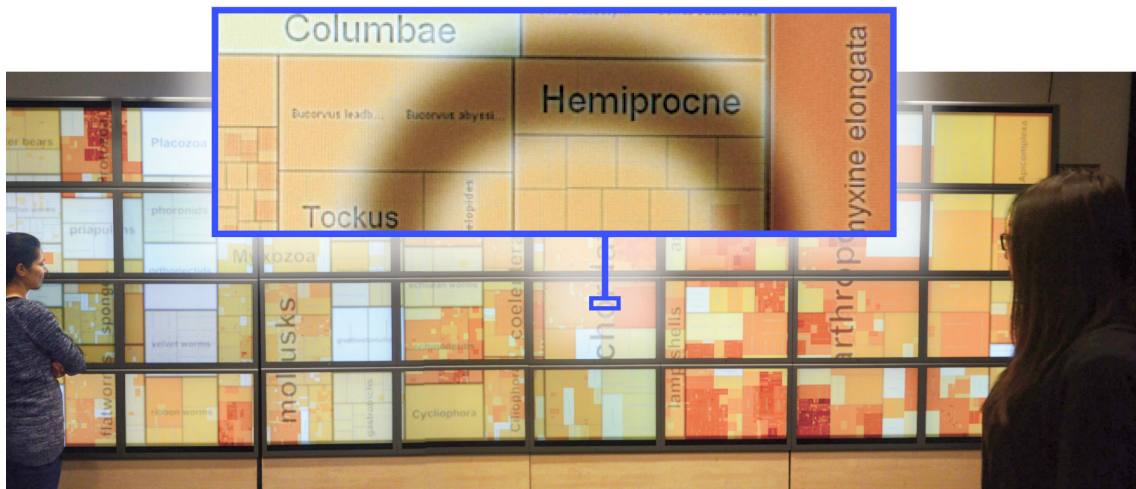
Figure 8. Hybrid image visualization software for the creation of visualizations for distant and close viewing on large displays.

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 is 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. We contribute a design space, discuss the perceptual rationale for our work, provide examples and a set of techniques for hybrid-image visualizations, and describe tools for designing hybrid-image visualizations. An example can be found in Figure 8.

KEYWORDS: Wall-Sized Displays, Perception, Hybrid Images

FUNCTIONAL DESCRIPTION

- Participants: Jean-Daniel Fekete, Petra Isenberg, Pierre Dragicevic, Wesley Willett, Romain Primet.
- Contact: Petra Isenberg
- URL: http://aviz.fr/Research/HybridImageVisualizations

6.8. Sparklificator

KEYWORDS: Information visualization - Data visualization - Visualization

SCIENTIFIC DESCRIPTION

Sparklificator is a general open-source jQuery library that eases the process of integrating word-scale visualizations into HTML documents.

FUNCTIONAL DESCRIPTION

Sparklificator's name comes from adding sparklines to a textual document. It is a general open-source jQuery library that eases the process of integrating wordscale visualizations into HTML documents. Sparklificator provides a range of options for adjusting the position (on top, to the right, as an overlay), size, and spacing of vizualisations within the text. The library includes default visualizations, including small line charts and bar charts, and can also be used to integrate custom word-scale visualizations created using webbased visualization toolkits such as D3.

- Participants: Pascal Goffin, Wesley Willett and Petra Isenberg
- Contact: Jean-Daniel Fekete
- URL: http://inria.github.io/sparklificator/

7. New Results

7.1. Swarm User Interfaces

Participants: Mathieu Le Goc [correspondant], Lawrence Kim, Ali Parsaei, Jean-Daniel Fekete, Pierre Dragicevic, Sean Follmer.

We introduce swarm user interfaces (Fig 3), a new class of human-computer interfaces comprised of many autonomous robots that handle both display and interaction. We describe the design of Zooids, a hardware and software system: a small wheel-propelled robot with position and touch sensing capabilities that can be freely arranged and repositioned on any horizontal surface, both through user manipulation and computer control. Zooids is an open-source open-hardware platform for developing tabletop swarm interfaces. We illustrate the potential of tabletop swarm user interfaces through a set of application scenarios developed with Zooids, and discuss general design considerations unique to swarm user interfaces.

More on the project Web page: http://www.aviz.fr/swarmui.

7.2. A Systematic Review of Experimental Studies on Data Glyphs

Participants: Johannes Fuchs, Petra Isenberg [correspondant], Anastasia Bezerianos, Daniel Keim.
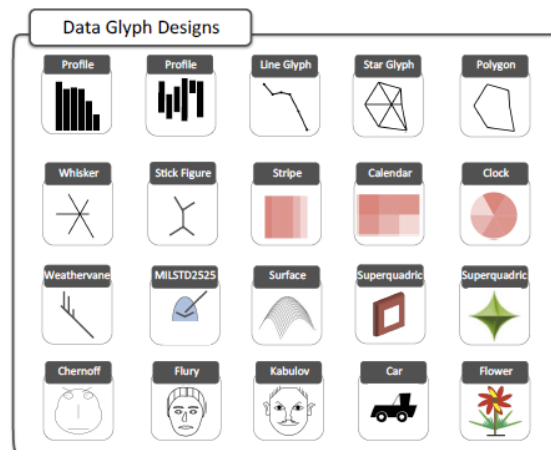

Figure 9. Overview of the glyphs reviewed in the study.

We systematically reviewed 64 user-study papers on data glyphs to help researchers and practitioners gain an informed understanding of tradeoffs in the glyph design space. The glyphs we considered were individual representations of multi-dimensional data points, often meant to be shown in small-multiple settings. Over the past 60 years many different glyph designs were proposed and many of these designs have been subjected to perceptual or comparative evaluations. Yet, a systematic overview of the types of glyphs and design variations tested, the tasks under which they were analyzed, or even the study goals and results did not yet exist. We provide such an overview by systematically sampling and tabulating the literature on data glyph studies, listing their designs, questions, data, and tasks. In addition we present a concise overview of the types of glyphs and their design characteristics analyzed by researchers in the past, and a synthesis of the study results. Based on our meta analysis of all results we further contribute a set of design implications and a discussion on open research directions.

7.3. Towards an Understanding of Mobile Touch Navigation in a Stereoscopic Viewing Environment for 3D Data Exploration

Participants: David López, Lora Oehlberg, Candemir Doger, Tobias Isenberg [correspondant].
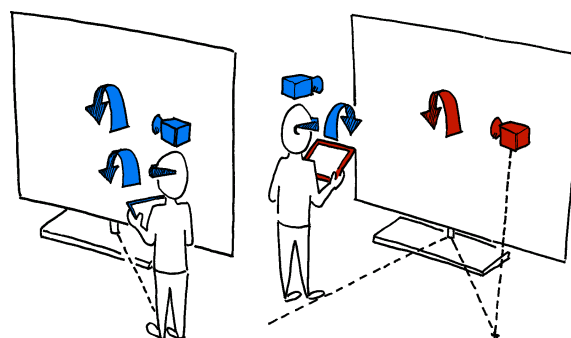
Figure 10. Illustration of the problem of mobility within a virtual environment, while interacting with a view on a tablet.

We discuss touch-based navigation of 3D visualizations in a combined monoscopic and stereoscopic viewing environment. We identify a set of interaction modes, and a workflow that helps users transition between these modes to improve their interaction experience. In our discussion we analyze, in particular, the control-display space mapping between the different reference frames of the stereoscopic and monoscopic displays. We show how this mapping supports interactive data exploration, but may also lead to conflicts between the stereoscopic and monoscopic views due to users' movement in space; we resolve these problems through synchronization. To support our discussion, we present results from an exploratory observational evaluation with domain experts in fluid mechanics and structural biology. These experts explored domain-specific datasets using variations of a system that embodies the interaction modes and workflows; we report on their interactions and qualitative feedback on the system and its workflow.

More on the project Web page: https://tobias.isenberg.cc/VideosAndDemos/Lopez2016TUM.

7.4. CAST: Effective and Efficient User Interaction for Context-Aware Selection in 3D Particle Clouds

Participants: Lingyun Yu, Konstantinos Efstathiou, Petra Isenberg, Tobias Isenberg [correspondant].
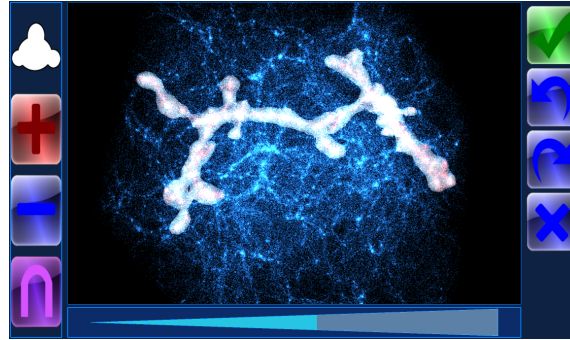
Figure 11. Illustration of the complexity metrics used in the study.

We present a family of three interactive Context-Aware Selection Techniques (CAST) for the analysis of large 3D particle datasets. For these datasets, spatial selection is an essential prerequisite to many other analysis tasks. Traditionally, such interactive target selection has been particularly challenging when the data subsets of interest were implicitly defined in the form of complicated structures of thousands of particles. Our new techniques SpaceCast, TraceCast, and PointCast improve usability and speed of spatial selection in point clouds through novel context-aware algorithms. They are able to infer a user's subtle selection intention from gestural input, can deal with complex situations such as partially occluded point clusters or multiple cluster layers, and can all be fine-tuned after the selection interaction has been completed. Together, they provide an effective and efficient tool set for the fast exploratory analysis of large datasets. In addition to presenting Cast, we report on a formal user study that compares our new techniques not only to each other but also to existing state-of-the-art selection methods. Our results show that Cast family members are virtually always faster than existing methods without tradeoffs in accuracy. In addition, qualitative feedback shows that PointCast and TraceCast were strongly favored by our participants for intuitiveness and efficiency.

More on the project Web page: https://tobias.isenberg.cc/VideosAndDemos/Yu2016CEE.

7.5. Hybrid Tactile/Tangible Interaction for 3D Data Exploration

Participants: Lonni Besançon [correspondant], Paul Issartel, Mehdi Ammi, Tobias Isenberg.
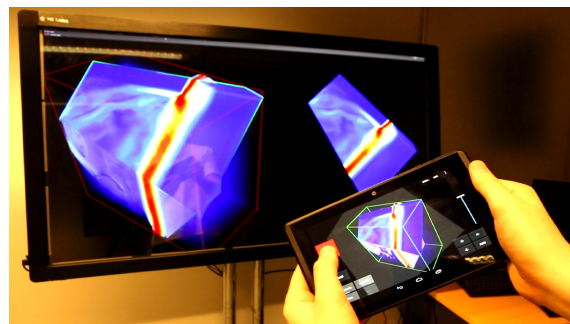
Figure 12. Picture of the hybrid interaction system.

We present the design and evaluation of an interface that combines tactile and tangible paradigms for 3D visualization. While studies have demonstrated that both tactile and tangible input can be efficient for a subset of 3D manipulation tasks, we reflect here on the possibility to combine the two complementary input types. Based on a field study and follow-up interviews, we present a conceptual framework of the use of these different interaction modalities for visualization both separately and combined—focusing on free exploration as well as precise control. We present a prototypical application of a subset of these combined mappings for fluid dynamics data visualization using a portable, position-aware device which offers both tactile input and tangible sensing. We evaluate our approach with domain experts and report on their qualitative feedback.

More on the project Web page: http://lonni.besancon.pagesperso-orange.fr/Projects/HybridInteraction/HybridInteraction.html.

7.6. A Tangible Volume for Portable 3D Interaction

Participants: Paul Issartel, Lonni Besançon [correspondant], Tobias Isenberg, Mehdi Ammi.
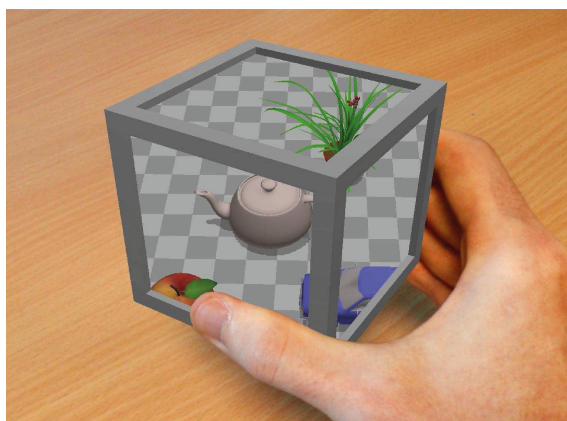
Figure 13. Image of the Cube.

We present a new approach to achieve tangible object manipulation with a single, fully portable and self-contained device. Our solution is based on the concept of a "tangible volume". We turn a tangible object into a handheld fish-tank display. The tangible volume represents a volume of space that can be freely manipulated within a virtual scene. This volume can be positioned onto virtual objects to directly grasp them, and to manipulate them in 3D space. We investigate this concept through two user studies. The first study evaluates the intuitiveness of using a tangible volume for grasping and manipulating virtual objects. The second study evaluates the effects of the limited field of view on spatial awareness. Finally, we present a generalization of this concept to other forms of interaction through the surface of the volume.

More on the project Web page: http://lonni.besancon.pagesperso-orange.fr/Projects/TangibleCube/TangibleCube.html.

7.7. Preference Between Allocentric and Egocentric 3D Manipulation in a Locally Coupled Configuration

Participants: Paul Issartel, Lonni Besançon [correspondant], Steven Franconeri.

We study user preference between allocentric and egocentric 3D manipulation on mobile devices, in a configuration where the motion of the device is applied to an object displayed on the device itself. We first evaluate this preference for translations and for rotations alone, then for full 6-DOF manipulation. We also investigate the role of contextual cues by performing this experiment in different 3D scenes. Finally, we look at the specific influence of each manipulation axis. Our results provide guidelines to help interface designers select an appropriate default mapping in this locally coupled configuration.

More on the project Web page: http://lonni.besancon.pagesperso-orange.fr/Projects/Mappings/Mappings.html.

7.8. Embedded Data Representations

Participants: Wesley Willett, Yvonne Jansen, Pierre Dragicevic [correspondant].
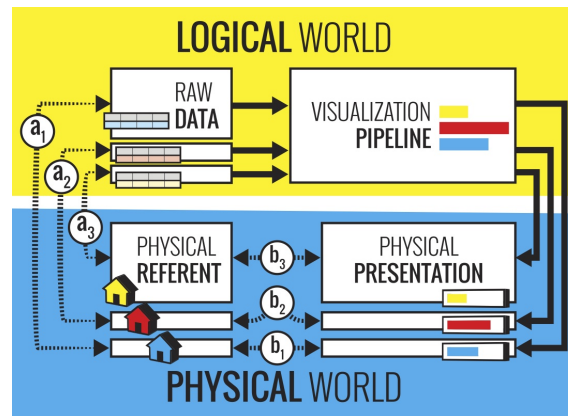
Figure 14. Conceptual model for situated and embedded data representations.

We introduces *embedded data representations* as the use of visual and physical representations of data that are deeply integrated with the physical spaces, objects, and entities to which the data refers [16]. Technologies like lightweight wireless displays, mixed reality hardware, and autonomous vehicles are making it increasingly easier to display data in-context. While researchers and artists have already begun to create embedded data representations, the benefits, trade-offs, and even the language necessary to describe and compare these approaches remain unexplored.

In this paper, we formalize the notion of physical data referents – the real-world entities and spaces to which data corresponds – and examine the relationship between referents and the visual and physical representations of their data. We differentiate situated representations, which display data in proximity to data referents, and embedded representations, which display data so that it spatially coincides with data referents. Drawing on examples from visualization, ubiquitous computing, and art, we explore the role of spatial indirection, scale, and interaction for embedded representations. We also examine the tradeoffs between non-situated, situated, and embedded data displays, including both visualizations and physicalizations. Based on our observations, we identify a variety of design challenges for embedded data representation, and suggest opportunities for future research and applications

7.9. Space-Time Cube Framework

Participants: Benjamin Bach, Pierre Dragicevic [correspondant], Dominique Archambault, Christophe Hurter, Sheelagh Carpendale.

We presented a descriptive model for visualizations of temporal data based on a generalized space-time cube framework [1]. Visualizations are described as operations on a conceptual space-time cube, which transform the cube's 3D shape into readable 2D visualizations. Operations include: extracting subparts of the cube, flattening it across space or time, or transforming the cube's geometry and content. We introduced a taxonomy of elementary space-time cube operations and explained how these operations can be combined and parameterized.

The generalized space-time cube has two properties: a) it is purely conceptual without the need to be implemented, and b) it applies to all datasets that can be represented in two dimensions plus time (e.g., geospatial, videos, networks, multivariate data). The proper choice of space-time cube operations depends on many factors, e.g., density or sparsity of a cube, hence we proposed a characterization of structures within space-time cubes, which allowed us to discuss strengths and limitations of operations. We also reviewed interactive systems that support multiple operations, allowing a user to customize his view on the data. With this framework, we hope to facilitate the description, criticism and comparison of temporal data visualizations, as well as encourage the exploration of new techniques and systems.

More on the project Web page: spacetimecubevis.com.

7.10. The Attraction Effect in Information Visualization

Participants: Evanthia Dimara [correspondant], Anastasia Bezerianos, Pierre Dragicevic.
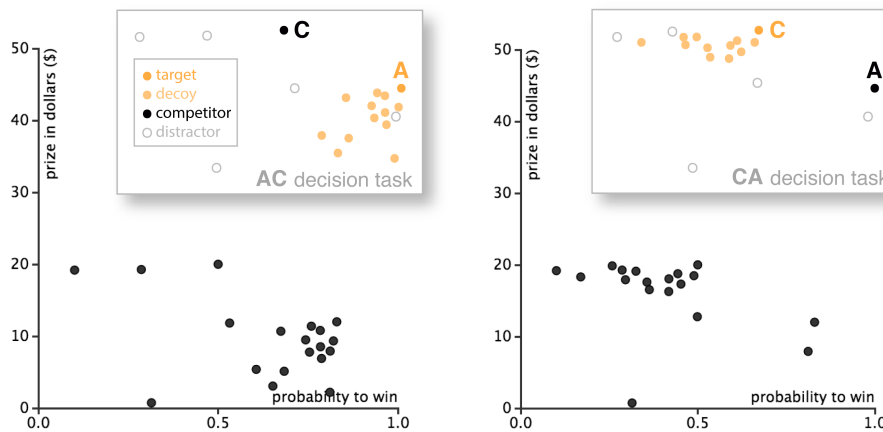
Figure 15. Illustration of the extension of the attraction effect in large datasets. Example of two matched decision tasks AC and CA in scatterplots.

The attraction effect is a well-studied cognitive bias in decision making research, where one's choice between two alternatives is influenced by the presence of an irrelevant (dominated) third alternative. We examine whether this cognitive bias, so far only tested with three alternatives and simple presentation formats such as numerical tables, text and pictures, also appears in visualizations. Since visualizations can be used to support decision making — e.g., when choosing a house to buy or an employee to hire — a systematic bias could have important implications. In a first crowdsource experiment, we indeed partially replicated the attraction effect with three alternatives presented as a numerical table, and observed similar effects when they were presented as a scatterplot. In a second experiment, we investigated if the effect extends to larger sets of alternatives, where the number of alternatives is too large for numerical tables to be practical. Our findings indicate that the bias persists for larger sets of alternatives presented as scatterplots. We discuss implications for future research on how to further study and possibly alleviate the attraction effect.

More on the project Web page: www.aviz.fr/decoy.

7.11. An Exploratory Study of Word-Scale Graphics in Data-Rich Text Documents

Participants: Pascal Goffin [correspondant], Jeremy Boy, Wesley Willett, Petra Isenberg.

Figure 16. Examples of word-scale graphics collected during the study.

We investigated the design and function of word-scale graphics and visualizations embedded in text documents. Word-scale graphics include both data-driven representations such as word-scale visualizations and sparklines, and non-data-driven visual marks. There has been little research attention on their design, function, and use so far. We present the results of an open ended exploratory study with nine graphic designers. The study resulted in a rich collection of different types of graphics, data provenance, and relationships between text, graphics, and data. Based on this corpus, we present a systematic overview of word-scale graphic designs, and examine how designers used them. We also discuss the designers' goals in creating their graphics, and characterize how they used word-scale graphics to visualize data, add emphasis, and create alternative narratives. We discuss implications for the design of authoring tools for word-scale graphics and visualizations building on these examples, and explore how new authoring environments could make it easier for designers to integrate them into documents.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR FITOC: From Individual To Collaborative Visual Analytics

Participants: Petra Isenberg [correspondant], Jean-Daniel Fekete, Pierre Dragicevic, Pascal Goffin, Wesley Willett.

The goal of this project is to help bringing collaboration to existing individual visual data analysis work. It is situated in the domain of information visualization, a subdomain of computer science, but views and tries to support data analysis as a social process. The work is motivated by the fact that a large amount of data analysis work is conducted by individuals in isolated tool, such as Excel, R, SPSS, Tableau, and others. Synthesis and sharing of the results then happens in another set of tools such as notes, email, or office documents. The research is situated in the domain of visualization which has a long tradition of building

tools and techniques for individual data analysis. Currently there are technological innovations under way to help people analyze data together, but there is still a disconnect between the two modes of data analysis (collaborative and individual). In this project, we want to find ways in which information can best be used and shared visually while transitioning between individual and collaborative data analysis activities.

The project ended in July, 2016.

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. CENDARI

Title: Collaborative EuropeaN Digital/Archival Infrastructure
Programm: FP7
Duration: February 2012 - January 2016
Coordinator: Trinity College - Dublin
Partners:

Consortium of European Research Libraries (United Kingdom)
Koninklijke Bibliotheek (Netherlands)
Fondazione Ezio Franceschini Onlus (Italy)
Freie Universitaet Berlin (Germany)
King's College London (United Kingdom)
"matematicki Institutnu, Beograd" (Serbia)
Narodni Knihovna Ceske Republiky (Czech Republic)
Societa Internazionale Per Lo Studio Del Medioevo Latino-S.I.S.M.E.L.Associazione (Italy)
The Provost, Fellows, Foundation Scholars & The Other Members of Board of The College of The Holy & Undivided Trinity of Queen Elizabeth Near Dublin (Ireland)
Georg-August-Universitaet Goettingen Stiftung Oeffentlichen Rechts (Germany)
The University of Birmingham (United Kingdom)
Universitaet Stuttgart (Germany)
Universita Degli Studi di Cassino E Del Lazio Meridionale (Italy)

Inria contact: L. Romary

'The Collaborative EuropeaN Digital Archive Infrastructure (CENDARI) will provide and facilitate access to existing archives and resources in Europe for the study of medieval and modern European history through the development of an 'enquiry environment'. This environment will increase access to records of historic importance across the European Research Area, creating a powerful new platform for accessing and investigating historical data in a transnational fashion overcoming the national and institutional data silos that now exist. It will leverage the power of the European infrastructure for Digital Humanities (DARIAH) bringing these technical experts together with leading historians and existing research infrastructures (archives, libraries and individual digital projects) within a programme of technical research informed by cutting edge reflection on the impact of the digital age on scholarly practice. The enquiry environment that is at the heart of this proposal will create new ways to discover meaning, a methodology not just of scale but of kind. It will create tools and workspaces that allow researchers to engage with large data sets via federated multilingual searches across heterogeneous resources while defining workflows enabling the creation of personalized research environments, shared research and teaching spaces, and annotation trails, amongst other features. This will be facilitated by multilingual authority lists of named entities (people, places, events) that will harness user involvement to add intelligence to the system. Moreover, it will develop new visual paradigms for the exploration of patterns generated by the system, from knowledge transfer and dissemination, to language usage and shifts, to the advancement and diffusion of ideas.'

8.3. International Initiatives

8.3.1. Inria International Partners

8.3.1.1. Informal International Partners

- Univ. of Konstanz, Jean-Daniel Fekete collaborates with Michael Behrischon network exploration based on matrices [4], [3].

- NYU, Jean-Daniel Fekete collaborates with Enrico Bertini and his students on multidimensional visualization and exploration [23]/

- Microsoft Research Redmond, Jean-Daniel Fekete collaborates with Nathalie Henry-Riche on the visualization of dynamic networks (see 6.3).

- Stanford University. Mathieu Le Goc, Jean-Daniel Fekete and Pierre Dragicevic collaborate with Sean Follmer on Swarm User Interfaces and the design of the Zooids (section 7.1).

- Univ of Calgary. Pierre Dragicevic collaborates with Wesley Willett on situated data visualization.

- Univ of Washington, Univ Chicago and Univ Zurich. Pierre Dragicevic collaborates with Matthew Kay, Steve Haroz and Chat Wacharamanotham on transparent statistical reporting.

- Microsoft Research, Redmond, University of Waterloo, University of Calgary. Petra Isenberg and Tobias Isenberg collaborate with Bongshin Lee, Mark Hancock, Diane Watson, and Sheelagh Carpendale on touch vs. mouse interaction

- Microsoft Research, Redmond. Petra Isenberg collaborates with Bongshin Lee on mobile visualization research.

- Univ. of Vienna, Austria: Petra Isenberg and Tobias Isenberg collaborate with Torsten Möller and Michael Sedlmair on visualization practices and evaluation of visualization.

- Univ. of Maryland, Baltimore County, USA: Petra Isenberg and Tobias Isenberg collaborate with Jian Chen on visualization practices and evaluation of visualization.

- Georgia Tech, USA: Petra Isenberg and Tobias Isenberg collaborate with John Stasko on visualization practices.

- Univ. Groningen, the Netherlands: Petra Isenberg and Tobias Isenberg collaborate with Lingyun Yu and Konstantinos Efstathiou on context-aware 3D selection.

- Univ. of Granada, Spain: Tobias Isenberg collaborates with Domingo Martín on non-photorealistic rendering.

- Techn. Univ. of Vienna, Austria: Tobias Isenberg collaborates with Ivan Viola on illustrative visualization.

- Univ. of Bergen, Norway: Tobias Isenberg collaborates with Stefan Bruckner on interactive visualization.

- Univ. of Ulm, Germany: Tobias Isenberg collaborates with Timo Ropinski on interactive visualization.

- Worms Univ. of Appl. Sciences, Germany: Tobias Isenberg collaborates with Alexander Wiebel on interactive visualization.

- Univ. Koblenz-Landau, Germany: Tobias Isenberg collaborates with Kai Lawonn on illustrative visualization.

- Univ. Magdeburg, Germany: Tobias Isenberg collaborates with Bernhard Preim on illustrative visualization.

8.4. International Research Visitors

8.4.1. Visits of International Scientists

8.4.1.1. Internships

- Sriram Karthik Badam, PhD, Univ. of Maryland, from Jun 2016 until Sep 2016
- Yanhong Wu, PhD, Hong-Kong Univ. of Science and Technology, from May to Sep. 2016

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the Organizing Committees

- Jean-Daniel Fekete co-organizing the LIVVIL 2016 workshop on "Logging Visualization and Visualizing Logs" at IEEE VIS.
- Jean-Daniel Fekete served on the best paper committee for EuroVis 2016
- Pierre Dragicevic was Doctoral Colloquium co-chair at IEEE VIS 2016.
- Petra Isenberg co-organized the Beliv 2016 workshop on "Beyond Time and Error: Novel Evaluation Techniques for Visualization" at IEEE VIS.
- Petra Isenberg served as Posters co-chair for VIS 2016
- Petra Isenberg served on the best short-paper committee for EuroVis 2016
- Tobias Isenberg was posters co-chair for EuroVis 2016.
- Tobias Isenberg was visualization area co-chair for ISVC 2016.

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

- Jean-Daniel Fekete was a member of the program committees for Graph Drawing 2016
- Jean-Daniel Fekete was a member of the program committees for EuroVis 2016
- Jean-Daniel Fekete was a member of the program committees for ACM AVI 2016
- Pierre Dragicevic was a member of the program committee for VIS 2016.
- Pierre Dragicevic was a member of the program committee for CHI 2017.
- Petra Isenberg was a member of the program committee for CHI 2016.
- Petra Isenberg was a member of the program committee for Beliv 2016.
- Petra Isenberg was a member of the program committee for EuroVA 2016.
- Petra Isenberg was a member of the program committee for InfoVis 2016.
- Tobias Isenberg was a member of the program committee for IEEE SciVis 2016.
- Tobias Isenberg was a member of the program committee for IEEE VISAP 2016.
- Tobias Isenberg was a member of the program committee for Expresssive 2016.
- Tobias Isenberg was a member of the program committee for IEEE 3DUI 2016.
- Pascal Goffin was a member of the program committee for 9. Forum Medientechnik 2016.

9.1.2.2. Reviewer

- Jean-Daniel Fekete: VIS, CHI
- Pierre Dragicevic: VIS, UIST, IHM
- Petra Isenberg: EuroVA, EuroVis, InfoVis, Nordic Forum for Human-Computer Interaction Research (NordiCHI), UIST Beliv (Workshop), DFG (Funding Agency)
- Tobias Isenberg: 3DUI, CHI, EuroVis, Expressive, ISS, PacificVis, SciVis, VISAP
- Christoph Kinkeldey: CHI

- Lonni Besancon: CHI EA, ISMI, SUI, CHI
- Mathieu Le Goc: CHI

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Tobias Isenberg is associate editor of Elsevier Computers & Graphics.
- Petra Isenberg and Tobias Isenberg were guest editors of a special issue of Sage Publishing's Information Visualization on visualization evaluation.

9.1.3.2. Reviewer - Reviewing Activities

- Petra Isenberg: Big Data Journal,CG&A, Journal of Graph Algorithms and Applications (JGAA), Information Visualization Journal (IVS), TVCG
- Tobias Isenberg: C&G, IV, TVCG
- Christoph Kinkeldey: JoEP (Journal of Economic Psychology), IJGI (ISPRS International Journal of Geo-Information), CP (Cartographic Perspectives), CaGIS (Cartography and Geographic Information Science)

9.1.4. Invited Talks

- Jean-Daniel Fekete: Bilan du programme ANR Kinsources, Paris, "Visualisation de réseaux dynamiques", July 8th, 2016
- Jean-Daniel Fekete: 2nd European Conference on Social Networks (EUSN), Paris, "Challenges in Social Network Visualization: Bigger, Dynamic, Multivariate.", June 16th, 2016
- Jean-Daniel Fekete: School in Geomatique (GIS), Lyon, "DataVis and InfoVis", May 31th, 2016.
- Jean-Daniel Fekete: Invited Talk, Aarhus University, "ProgressiVis: a New Workflow Model for Scalability in Information Visualization", Apr 20, 2016
- Jean-Daniel Fekete: Invited Talk, Inria Lille, "ProgressiVis: a New Workflow Model for Scalability in Information Visualization", Apr 15, 2016
- Pierre Dragicevic: "Bad Stats are Miscommunicated Stats". Invited talk at RWTH Aachen, 5 Feb 2016.
- Pierre Dragicevic: "Statistical Dances: Why No Statistical Analysis is Reliable and What To Do About It". Keynote talk at the BioVis 2016 symposium, 23 Oct 2016.
- Tobias Isenberg: "Tactile Navigation and Selection for 3D Data Exploration". University of Granada, Spain, May 25, 2016.
- Christoph Kinkeldey: "Challenges in Visually-Supported Reasoning with Uncertainty". VIA group, Telecom ParisTech, 23 June 2016.

9.1.5. Leadership within the Scientific Community

- Jean-Daniel Fekete is the chair of the Steering Community of the IEEE Information Visualization conference.
- Jean-Daniel Fekete is a member of the Visualization Steering Committee for the IEEE VIS conference
- Jean-Daniel Fekete is a member of the Steering Community of the EuroVis conference.
- Jean-Daniel Fekete is a member of the publication board of Eurographics.
- Tobias Isenberg is a member of the Steering Committee of the Expressive conference.
- Tobias Isenberg is a member of the Executive Committee of the Visualization and Computer Graphics Technical Committee of the IEEE Computer Society.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

- "Visual Analytics" taught by Jean-Daniel Fekete at Ecole Centrale Paris, France
- "Block Course Visual Analytics" taught by Petra Isenberg at the University of Dresden, Germany
- "Visual Analytics" taught by Petra Isenberg at Ecole Centrale Paris, France
- "Introduction to Human-Computer Interaction" taught by Petra Isenberg at Ecole Centrale Paris, France
- "Information Visualization" co-taught by Petra Isenberg at Université Paris Sud
- "Non-Photorealistic Rendering" taught by Tobias Isenberg at the University of Granada, Spain
- "Introduction to Computer Graphics" taught by Tobias Isenberg at Polytech Paris-Sud
- "Photorealistic Rendering" taught by Tobias Isenberg at Polytech Paris-Sud and Université Paris-Saclay

9.2.2. Supervision

PhD: Evanthia Dimara, Information Visualization for Decision Making, Université Paris-Sud, 2014, Pierre Dragicevic and Anastasia Bezerianos

PhD: Mathieu Le Goc, Supporting Versatility in Tangible User Interfaces Using Collections of Small Actuated Objects, Université Paris-Sud, Defense date 15 Dec 2016, Pierre Dragicevic and Jean-Daniel Fekete

PhD: Pascal Goffin, From Individual to Collaborative Work, Université Paris-Sud, defended October 2016, Petra Isenberg and Jean-Daniel Fekete

PhD: Lonni Besançon, An Interaction Continuum for Scientific Visualization, Université Paris-Sud, 2014, Tobias Isenberg

9.2.3. Juries

- Jean-Daniel Fekete: Julien Tierny, HdR (Tenure) "Contributions à l'Analyse Topologique de Données pour la Visualisation Scientifique" , LIP6, Paris, Apr 29, 2016.
- Jean-Daniel Fekete: Matthias Nielsen, "Interactive Visual Analytics of Big Data A Web-Based Approach", Aarhus, Denmark, Apr 20, 2016
- Jean-Daniel Fekete: Mi-term PhD evaluation committee of Lobo María-Jesús, Saclay, May 2016.
- Jean-Daniel Fekete: Hiring Committee, Univ. Toulouse 1, France
- Pierre Dragicevic: Commission Scientifique Inria Saclay
- Pierre Dragicevic: PhD committee of Chat Wacharamanotham: Input Accuracy for Touch Surfaces. RWTH Aachen University, defended 5 February 2016.
- Pierre Dragicevic: Mi-term PhD evaluation committee of Nolwenn Maudet. 8 Sep 2016.
- Pierre Dragicevic: Mi-term PhD evaluation committee of Rafael Morales. 26 May 2016.
- Pierre Dragicevic: Post-doctoral assessment committee, project "Shareable Dynamic Media in Design and Knowledge Work", Aarhus University. Nov 2016.
- Petra Isenberg: Jury CR2 Inria Saclay

9.3. Popularization

- Pierre Dragicevic and Yvonne Jansen: the Curated List of Physical Visualizations is continuously being updated.
- Elsa Ferreira's interview with Mathieu Le Goc, Zooids: mais qui sont ces robots mignons?, Makery Magazine, 26 Nov 2016.

- Stéphanie Dupont's interview with Mathieu Le Goc, Jean-Daniel Fekete and Pierre Dragicevic: Nouvelle interface homme-machine: des mini robots pour l'affichage et l'interaction. 26 Oct 2016.

- Petra Isenberg, Tobias Isenberg, and collaborators published an open dataset of IEEE VIS publications at http://vispubdata.org

- Petra Isenberg, Tobias Isenberg, and collaborators published an online tool to explore keywords of the IEEE Visualization conference at http://keyvis.org

- Pascal Goffin: "Enrichir le mot écrit – Intégrer texte et données avec les Word-Scale Visualizations (visualisations mots)". EnsadLab/Spatial Media/Hist3D Colloquium: Les images du numérique. Histoires et futurs des images produites par ordinateur. Paris, France, 21 Mar 2016.

- Mathieu Le Goc and Zooids co-authors: Material for Zooids.

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] B. BACH, P. DRAGICEVIC, D. ARCHAMBAULT, C. HURTER, S. CARPENDALE.*A Descriptive Framework for Temporal Data Visualizations Based on Generalized Space-Time Cubes*, in "Computer Graphics Forum", April 2016 [*DOI :* 10.1111/CGF.12804], https://hal-enac.archives-ouvertes.fr/hal-01303506.

[2] B. BACH, C. SHI, N. HEULOT, T. MADHYASTHA, T. GRABOWSKI, P. DRAGICEVIC.*Time Curves: Folding Time to Visualize Patterns of Temporal Evolution in Data*, in "IEEE Transactions on Visualization and Computer Graphics", January 2016, vol. 22, n° 1 [*DOI :* 10.1109/TVCG.2015.2467851], https://hal.inria.fr/hal-01205821.

[3] M. BEHRISCH, B. BACH, N. HENRY RICHE, T. SCHRECK, J.-D. FEKETE.*Matrix Reordering Methods for Table and Network Visualization*, in "Computer Graphics Forum", 2016, vol. 35, 24 [*DOI :* 10.1111/CGF.12935], https://hal.inria.fr/hal-01326759.

[4] M. BEHRISCH, B. BACH, M. HUND, M. DELZ, L. VON RÜDEN, J.-D. FEKETE, T. SCHRECK.*Magnostics: Image-based Search of Interesting Matrix Views for Guided Network Exploration*, in "IEEE Transactions on Visualization and Computer Graphics", 2017, 1 [*DOI :* 10.1109/TVCG.2016.2598467], https://hal.inria.fr/hal-01377861.

[5] L. BESANÇON, P. ISSARTEL, M. AMMI, T. ISENBERG.*Hybrid Tactile/Tangible Interaction for 3D Data Exploration*, in "IEEE Transactions on Visualization and Computer Graphics", January 2017, vol. 23, n° 1, p. 881–890 [*DOI :* 10.1109/TVCG.2016.2599217], https://hal.inria.fr/hal-01372922.

[6] N. BOUKHELIFA, A. BEZERIANOS, W. CANCINO, E. LUTTON.*Evolutionary Visual Exploration: Evaluation of an IEC Framework for Guided Visual Search*, in "Evolutionary Computation", 2016 [*DOI :* 10.1162/EVCO_A_00161], https://hal.inria.fr/hal-01218959.

[7] J. BOY, L. EVEILLARD, F. DETIENNE, J.-D. FEKETE.*Suggested Interactivity: Seeking Perceived Affordances for Information Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", January 2016, vol. 22, n° 1, p. 639-648 [*DOI :* 10.1109/TVCG.2015.2467201], https://hal.inria.fr/hal-01188973.

[8] E. DIMARA, A. BEZERIANOS, P. DRAGICEVIC.*The Attraction Effect in Information Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", 2017, vol. 23, nº 1 [*DOI :* 10.1109/TVCG.2016.2598594], https://hal.inria.fr/hal-01355750.

[9] J. FUCHS, P. ISENBERG, A. BEZERIANOS, D. KEIM.*A Systematic Review of Experimental Studies on Data Glyphs*, in "IEEE Transactions on Visualization and Computer Graphics", 2017 [*DOI :* 10.1109/TVCG.2016.2549018], https://hal.inria.fr/hal-01378429.

[10] P. GOFFIN, J. BOY, W. WILLETT, P. ISENBERG.*An Exploratory Study of Word-Scale Graphics in Data-Rich Text Documents*, in "IEEE Transactions on Visualization and Computer Graphics", 2017 [*DOI :* 10.1109/TVCG.2016.2618797], https://hal.inria.fr/hal-01389998.

[11] P. ISENBERG, F. HEIMERL, S. KOCH, T. ISENBERG, P. XU, C. D. STOLPER, M. SEDLMAIR, J. CHEN, T. MÖLLER, J. STASKO.*vispubdata.org: A Metadata Collection about IEEE Visualization (VIS) Publications*, in "IEEE Transactions on Visualization and Computer Graphics", 2017, vol. 23 [*DOI :* 10.1109/TVCG.2016.2615308], https://hal.inria.fr/hal-01376597.

[12] P. ISENBERG, T. ISENBERG, M. SEDLMAIR, J. CHEN, T. MÖLLER.*Visualization as Seen Through its Research Paper Keywords*, in "IEEE Transactions on Visualization and Computer Graphics", January 2017, vol. 23, nº 1, p. 771–780 [*DOI :* 10.1109/TVCG.2016.2598827], https://hal.inria.fr/hal-01370750.

[13] K. LAWONN, S. GLASSER, A. VILANOVA, B. PREIM, T. ISENBERG.*Occlusion-free Blood Flow Animation with Wall Thickness Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", January 2016, vol. 22, nº 1, p. 728–737 [*DOI :* 10.1109/TVCG.2015.2467961], https://hal.inria.fr/hal-01178054.

[14] D. LÓPEZ, L. OEHLBERG, C. DOGER, T. ISENBERG.*Towards an Understanding of Mobile Touch Navigation in a Stereoscopic Viewing Environment for 3D Data Exploration*, in "IEEE Transactions on Visualization and Computer Graphics", May 2016, vol. 22, nº 5, p. 1616-1629 [*DOI :* 10.1109/TVCG.2015.2440233], https://hal.inria.fr/hal-01174618.

[15] B. PREIM, A. BAER, D. CUNNINGHAM, T. ISENBERG, T. ROPINSKI.*A Survey of Perceptually Motivated 3D Visualization of Medical Image Data*, in "Computer Graphics Forum", June 2016, vol. 35, nº 3, p. 501-525 [*DOI :* 10.1111/CGF.12927], https://hal.inria.fr/hal-01310290.

[16] W. WILLETT, Y. JANSEN, P. DRAGICEVIC.*Embedded Data Representations*, in "IEEE Transactions on Visualization and Computer Graphics", 2017, p. 461 - 470 [*DOI :* 10.1109/TVCG.2016.2598608], https://hal.inria.fr/hal-01377901.

[17] L. YU, K. EFSTATHIOU, P. ISENBERG, T. ISENBERG.*CAST: Effective and Efficient User Interaction for Context-Aware Selection in 3D Particle Clouds*, in "IEEE Transactions on Visualization and Computer Graphics", January 2016, vol. 22, nº 1, p. 886–895 [*DOI :* 10.1109/TVCG.2015.2467202], https://hal.inria.fr/hal-01178051.

[18] E. ZGRAGGEN, A. GALAKATOS, A. CROTTY, J.-D. FEKETE, T. KRASKA.*How Progressive Visualizations Affect Exploratory Analysis*, in "IEEE Transactions on Visualization and Computer Graphics", 2017 [*DOI :* 10.1109/TVCG.2016.2607714], https://hal.inria.fr/hal-01377896.

Articles in Non Peer-Reviewed Journal

[19] H. LAM, P. ISENBERG, T. ISENBERG, M. SEDLMAIR.*BELIV 2014 Special Issue*, in "Information Visualization", October 2016, vol. 15, n⁰ 4, p. 286-287 [*DOI :* 10.1177/1473871616638137], https://hal.inria.fr/hal-01298534.

International Conferences with Proceedings

[20] L. BESANÇON, M. AMMI, T. ISENBERG.*Pressure-Based Gain Factor Control for Mobile 3D Interaction using Locally-Coupled Devices*, in "CHI 2017 - ACM CHI Conference on Human Factors in Computing Systems", Denver, United States, May 2017 [*DOI :* 10.1145/3025453.3025890], https://hal.inria.fr/hal-01436172.

[21] L. BESANÇON, P. ISSARTEL, M. AMMI, T. ISENBERG.*Mouse, Tactile, and Tangible Input for 3D Manipulation*, in "Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)", Denver, United States, May 2017 [*DOI :* 10.1145/3025453.3025863], https://hal.inria.fr/hal-01436206.

[22] P. ISSARTEL, L. BESANÇON, T. ISENBERG, M. AMMI.*A Tangible Volume for Portable 3D Interaction*, in "International Symposium on Mixed and Augmented Reality", Merida, Mexico, September 2016, https://hal.inria.fr/hal-01423533.

[23] J. KRAUSE, A. DASGUPTA, J.-D. FEKETE, E. BERTINI.*SeekAView: An Intelligent Dimensionality Reduction Strategy for Navigating High-Dimensional Data Spaces*, in "LDAV 2016 - IEEE 6th Symposium on Large Data Analysis and Visualization", Baltimore, MD, United States, IEEE, October 2016, https://hal.inria.fr/hal-01377974.

[24] *Best Paper*
M. LE GOC, L. H. KIM, A. PARSAEI, J.-D. FEKETE, P. DRAGICEVIC, S. FOLLMER.*Zooids: Building Blocks for Swarm User Interfaces*, in "Proceedings of the Symposium on User Interface Software and Technology (UIST)", New York, NY, United States, ACM, October 2016, p. 97 - 109 [*DOI :* 10.1145/2984511.2984547], https://hal.inria.fr/hal-01391281.

Conferences without Proceedings

[25] *Best Paper*
S. K. BADAM, C. KINKELDEY, P. ISENBERG.*Haztrailz: Exploratory Analysis of Trajectory and Sensor Data: VAST 2016 Mini Challenge 2 Award: Honorable Mention for Clear Analysis Strategy*, in "IEEE VIS 2016", Baltimore, United States, November 2016, https://hal.inria.fr/hal-01397007.

[26] S. HURON, S. CARPENDALE, J. BOY, J.-D. FEKETE.*Using VisKit: A Manual for Running a Constructive Visualization Workshop*, in "Pedagogy of Data Visualization Workshop at IEEE VIS 2016", Baltimore, MD, United States, October 2016, https://hal.inria.fr/hal-01384388.

[27] T. ISENBERG.*Interactive NPAR: What type of tools should we create?*, in "Proceedings of Expressive", Lisbon, Portugal, Eurographics Association, May 2016, p. 89-96 [*DOI :* 10.2312/EXP.20161067], https://hal.inria.fr/hal-01298075.

[28] P. ISSARTEL, L. BESANÇON, F. GUÉNIAT, T. ISENBERG, M. AMMI.*Preference Between Allocentric and Egocentric 3D Manipulation in a Locally Coupled Configuration*, in "ACM Symposium on Spatial User

Interaction", Tokyo, Japan, ACM, October 2016 [*DOI :* 10.1145/2983310.2985750], https://hal.archives-ouvertes.fr/hal-01373269.

[29] M. LE GOC, P. DRAGICEVIC, S. HURON, J. BOY, J.-D. FEKETE.*A Better Grasp on Pictures Under Glass: Comparing Touch and Tangible Object Manipulation using Physical Proxies*, in "Proceedings of the International Working Conference on Advanced Visual Interfaces (AVI)", Bari, Italy, June 2016 [*DOI :* 10.1145/2909132.2909247], https://hal.inria.fr/hal-01323474.

Scientific Books (or Scientific Book chapters)

[30] P. DRAGICEVIC.*Fair Statistical Communication in HCI*, in "Modern Statistical Methods for HCI", Springer, 2016, p. 291 - 330 [*DOI :* 10.1007/978-3-319-26633-6_13], https://hal.inria.fr/hal-01377894.

[31] T. ISENBERG.*Interactive Exploration of Three-Dimensional Scientific Visualizations on Large Display Surfaces*, in "Collaboration Meets Interactive Spaces", C. ANSLOW, P. CAMPOS, J. JORGE (editors), Springer, December 2016 [*DOI :* 10.1007/978-3-319-45853-3_6], https://hal.inria.fr/hal-01379938.

Books or Proceedings Editing

[32] G. BEBIS, R. BOYLE, B. PARVIN, D. KORACIN, F. PORIKLI, S. SKAFF, A. ENTEZARI, J. MIN, D. IWAI, A. SADAGIC, C. SCHEIDEGGER, T. ISENBERG (editors). *Advances in Visual Computing: Proceedings of the 12th International Symposium (ISVC, December 12–14, Las Vegas, NV, USA), Part 1*, December 2016 [*DOI :* 10.1007/978-3-319-50835-1], https://hal.inria.fr/hal-01415869.

[33] G. BEBIS, R. BOYLE, B. PARVIN, D. KORACIN, F. PORIKLI, S. SKAFF, A. ENTEZARI, J. MIN, D. IWAI, A. SADAGIC, C. SCHEIDEGGER, T. ISENBERG (editors). *Advances in Visual Computing: Proceedings of the 12th International Symposium (ISVC, December 12–14, Las Vegas, NV, USA), Part 2*, December 2016 [*DOI :* 10.1007/978-3-319-50832-0], https://hal.inria.fr/hal-01415875.

[34] M. SEDLMAIR, P. ISENBERG, T. ISENBERG, N. MAHYAR, H. LAM (editors). *Proceedings of the Sixth Workshop on "Beyond Time and Errors:Novel Evaluation Methods for Visualization" (BELIV 2016, October 24, Baltimore, Maryland, USA)*, Beyond Time and Errors—Novel Evaluation Methods for Visualization (BELIV), Baltimore, Maryland, United States, October 2016 [*DOI :* 10.1145/2993901], https://hal.inria.fr/hal-01375428.

Research Reports

[35] L. BESANÇON, P. ISSARTEL, M. AMMI, T. ISENBERG.*Usability Comparison of Mouse, Touch and Tangible Inputs for 3D Data Manipulation*, Inria, April 2016, https://hal.inria.fr/hal-01305726.

[36] E. DIMARA, A. BEZERIANOS, P. DRAGICEVIC.*Testing the Attraction Effect on Two Information Visualization Datasets*, Inria Saclay Ile de France, March 2016, n[o] RR-8895, 13, https://hal.inria.fr/hal-01295624.

[37] T. TSANDILAS, P. DRAGICEVIC.*Accounting for Chance Agreement in Gesture Elicitation Studies*, LRI - CNRS, University Paris-Sud, February 2016, n° 1584, 5, https://hal.archives-ouvertes.fr/hal-01267288.

Other Publications

[38] J.-D. FEKETE, R. PRIMET.*Progressive Analytics: A Computation Paradigm for Exploratory Data Analysis*, July 2016, working paper or preprint, https://hal.inria.fr/hal-01361430.

[39] P. ISSARTEL, L. BESANÇON, T. ISENBERG, M. AMMI.*A Tangible Volume for Portable 3D Interaction*, March 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01284985.

[40] P. ISSARTEL, F. GUÉNIAT, T. ISENBERG, M. AMMI.*Analysis of Locally Coupled 3D Manipulation Mappings Based on Mobile Device Motion*, 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01290738.

[41] Y. JANSEN, K. HORNBAEK, P. DRAGICEVIC.*What Did Authors Value in the CHI'16 Reviews They Received?*, May 2016, p. 596 - 608, Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems [*DOI :* 10.1145/2851581.2892576], https://hal.inria.fr/hal-01405042.

[42] M. KAY, S. HAROZ, S. GUHA, P. DRAGICEVIC.*Special Interest Group on Transparent Statistics in HCI*, May 2016, p. 1081–1084, Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems [*DOI :* 10.1145/2851581.2886442], https://hal.inria.fr/hal-01405018.

[43] M. SERENO, M. AMMI, T. ISENBERG, L. BESANÇON.*Tangible Brush: Performing 3D Selection with Portable and Position-aware Devices*, October 2016, IEEE VIS2016, Poster, https://hal.inria.fr/hal-01372925.

References in notes

[44] S. K. CARD, J. D. MACKINLAY, B. SHNEIDERMAN (editors). *Readings in information visualization: using vision to think*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1999.

[45] B. BACH, N. HENRY RICHE, R. FERNANDEZ, E. GIANNISAKIS, B. LEE, J.-D. FEKETE.*NetworkCube: Bringing Dynamic Network Visualizations to Domain Scientists*, October 2015, Posters of the Conference on Information Visualization (InfoVis), Poster, https://hal.inria.fr/hal-01205822.

[46] J. BERTIN.*Sémiologie graphique : Les diagrammes - Les réseaux - Les cartes*, Les réimpressions, Editions de l'Ecole des Hautes Etudes en Sciences, Paris, France, 1967.

[47] A. BEZERIANOS, P. DRAGICEVIC, R. BALAKRISHNAN.*Mnemonic rendering: an image-based approach for exposing hidden changes in dynamic displays*, in "UIST '06: Proceedings of the 19th annual ACM symposium on User interface software and technology", New York, NY, USA, ACM, 2006, p. 159–168, http://doi.acm.org/10.1145/1166253.1166279.

[48] J.-D. FEKETE.*Reorder.js: A JavaScript Library to Reorder Tables and Networks*, October 2015, IEEE VIS 2015, Poster, https://hal.inria.fr/hal-01214274.

[49] J.-D. FEKETE, C. PLAISANT.*Interactive Information Visualization of a Million Items*, in "Proc. IEEE Symposium on Information Visualization 2002 (InfoVis 2002)", Boston, USA, IEEE Press, October 2002, p. 117-124.

[50] J.-D. FEKETE, C. PLAISANT.*Les leçons tirées des deux compétitions de visualisation d'information*, in "Proceedings of IHM 2004", Namur, Belgium, International Conference Proceedings Series, ACM Press, September 2004, p. 7-12.

[51] J. J. GIBSON.*The Ecological Approach to Visual Perception*, Lawrence Erlbaum Associates, New Jersey, USA, 1979.

[52] Y. GUIARD, Y. DU, J.-D. FEKETE, M. BEAUDOUIN-LAFON, C. APPERT, O. CHAPUIS.*Shakespeare's Complete Works as a Benchmark for Evaluating Multiscale Document-Navigation Techniques*, in "Proceedings of BEyond time and errors: novel evaLuation methods for Information Visualization (BELIV'06)", Venice, Italy, ACM Press, May 2006, p. 65-70.

[53] N. HENRY RICHE, J.-D. FEKETE.*Evaluating Visual Table Data Understanding*, in "Proceedings of BEyond time and errors: novel evaLuation methods for Information Visualization (BELIV'06)", Venice, Italy, ACM Press, May 2006, 6.

[54] B. LEE, C. PLAISANT, C. SIMS PARR, J.-D. FEKETE, N. HENRY RICHE.*Task taxonomy for graph visualization*, in "BELIV '06: Proceedings of the 2006 AVI workshop on BEyond time and errors", New York, NY, USA, ACM, 2006, p. 1–5, http://doi.acm.org/10.1145/1168149.1168168.

[55] C. PLAISANT, J.-D. FEKETE, G. GRINSTEIN.*Promoting Insight-Based Evaluation of Visualizations: From Contest to Benchmark Repository*, in "IEEE Transactions on Visualization and Computer Graphics", 2008, vol. 14, n° 1, p. 120–134, http://doi.ieeecomputersociety.org/10.1109/TVCG.2007.70412.

[56] A. TRIESMAN.*Preattentive Processing in Vision*, in "Computer Vision, Graphics, and Image Processing", August 1985, vol. 31, n° 2, p. 156-177.

[57] E. TUFTE.*The Visual Display of Quantitative Information*, Graphics Press, 1983.

[58] J. W. TUKEY.*Exploratory Data Analysis*, Addison-Wesley, 1977.

Team CEDAR

Rich Data Exploration at Cloud Scale

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

RESEARCH CENTER
Saclay - Île-de-France

THEME
Data and Knowledge Representation and Processing

Table of contents

Team CEDAR

Creation of the Team: 2016 January 01

Keywords:

Computer Science and Digital Science:

3.1.1. - Modeling, representation
3.1.2. - Data management, quering and storage
3.1.3. - Distributed data
3.1.6. - Query optimization
3.1.7. - Open data
3.1.8. - Big data (production, storage, transfer)
3.1.9. - Database
3.2.1. - Knowledge bases
3.2.3. - Inference
3.2.4. - Semantic Web
3.2.5. - Ontologies
3.3. - Data and knowledge analysis
3.3.1. - On-line analytical processing
3.3.2. - Data mining
3.3.3. - Big data analysis
8.1. - Knowledge

Other Research Topics and Application Domains:

1.1.6. - Genomics
8.5.1. - Participative democracy
9.4.5. - Data science
9.7.2. - Open data

1. Members

Research Scientists
Ioana Manolescu [Team leader, Inria, Senior Researcher, HDR]
Michael Thomazo [Inria, Researcher]

Faculty Member
Yanlei Diao [Ecole Polytechnique, Professor, HDR]

Technical Staff
Oscar Santiago Mendoza Rivera [Inria]
Swen Ribeiro [Inria]

PhD Students
Damian Bursztyn [Inria]
Tien Duc Cao [Inria, from Apr 2016]
Sejla Cebiric [Inria]
Raphael Bonaque [Inria, until Sep 2016]
Enhui Huang [École Polytechnique, from Sep 2016]

Post-Doctoral Fellows
  Alessandro Solimando [Inria, until Sep 2016]
  Fei Song [Inria, from Dec 2016]
Visiting Scientists
  Davide Lanti [FUB, from Mar 2016]
  Rana Alotaibi [UCSD, August to October 2016]
  Xavier Tannier [UPS, Associate Professor, from May 2016, HDR]
  François Goasdoué [U. de Rennes,Professor, HDR]
Administrative Assistants
  Tiffany Caristan [Inria, from Jun 2016]
  Helena Kutniak [Inria]
  Maeva Jeannot [Inria]

2. Overall Objectives

2.1. Overall Objectives

Today data is being generated at an unprecedented rate, so much that 90created in the past two years. Such significant increase of data volume is due to the new ways that we gather data: from software tools that record system and user activities; from sensors and scientific instruments that monitor our built and natural environment; from medical instruments that enable genomic diagnosis of patients; and from user-initiated sources on the Web or social networks. Data often comes with semantics, enriching its interpretation and enhancing its value. Importantly, we observe that in today's data-intensive application, variety is the norm, and is likely to re- main so for a while. This is because different applications are best served by different kinds of data: traditional commerce-oriented applications use relational databases, Web content management systems handle semistruc- tured documents, sensors provide numerical streams, science applications manipulate arrays, highly heterogeneous data sets is often exported in RDF graphs, software system logs consist of structured text etc. At the scale and speed of consumption of today's Big Data, unifying data across such formats into a single architecture (approach formerly known as extract-transform-load in a data warehouse context) is no longer feasible. Instead, Cedar aims at inventing expressive models and highly efficient data management tools, focused from the start on Big Data variety. Our tools will be designed for deployment in the cloud, and validated at large scale.

3. Research Program

3.1. Scalable Heterogeneous Stores

Big Data applications increasingly involve *diverse* data sources, such as: structured or unstructured documents, data graphs, relational databases etc. and it is often impractical to load (consolidate) diverse data sources in a single repository. Instead, interesting data sources need to be exploited "as they are", with the added value of the data being realized especially through the ability to combine (join) together data from several sources. Systems capable of exploiting diverse Big Data in this fashion are usually termed *polystores*. A current limitation of polystores is that data stays captive of its original storage system, which may limit the data exploitation performance. We work to devise highly efficient storage systems for heterogeneous data across a variety of data stores.

3.2. Semantic Query Answering

In the presence of data semantics, query evaluation techniques are insufficient as they only take into account the database, but do not provide the reasoning capabilities required in order to reflect the semantic knowledge. In contrast, (ontology-based) query answering takes into account both the data and the semantic knowledge in order to compute the full query answers, blending query evaluation and semantic reasoning.

We aim at designing efficient semantic query answering algorithms, both building on cost-based reformulation algorithms developed in the team and exploring new approaches mixing materialization and reformulation.

3.3. Multi-Model Querying

As the world's affairs get increasingly more digital, a large and varied set of data sources becomes available: they are either structured databases, such as government-gathered data (demographics, economics, taxes, elections, ...), legal records, stock quotes for specific companies, un-structured or semi-structured, including in particular graph data, sometimes endowed with semantics (see e.g. the Linked Open Data cloud). Modern data management applications, such as data journalism, are eager to combine in innovative ways both static and dynamic information coming from structured, semi-structured, and un-structured databases and social feeds. However, current content management tools for this task are not suited for the task, in particular when they require a lenghy rigid cycle of data integration and consolidation in a warehouse. Thus, we see a need for flexible tools allowing to interconnect various kinds of data sources and to query them together.

3.4. Interactive Data Exploration at Scale

In the Big Data era we are faced with an increasing gap between the fast growth of data and the limited human ability to comprehend data. Consequently, there has been a growing demand of data management tools that can bridge this gap and help users retrieve high-value content from data more effectively. To respond to such user information needs, we aim to build interactive data exploration as a new database service, using an approach called "explore-by-example".

3.5. Exploratory Querying of Semantic Graphs

Semantic graphs including data and knowledge are hard to apprehend for users, due to the complexity of their structure and oftentimes to their large volumes. To help tame this complexity, in prior research (2014), we have presented a full framework for RDF data warehousing, specifically designed for heterogeneous and semantic-rich graphs. However, this framework still leaves to the users the burden of chosing the most interesting warehousing queries to ask. More user-friendly data management tools are needed, which help the user discover the interesting structure and information hidden within RDF graphs.

3.6. Representative Semantic Query Answering

Top-k search is a classical topic, studied in relational databases, semantic web, recommandation systems,... It is extremely useful, among other, when a human user face a large number of query results, allowing the user to reformulate the query if necessary. However, we argue that top-k search incurs a bias on the perception of the set of results which is out of the control of the user. Our goal is to provide the user with k answers as well which are chosen so as to represent the diversity of the answer set. We will first consider this problem in the setting of relational or RDF databases. We will then extend to more heterogeneous sources, including in particular plain text.

4. Application Domains

4.1. Computational Journalism

Modern journalism increasingly relies on content management technologies in order to represent, store, and query source data and media objects themselves. Writing news articles increasingly requires consulting several sources, interpreting their findings in context, and crossing links between related sources of information. CEDARresearch results directly applicable to this area provide techniques and tools for rich Web content warehouse management. This work will be funded by the ANR ContentCheck project, and a Google Award on Even Thread Extraction. We work in collaboration with Le Monde's "Les Décodeurs" team to investigate these topics.

4.2. Open Data Intelligence

The Web is a vast source of information, to which more is added every day either in unstructured form (Web pages) or, increasingly, as partially structured sources of information, in particular as Open Data sets, which can be seen as connected graphs of data, most frequently described in the RDF data format recommended by the W3C. Further, RDF data is also the most appropriate format for representing structured information extracted automatically from Web pages, such as the DBPedia database extracted from Wikipedia or Google's InfoBoxes. We work on this topic within the 4-year project ODIN started in 2014.

4.3. Hybrid Data Warehousing

Increasingly many modern applications need to exploit data from a variety of formats, including relations, text, trees, graphs etc. The recent development of data management systems aimed at "Big Data", including NoSQL platforms, large-scale distributed systems etc. provides enteprise architects with many systems to chose from. This makes it hard to decide which part of the application data to handle in which system, especially given that each system is best at handling a specific kind of data and a certain class of operations. CEDARinvestigates principled techniques for distributing an application's data sources across a variety of systems and data models, based on materialized views. We test our ideas in this area within the Datalyse project.

5. Highlights of the Year

5.1. Highlights of the Year

ERC Proposal Accepted

Y. Diao's ERC Consolidator proposal "Charting a New Horizon of Big and Fast Data Analysis through Integrated Algorithm Design" has been accepted by the EU.

Awards

- A team of five including the team's PhD student Tien Duc Cao has won the first place at the Start-up Week-End in Artificial Intelligence (SWAI) in November 2016 (https://twitter.com/i/moments/796004617410711552, http://swai.fr/).
- Šejla Čebirić has been awarded the Google Anita Borg Scholarship.
- The paper "On the Complexity of Evaluating Regular Path Queries over Linear Existential Rules." by M. Bienvenu and M. Thomazo received the best paper award at the RR'16 conference .

BEST PAPERS AWARDS :

[3] **10th International Conference on Web Reasoning and Rule Systems**. M. BIENVENU, M. THOMAZO.

6. New Software and Platforms

6.1. New Software

6.1.1. *CliqueSquare*

CliqueSquare allows storing and querying very large volumes of RDF data in a massively parralel fashion in a Hadoop cluster. The system uses its own partitioning and storage model for the RDF triples in the cluster.

CliqueSquare evaluates queries expressed in a dialect of the SPARQL query language. It is particularly efficient when processing complex queries, because it is capable of translating them into MapReduce programs guaranteed to have the minimum number of successive jobs. Given the high overhead of a MapReduce job, this advantage is considerable.

6.1.2. *Compact*

Compact reformulates conjunctive queries in the setting of ontology-based query anwering. It takes as input a conjunctive query and an ontology, and outputs a first-order rewriting of that query whenever it exists (without termination guarantee when it does not exists). To ease its use and dissemination, a novel version has been implemented by M. Thomazo based on the framework GRAAL, developed within the Inria Sophia-Antipolis team GraphIK by C. Sipieter, an engineer funded by an ADT. It will in particulary ease the integration with Semantic Web standards, as well as the use of query optimization techniques developed within Cedar for RDFS and DL-Lite$_\mathcal{R}$ to more general ontology languages.

6.1.3. *RDF-Commons*

RDF-Commons is a set of modules providing the abilities to *i)* load and store RDF data in a DBMS *ii)* parse RDF conjunctive queries *iii)* encode URIs and literals into integers *iv)* encode RDF conjunctive queries *v)* build statistics on RDF data *vi)* estimate the cost of the evaluation of a conjunctive query *vii)* saturate the RDF data, with respect to an RDF Schema *viii)*reformulate a conjunctive query with respect to an RDF Schema (ix) propose algebraic plans.

The algebraic plan part has been developed by A. Solimando and D. Bursztyn. An ADT funding for two years has been granted to consolidate and extend the development of RDF-Commons. The hiring process is ongoing.

6.1.4. *RDFSummary*

RDF Summary is a standalone Java software capable of building summaries of RDF graphs. Summaries are compact graphs (typically several orders of magnitude smaller than the original graph), which can be used to get acquainted quickly with a given graph, they can also be used to perform static query analysis, infer certain things about the answer of a query on a graph, just by considering the query and the summary.

6.1.5. *Tatooine*
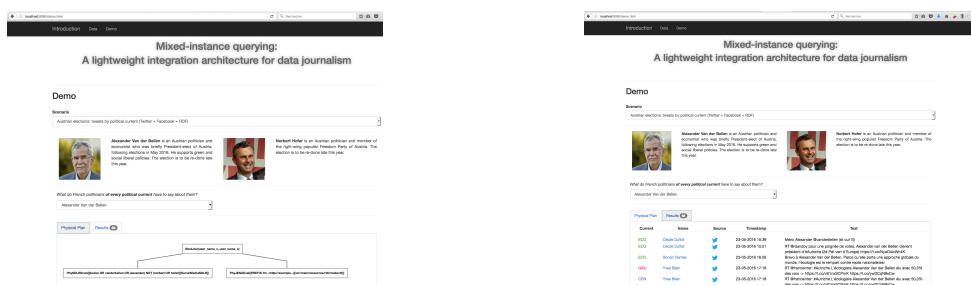
Figure 1. Tweet enrichment in Tatooine: evaluation plan (left) and results (right).

We developed lightweight data integration system called Tatooine, based on our discussions with our journalist partners in the ANR ContentCheck project from the team "Les Décodeurs". Tatooine allows to exploit heterogeneous data sources of different data models, which we view as a mixed data instance, by querying them together; Tatooine combines data from various sources within an integrated engine complemented by information extraction and data visualization modules. Figure 1 illustrates the functioning of Tatooine through screen captures: a set of tweets (JSON documents stored in SOLR) obtained through a full-text search are combined with information about their authors (RDF metadata stored in Jena TDB) and the results are presented to the users highlighting the political affiliation of the tweet authors.

7. New Results

7.1. Scalable Heterogeneous Stores

To improve data querying performance within polystores (Section 3.1), we developed Estocada, a novel system capable of exploiting side-by-side a practically unbound variety of data management system, all the while guaranteeing the soundness and completeness of the store, and striving to extract the best performance out of the various DMSs. Estocada leverages recent advances in the area of query rewriting under constraints, which we use to capture the various data models and describe the fragments stored within each data management system. Estocada was demonstrated at the IEEE ICDE conference [12]; recent experimental results demonstrated performance improvements by many orders of magnitude brought by the fragments Estocada supports, with respect to the setting where data is stored only in the system it originates from. This work continues, in collaboration with Alin Deutsch and Rana Alotaibi from UCSD.

7.2. Semantic Query Answering

This is a core topic for the team, in which the year has been particularly fruitful.

First, we investigated efficient query answering techniques in knowledge bases. A large and useful set of ontologies enjoys FOL (first-order logic) reducibility of query answering, that is: answering a query q can be reduced to evaluating a certain first-order logic (FOL) formula (obtained from the query and ontology) against only the explicit facts. We devised a novel query optimization framework for ontology-based data access settings enjoying FOL reducibility. Our framework is based on searching within a set of alternative equivalent FOL queries, that is, FOL reformulations, one with minimal evaluation cost when evaluated through a relational database system. We applied this framework to the DL-Lite$_\mathcal{R}$ Description Logic underpinning the W3C's OWL2 QL ontology language, and demonstrated through experiments its performance benefits when two leading SQL systems, one open-source and one commercial, are used for evaluating the FOL query reformulations. This work has lead to a major publication in the PVLDB journal [13], and a demonstration at the Semantic Web conference [4], while the complete details appear in [16] and the PhD thesis of the student author. [2].

Second, we initiated a study of extensions of conjunctive queries to conjunctive regular path queries. The first step has been to study regular path queries under linear existential rules, generalizing previous work on DL-Lite$_\mathcal{R}$, which is at the core of the Semantic Web OWL 2 QL profile. Regular path queries are queries that check for a path between two individuals, which is labeled by a word belonging to a given regular language. Such navigational languages are very popular for graph-based data representation, such as RDF. We have studied the complexity for this query language, and shown that it is NL-complete in data complexity, and EXPTIME-complete in combined complexity (and PTIME complete with bounded arity). This work has received the best paper award at RR'16 , and is currently being extented to conjunctive regular path queries.

Last, we studied the expressivity of several variants of Datalog, the classical language for deductive databases. In particular, we have studied its expressivity when given access (or not) to input negation (the ability to check if an extensional atoms hold or not) and to a linear order. We provided a complete Venn diagram regarding the expressivity of all the variants when considering homomorphism-closed query. The trickiest (and most surprising) points is the existence of polynomial-time computable homomorphism closed queries that are not expressible within Datalog with linear order but without input negation. These results have been published at IJCAI'16 [7].

7.3. Multi-model Querying

We have proposed a lightweight data integration architecture implemented within Tatooine (see Section 6.1.5); the system was demonstrated on a data journalism use case at the prestigious VLDB conference [9].

A separate effort in the area of multi-model querying considered querying databases of interconnected documents, users and concepts, by means of keywords. In this context, it is important that query results reflect not only the keywords present in documents but also the links between users and documents (so as to return to one user first the results authored in his social neighborhood), links between documents (for instance when a tweet answers another or an article has a link to another), and last but not least semantic information which allows interconnecting and interpreting terms mentioned in text. This research was finalized as part of the PhD of Raphaël Bonaque [1] and appeared at the EDBT conference 2016 [11].

7.4. Interactive Data Exploration at Scale

In the work with Enhui Huang (PhD student at Ecole Polytechnique), we seek to minimize the number of samples presented to the user for reviewing in order to build an accurate model of the user interest. In particular, as the dimensionality of the data space increases, the number of samples needed to build an accurate user interest model increases fast. We examine a range of popular feature selection techniques for data exploration, and for the best-performing feature selection technique, Gradient boosting regression trees (GBRT), we propose optimizations to overcome the issue of unbalanced training data and to dynamically determine the number of relevant features to select. Experimental results show that our optimized GBRT improves F-measure from nearly 0 without feature selection, to high F-measure (>0.8), by adaptively choosing the number of relevant features.

This work is currently under submission to a database conference.

7.5. Exploratory Querying of Semantic Graphs

We have started work with an intern (Zheng Zhang) toward automatically exploring the structure of an RDF graph and visualizing it with the help of a D3.js (https://d3js.org/) visualization library. These initial steps should serve to guide the beginning of an interactive exploration of the RDF graph in order to identify interesting analytical queries to be asked and evaluated. This work continues.

Separately, with a different intern (Javier Letelier), we have investigated efficient algorithms for keyword search in an RDF graph, exploiting structural and semantic knowledge about the graph; such knowledge is organized as an RDF summary which is an RDF graph itself. The algorithm was implemented and integrated as a text search tool within the Tatooine prototype; the work is ongoing.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

- AIDE ("A New Database Service for Interactive Exploration on Big Data") is an ANR "Young Researcher" project led by Y. Diao, to start at the end of 2016.
- CBOD ("Cloud-Based Organizational Design") is a 4-year ANR started in 2014, coordinated by prof. Ahmed Bounfour from UPS. Its goal is to study and model the ways in which cloud computing impacts the behavior and operation of companies and organizations, with a particular focus on the cloud-based management of data, a crucial asset in many companies.
- ContentCheck (2015-2018) is an ANR project in collaboration with U. Rennes 1 (F. Goasdoué), INSA Lyon (P. Lamarre), the LIMSI lab from U. Paris Sud, and the Le Monde newspaper, in particular their fact-checking team Les Décodeurs. Its aim is to investigate content management models and tools for journalistic fact-checking.

- Datalyse is funded for 3.5 years as part of the *Investissement d'Avenir - Cloud & Big Data* national program. The project is led by the Grenoble company Eolas, a subsidiary of Business & Decision. It is a collaboration with LIG Grenoble, U. Lille 1, U. Montpellier, and Inria Rhône-Alpes aiming at building scalable and expressive tools for Big Data analytics. The project has ended in November 2016.

8.1.2. LabEx, IdEx

- Structured, Social and Semantic Search (S4) is a 3-year project started in October 2013, financed by the *LabEx (Laboratoire d'Excellence)* DIGICOSME. The project aims at developing a data model for rich structured content enriched with semantic annotations and authored in a distributed setting, as well as efficient algorithms for top-k search on such content. The project has ended in September 2016.

- CloudSelect is a three-years project started in October 2015. It is financed by the *Institut de la Société Numérique* (ISN) of the IDEX Paris-Saclay; it funds the PhD scholarship of S. Cebiric. The project is a collaboration with A. Bounfour from the economics department of Université Paris Sud. The project aims at exploring technical and business-oriented aspects of data mobility across clo ud services, and from the cloud to outside the cloud. Research contributing to this project is carried in collaboration with U. California in San Diego (UCSD) (see Section 3.1).

8.1.3. Others

- ODIN is a four-year project started in 2014, funded by the Direction Générale de l'Armement, between the SemSoft company, IRISA Rennes and Cedar. The project aims to develop a complete framework for analytics on Web data, in particular taking into account uncertainty, based on Semantic Web technologies such as RDF.

- Google Award I. Manolescu has received a Google Award in collaboration with X. Tannier from LIMSI/CNRS and Université de Paris-Sud. The award is given within a call specifically dedicated to computing tools for computational journalism. The project given the award focuses on "Event Thread Extraction for Viewpoint Analysis"; the project has finished at the end of 2016.

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

The permanent members of the team participate to build a proposal called GDMA (Graph Data Management and Analytics, for an European Joint Doctorate within the Initial Training Network (ITN) chapter of Europe's H2020 program, with the University of Aalborg (Denmark), Université Libre de Bruxelles, Universitat Politecnica de Catalunya, and University of Ioannina (Greece). If successful the project would involve six PhD thesis co-supervised in Cedar and starting in 2018, three students mostly residing with us, and three abroad working with our partners from Aalborg and Brussels.

I. Manolescu has submitted a Marie-Curie proposal titled IDEAA (An interactive toolbox to help citizens understand and build a viewpoint on specific issues by monitoring, analysing, and interlinking public data from EU institutions) to host a junior researcher (Mirjana Mazuran from Politecnico di Milano) for two years.

8.3. International Initiatives

8.3.1. Inria International Partners

8.3.1.1. Informal International Partners

We continue collaborating with U. California in San Diego (UCSD) following the OAKSAD associated team (2013-2015), in particular in the Estocada project (Section 7.1).

8.4. International Research Visitors

8.4.1. Visits of International Scientists

Several international guests gave seminars in our group:

- L. Ach and M. Rezk (Rakuten)
- D. Calvanese (University of Bolzano)
- R. Cheng (Hong Kong University)
- M. Franklin (University of Berkeley)
- R. Kontchakov, S. Kikot, M. Zakharyaschev (Birbeck University College)
- Y. Papakonstantinou (University of California in San Diego)
- V. Vianu (University of California in San Diego)

8.4.1.1. Internships

- R. Alotaibi visited the team for two months working on scalable heterogeneous stores with D. Bursztyn and I. Manolescu.
- D. Lanti visited the team for five months, working on efficient semantic query answering with D. Bursztyn.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Selection

9.1.1.1. Chair of Conference Program Committees

Y. Diao has been the Program Committee Chair of the ACM Symposium on Cloud Computing (SoCC) 2016, Chair of the ACM Symposium on Cloud Computing, October 2016 and Chair of the Inaugural Paris Big Data Summit, March 2016.

I. Manolescu has been the Program Committee Chair of the International Conference on Scientific and Statistical Database Management (SSDBM) 2016, and the Vision Papers track chair in the Extending Database Technologies (EDBT) conference 2016.

9.1.1.2. Member of the Conference Program Committees

- I. Manolescu has been a member of program committee of the IEEE Conference on Data Engineering (ICDE) 2016, of the Workshop on Data Engineering for the Semantic Web (DESWeb, in conjunction with ICDE) 2016 and of the Workshop on Semantic Big Data (SBD) 2016 in conjunction with the ACM SIGMOD conference.
- M. Thomazo has been member of the PC of ICCS'16, IJCAI'16 and PRIMA'16.

9.1.1.3. Reviewer

- M. Thomazo has reviewed for RR'16, VLDB'16 and ICDT'17.

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

- Y. Diao is Editor-in-Chief of the ACM SIGMOD Record, Associate Editor of ACM Transactions on Databases (TODS) and Editor of the Proceedings of Very Large Databases (PVLDB), 2016
- I. Manolescu has been an Associate Editor for the ACM Transactions of the Web (TWeb) and a member of the editorial board of PVLDB 2016.

9.1.2.2. Reviewer - Reviewing Activities

- Y. Diao has reviewed for the EDBT vision track
- I. Manolescu has reviewed for the ACM Transactions on Database Systems (TODS).
- M. Thomazo has reviewed for the Journal of Web Semantics.

9.1.3. Invited Talks

- I. Manolescu has given an invited talk "Estocada: Flexible Hybrid Stores" at the INESC institute in Portugal.
- I. Manolescu has given an invited talk "CliqueSquare: Flat Plans for Massively Parallel RDF Queries" at the Aalborg University, in Denmark.
- M. Thomazo has presented FactMinder, a tool developed in the prequel of Cedar, to a panel of computer scientists and journalists at the Tech&Check event organized by Duke University

9.1.4. Leadership within the Scientific Community

I. Manolescu has been a coordinator of Task 1 (Scalable and Secure Data Processing) of the DataSense axis of the DigiCosme Labex (*Laboratoire d'Excellence*). I. Manolescu has been a member of the ACM SIGMOD Jim Gray PhD Award Committee, of the IEEE TCDE Awards Committee. She has also been a member of the EDBT (Extending Database Technology) Association and of the Steering Committee of BDA (*Bases de Données Avancées*). I. Manolescu has joined the Board of Trustees of the Proceedings of VLDB (PVLDB) Endowment in 2016.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

- Master: Y. Diao has taught the course "Systems for Big Data Analytics" in the Data Science Master Program of Université Paris Saclay
- Master: I. Manolescu, Architectures for Massive Data Management, 30h, M2, Université Paris-Saclay, France.
- Master: I. Manolescu, Database Management Systems, 52h, M1, École Polytechnique, France.
- Doctorat: I. Manolescu, Scalable Tools for Linked Data Analytics, 20h, Aalborg University, Denmark.
- Master: M. Thomazo, Database Management Systems, 20h, M1, École Polytechnique, France
- Master: M. Thomazo, Introduction to Big Data Systems, 40h, M2, Université Paris-Saclay, France

9.2.2. Supervision

Raphael Bonaque: "Structured, Social and Semantic Search", Université Paris Saclay, September 30, 2016, Bogdan Cautiş, François Goasdoué, and Ioana Manolescu.

Damian Bursztyn: "Scalable Techniques for Web Data Management", December 15, 2016, François Goasdoué and Ioana Manolescu.

Tien Duc Cao: "Extraction et interconnexion de connaissances appliquée aux données journalistiques", since October 2016, Ioana Manolescu and Xavier Tannier (LIMSI/CNRS and Université de Paris Sud)

Sejla Čebirić: "CloudSelect: Data Mobility Within, Across and Outside Clouds", since September 2015, F. Goasdoué and I. Manolescu.

Enhui Huang: Interactive Data Exploration at Scale, since October 2016, Y. Diao.

Y. Diao has as well supervised 5-10 students each semester for the 3A research project and been mentor and chair of the defense of 17 student summer internship projects.

9.2.3. *Juries*

I. Manolescu has been a member of the PhD committee of Raphaël Bonaque and Damian Burstzyn, and a referee and jury member for the PhD of Gonçalo Simões at INESC-ID, Portugal.

9.3. Popularization

- M. Thomazo has presented a game based on RDF graphs and social networks as part of Fête de la Science in Inria Saclay, in October.
- The ANR ContentCheck project on **content management techniques for journalistic fact-checking** has attracted news attention in a series of general-audience articles published by Ouest France, Le Devoir (Canada), Rue89, Le Monde, Inria and CNRS, among others:
  - http://www.ouest-france.fr/leditiondusoir/data/569/reader/reader.html#!preferred/1/package/569/pub/570/page/8
  - http://www.ledevoir.com/politique/canada/450937/sur-la-piste-du-mensonge
  - http://rue89.nouvelobs.com/2015/10/26/algorithmes-antimensonge-fin-bobards-politique-261827
  - http://data.blog.lemonde.fr/2015/10/23/le-fact-checking-peut-il-sautomatiser/
  - http://www.inria.fr/centre/saclay/actualites/un-logiciel-de-fact-checking-pour-comprendre-le-monde-qui-nous-entoure
  - https://lejournal.cnrs.fr/articles/un-logiciel-qui-decrypte-la-politique

The project has also lead to an interview by Science Po students studying in the Communication Master of the school, around automated fact-checking topics: https://inciveritas.wordpress.com/menu-page/lautomatisation-du-fact-checking/.

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] R. BONAQUE.*Top-k search over rich web content*, Université Paris-Saclay, September 2016, https://tel.archives-ouvertes.fr/tel-01418124.

[2] D. BURSZTYN.*Efficient Big Data query answering in the presence of constraints*, Université Paris-Saclay, December 2016, https://tel.archives-ouvertes.fr/tel-01449287.

International Conferences with Proceedings

[3] *Best Paper*
M. BIENVENU, M. THOMAZO.*On the Complexity of Evaluating Regular Path Queries over Linear Existential Rules*, in "10th International Conference on Web Reasoning and Rule Systems", Aberdeen, United Kingdom, 10th International Conference on Web Reasoning and Rule Systems, September 2016, https://hal.inria.fr/hal-01341787.

[4] D. BURSZTYN, F. GOASDOUÉ, I. MANOLESCU.*Optimizing FOL reducible query answering: understanding performance challenges*, in "ISWC 2016: The 15th International Semantic Web Conference", Kobe, Japan, October 2016, https://hal.inria.fr/hal-01400568.

[5] J. CAMACHO-RODRÍGUEZ, D. COLAZZO, M. HERSCHEL, I. MANOLESCU, S. ROY CHOWDHURY.*Reuse-based Optimization for Pig Latin*, in "25th ACM International on Conference on Information and Knowledge Management", Indianapolis, France, Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, October 2016, p. 2215 - 2220 [*DOI : 10.1145/2983323.2983669*], https://hal.inria.fr/hal-01425321.

[6] M. KRÖTZSCH, T. MASOPUST, M. THOMAZO.*On the Complexity of Universality for Partially Ordered NFAs*, in "41st International Symposium on Mathematical Foundations of Computer Science", Krakow, Poland, August 2016 [*DOI : 10.4230/LIPICS.MFCS.2016.62*], https://hal.inria.fr/hal-01334958.

[7] S. RUDOLPH, M. THOMAZO.*Expressivity of Datalog Variants – Completing the Picture*, in "25th International Joint Conference on Artificial Intelligence", New-York, United States, July 2016, https://hal.inria.fr/hal-01302832.

Conferences without Proceedings

[8] R. B. AL-OTAIBI, F. BUGIOTTI, D. BURSZTYN, A. DEUTSCH, I. MANOLESCU, S. ZAMPETAKIS.*Estocada: Stockage Hybride et Ré-écriture sous Contraintes d'Intégrité*, in "BDA: Conférence sur la Gestion de Données", Poitiers, France, November 2016, https://hal.archives-ouvertes.fr/hal-01355933.

[9] R. BONAQUE, T. D. CAO, B. CAUTIS, F. GOASDOUÉ, J. LETELIER, I. MANOLESCU, O. MENDOZA, S. RIBEIRO, X. TANNIER, M. THOMAZO.*Mixed-instance querying: a lightweight integration architecture for data journalism*, in "VLDB", New Delhi, India, September 2016, https://hal.inria.fr/hal-01321201.

[10] R. BONAQUE, B. CAUTIS, F. GOASDOUÉ, I. MANOLESCU.*Recherche Sociale, Structurée et Sémantique*, in "32ème Conférence sur la Gestion de Données - Principes, Technologies et Applications", Poitiers, France, November 2016, https://hal.inria.fr/hal-01426532.

[11] R. BONAQUE, B. CAUTIS, F. GOASDOUÉ, I. MANOLESCU.*Social, Structured and Semantic Search*, in "International Conference on Extending Database Technology", Bordeaux, France, March 2016, https://hal.inria.fr/hal-01277939.

[12] F. BUGIOTTI, D. BURSZTYN, A. DEUTSCH, I. MANOLESCU, S. ZAMPETAKIS.*Flexible Hybrid Stores: Constraint-Based Rewriting to the Rescue*, in "32nd IEEE International Conference on Data Engineering", Helsinki, Finland, May 2016, https://hal.inria.fr/hal-01321138.

[13] D. BURSZTYN, F. GOASDOUÉ, I. MANOLESCU.*Teaching an RDBMS about ontological constraints*, in "Very Large Data Bases", New Delhi, India, September 2016, https://hal.inria.fr/hal-01354592.

[14] Š. ČEBIRIĆ, F. GOASDOUÉ, I. MANOLESCU.*Query-Oriented Summarization of RDF Graphs*, in "BDA (Bases de Données Avancées)", Poitiers, France, November 2016, https://hal.inria.fr/hal-01363625.

Books or Proceedings Editing

[15] E. PITOURA, S. MAABOUT, K. GEORGIA, A. MARIAN, T. LETIZIA, I. MANOLESCU, K. STEFANIDIS (editors). *Proceedings of the 19th International Conference on Extending Database Technology, EDBT*, OpenProceedings.org, Bordeaux, France, March 2016, https://hal.archives-ouvertes.fr/hal-01285191.

Research Reports

[16] D. BURSZTYN, F. GOASDOUÉ, I. MANOLESCU.*Efficient query answering in the presence of DL-LiteR constraints*, Inria Saclay ; Inria, August 2016, n^o RR-8714, https://hal.inria.fr/hal-01143498.

[17] J. CAMACHO-RODRÍGUEZ, D. COLAZZO, M. HERSCHEL, I. MANOLESCU, S. ROY CHOWD-HURY.*PigReuse: A Reuse-based Optimizer for Pig Latin*, Inria Saclay, August 2016, https://hal.inria.fr/hal-01353891.

[18] Š. ČEBIRIĆ, F. GOASDOUÉ, I. MANOLESCU.*Query-Oriented Summarization of RDF Graphs*, Inria Saclay ; Université Rennes 1, June 2016, n^o RR-8920, https://hal.inria.fr/hal-01325900.

Project-Team COMETE

Concurrency, Mobility and Transactions

IN COLLABORATION WITH: Laboratoire d'informatique de l'école polytechnique (LIX)

IN PARTNERSHIP WITH:
CNRS

Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Security and Confidentiality

Table of contents

Project-Team COMETE

Creation of the Project-Team: 2008 January 01

Keywords:

Computer Science and Digital Science:

2.1.1. - Semantics of programming languages
2.1.5. - Constraint programming
2.1.6. - Concurrent programming
2.1.8. - Synchronous languages
2.4.1. - Analysis
2.4.2. - Model-checking
4.5. - Formal methods for security
4.8. - Privacy-enhancing technologies

Other Research Topics and Application Domains:

6.1. - Software industry
6.6. - Embedded systems
9.4.1. - Computer science
9.8. - Privacy

1. Members

Research Scientists
Catuscia Palamidessi [Team leader, Inria, Senior Researcher]
Konstantinos Chatzikokolakis [CNRS, Researcher]
Frank Valencia [CNRS, Researcher]

Technical Staff
Yamil Salim Perchy [Inria, granted by ERC (European Research Council Excecutive Agency)]

PhD Students
Tymofii Prokopenko [Inria, from Sep 2016]
Michell Guzman [Inria]
Joris Lamare [Inria]
Yamil Salim Perchy [Inria, until Nov 2016]

Post-Doctoral Fellow
Ehab Elsalamouny [Inria, until Nov 2016]

Visiting Scientists
Mario Ferreira Alvim Junior [Assistant Professor, Federal University of Minas Gerais, Dec 2016]
Annabelle Mciver [Associate Professor, Macquarie University, Dec 2016]
Charles Carroll Morgan [Professor, University of New South Wales, Dec 2016]
Geoffrey Smith [Professor, Florida International University, USA, Dec 2016]
Yusuke Kawamoto [Research Scientist, AIST, Japan, May, Sep, Nov 2016]
Linda Brodo [University of Sassari, September-October 2016]
Moreno Falashi [University of Siena, September-October 2016]
Carlos Olarte [Universidade Federal do Rio Grande do Norte, Dec 2016]

Administrative Assistants
Marie Enee [Inria, until Nov 2016]

Hélèna Kutniak [Inria]
Others
Anna Pazii [Inria, Intern, from Jul 2016]
Thomas Ragel [Inria, Intern, from Jun 2016 until Jul 2016]
Susheel Suresh [Ecole Polytechnique, Intern, from May 2016 until Aug 2016]
Tymofii Prokopenko [Inria, Intern, from Apr 2016 until Aug 2016]

2. Overall Objectives

2.1. Overall Objectives

Our times are characterized by the massive presence of highly *distributed systems* consisting of diverse and specialized devices, forming heterogeneous networks, and providing different services and applications. Revolutionary phenomena such as *social networks* and *cloud computing* are examples of such systems.

In Comète we study emerging concepts of this new era of computing. *Security* and *privacy* are some of the fundamental concerns that arise in this setting. In particular, in the modern digital world the problem of keeping information secret or confidential is exacerbated by orders of magnitude: the frequent interaction between users and electronic devices, and the continuous connection between these devices and the internet, offer malicious agents the opportunity to gather and store huge amount of information, often without the individual even being aware of it. Mobility is an additional source of vulnerability, since tracing may reveal significant information. To avoid these kinds of hazards, *security protocols* and various techniques for privacy protection have been designed. However, the properties that they are supposed to ensure are rather subtle, and, furthermore, it is difficult to foresee all possible expedients that a potential attacker may use. As a consequence, even protocols that seem at first "obviously correct" are later (often years later) found to be prone to attacks.

In addition to the security problems, the problems of correctness, robustness and reliability are made more challenging by the complexity of these systems, since they are highly concurrent and distributed. Despite being based on impressive engineering technologies, they are still prone to faulty behavior due to errors in the software design.

To overcome these drawbacks, we need to develop formalisms, reasoning techniques, and verification methods, to specify systems and protocols, their intended properties, and to guarantee that these intended properties of correctness and security are indeed satisfied.

In Comète we study formal computational frameworks for specifying these systems, theories for defining the desired properties of correctness and security and for reasoning about them, and methods and techniques for proving that a given system satisfies the intended properties.

3. Research Program

3.1. Probability and information theory

Participants: Konstantinos Chatzikokolakis, Catuscia Palamidessi, Ehab Elsalamouny, Tymofii Prokopenko, Joris Lamare.

Much of the research of Comète focuses on security and privacy. In particular, we are interested in the problem of the leakage of secret information through public observables.

Ideally we would like systems to be completely secure, but in practice this goal is often impossible to achieve. Therefore, we need to reason about the amount of information leaked, and the utility that it can have for the adversary, i.e. the probability that the adversary is able to exploit such information.

The recent tendency is to use an information theoretic approach to model the problem and define the leakage in a quantitative way. The idea is to consider the system as an information-theoretic *channel*. The input represents the secret, the output represents the observable, and the correlation between the input and output (*mutual information*) represents the information leakage.

Information theory depends on the notion of entropy as a measure of uncertainty. From the security point of view, this measure corresponds to a particular model of attack and a particular way of estimating the security threat (vulnerability of the secret). Most of the proposals in the literature use Shannon entropy, which is the most established notion of entropy in information theory. We, however, consider also other notions, in particular Rényi min-entropy, which seems to be more appropriate for security in common scenarios like one-try attacks.

3.2. Expressiveness of Concurrent Formalisms

Participants: Catuscia Palamidessi, Frank Valencia.

We study computational models and languages for distributed, probabilistic and mobile systems, with a particular attention to expressiveness issues. We aim at developing criteria to assess the expressive power of a model or formalism in a distributed setting, to compare existing models and formalisms, and to define new ones according to an intended level of expressiveness, also taking into account the issue of (efficient) implementability.

3.3. Concurrent constraint programming

Participants: Michell Guzman, Yamil Salim Perchy, Frank Valencia.

Concurrent constraint programming (ccp) is a well established process calculus for modeling systems where agents interact by posting and asking information in a store, much like in users interact in *social networks*. This information is represented as first-order logic formulae, called constraints, on the shared variables of the system (e.g., $X > 42$). The most distinctive and appealing feature of ccp is perhaps that it unifies in a single formalism the operational view of processes based upon process calculi with a declarative one based upon first-order logic. It also has an elegant denotational semantics that interprets processes as closure operators (over the set of constraints ordered by entailment). In other words, any ccp process can be seen as an idempotent, increasing, and monotonic function from stores to stores. Consequently, ccp processes can be viewed as: computing agents, formulae in the underlying logic, and closure operators. This allows ccp to benefit from the large body of techniques of process calculi, logic and domain theory.

Our research in ccp develops along the following two lines:

1. **(a)** The study of a bisimulation semantics for ccp. The advantage of bisimulation, over other kinds of semantics, is that it can be efficiently verified.

2. **(b)** The extension of ccp with constructs to capture emergent systems such as those in social networks and cloud computing.

3.4. Model checking

Participants: Konstantinos Chatzikokolakis, Catuscia Palamidessi.

Model checking addresses the problem of establishing whether a given specification satisfies a certain property. We are interested in developing model-checking techniques for verifying concurrent systems of the kind explained above. In particular, we focus on security and privacy, i.e., on the problem of proving that a given system satisfies the intended security or privacy properties. Since the properties we are interested in have a probabilistic nature, we use probabilistic automata to model the protocols. A challenging problem is represented by the fact that the interplay between nondeterminism and probability, which in security presents subtleties that cannot be handled with the traditional notion of a scheduler,

4. Application Domains

4.1. Security and privacy

Participants: Konstantinos Chatzikokolakis, Catuscia Palamidessi, Ehab Elsalamouny, Tymofii Prokopenko, Joris Lamare.

The aim of our research is the specification and verification of protocols used in mobile distributed systems, in particular security protocols. We are especially interested in protocols for *information hiding*.

Information hiding is a generic term which we use here to refer to the problem of preventing the disclosure of information which is supposed to be secret or confidential. The most prominent research areas which are concerned with this problem are those of *secure information flow* and of *privacy*.

Secure information flow refers to the problem of avoiding the so-called *propagation* of secret data due to their processing. It was initially considered as related to software, and the research focussed on type systems and other kind of static analysis to prevent dangerous operations, Nowadays the setting is more general, and a large part of the research effort is directed towards the investigation of probabilistic scenarios and treaths.

Privacy denotes the issue of preventing certain information to become publicly known. It may refer to the protection of *private data* (credit card number, personal info etc.), of the agent's identity (*anonymity*), of the link between information and user (*unlinkability*), of its activities (*unobservability*), and of its *mobility* (*untraceability*).

The common denominator of this class of problems is that an adversary can try to infer the private information (*secrets*) from the information that he can access (*observables*). The solution is then to obfuscate the link between secrets and observables as much as possible, and often the use randomization, i.e. the introduction of *noise*, can help to achieve this purpose. The system can then be seen as a *noisy channel*, in the information-theoretic sense, between the secrets and the observables.

We intend to explore the rich set of concepts and techniques in the fields of information theory and hypothesis testing to establish the foundations of quantitive information flow and of privacy, and to develop heuristics and methods to improve mechanisms for the protection of secret information. Our approach will be based on the specification of protocols in the probabilistic asynchronous π-calculus, and the application of model-checking to compute the matrices associated to the corresponding channels.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. *Notable New Projects and Contracts*

- New ANR project REPAS: Reliable and Privacy-Aware Software Systems via Bisimulation Metrics (Section 9.3.4.1)
- New industrial contract with Renault: Protection techniques for location data (Section 8.1.1)

6. New Software and Platforms

6.1. libqif - A Quantitative Information Flow C++ Toolkit Library

Participants: Konstantinos Chatzikokolakis [correspondant], Susheel Suresh, Tymofii Prokopenko.

https://github.com/chatziko/libqif

The goal of libqif is to provide an efficient C++ toolkit implementing a variety of techniques and algorithms from the area of quantitative information flow and differential privacy. We plan to implement all techniques produced by Comète in recent years, as well as several ones produced outside the group, giving the ability to privacy researchers to reproduce our results and compare different techniques in a uniform and efficient framework.

Some of these techniques were previously implemented in an ad-hoc fashion, in small, incompatible with each-other, non-maintained and usually inefficient tools, used only for the purposes of a single paper and then abandoned. We aim at reimplementing those – as well as adding several new ones not previously implemented – in a structured, efficient and maintainable manner, providing a tool of great value for future research. Of particular interest is the ability to easily re-run evaluations, experiments and case-studies from all our papers, which will be of great value for comparing new research results in the future.

The library was under constant development in 2016 with several new features added this year. The project's git repository shows for this year 77 commits by 2 contributors, containing 5697 line additions and 4067 line removals. Some of the techniques already implemented are:

- Standard leakage measures: Shannon, min-entropy, guessing entropy
- Measures from the g-leakage framework [26]
- Channel factorization
- Hyper distribution produced by a channel run under a prior
- Standard differential privacy mechanisms from the literature
- The planar Laplace mechanism of [27]
- The planar Geometric mechanism
- The tight-constraints mechanism of [29] (also with equality constraints)
- Optimal mechanism construction under DP
- The standard Kantorovich metric as well as the multiplicative variant from [28]
- Additive capacity for specific prior over all gain functions [2]
- All operations are supported for both doubles (for precision) and floats (for memory efficiency)
- All operations involving only rational quantities are supported using arbitrary precision rational arithmetic, allowing to obtain exact results
- Native linear programing for rationals
- Simple installation in OSX via Homebrew

Many more are scheduled to be added in the near future.

6.2. D-SPACES - constraint systems with space and extrusion operators

Participants: Frank Valencia, Yamil Salim Perchy.

http://www.lix.polytechnique.fr/~perchy/d-spaces/

D-SPACES is an implementation of constraint systems with space and extrusion operators. Constraint systems are algebraic models that allow for a semantic language-like representation of information in systems where the concept of space is a primary structural feature. We give this information mainly an epistemic interpretation and consider various agents as entities acting upon it. D-SPACES is coded as a c++11 library providing implementations for constraint systems, space functions and extrusion functions. The interfaces to access each implementation are minimal and thoroughly documented. D-SPACES also provides property-checking methods as well as an implementation of a specific type of constraint systems (a boolean algebra). This last implementation serves as an entry point for quick access and proof of concept when using these models. In [22] an illustrative example of using the library is given, in the form of a small social network where users post their beliefs and utter their opinions.

6.3. Trace Slicer for Timed Concurrent Constraint Programming

Participants: Catuscia Palamidessi, Carlos Olarte.

http://subsell.logic.at/slicer/

Concurrent Constraint Programming (CCP) is a declarative model for concurrency aimed at specifying reactive systems, i.e. systems that continuously interact with the environment. Some previous works have developed (approximated) declarative debuggers for CCP languages. However, the task of debugging concurrent programs remains difficult. This tool is a companion for the existing debugging techniques. Slicing in our proposal consists of considering partial computations, which show the presence of bugs. Often, the quantity of information in a trace is overwhelming, and the user gets easily lost, since she cannot focus on the sources of the bugs. Our slicer allows for marking part of the state of the computation and assists the user to eliminate most of the redundant information in order to highlight the errors. See [19] for further details.

7. New Results

7.1. 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.

7.1.1. *Axioms for Information Leakage*

Quantitative information flow aims to assess and control the leakage of sensitive information by computer systems. A key insight in this area is that no single leakage measure is appropriate in all operational scenarios; as a result, many leakage measures have been proposed, with many different properties. To clarify this complex situation, we studied in [17] information leakage axiomatically, showing important dependencies among different axioms. We also established a completeness result about the g-leakage family, showing that any leakage measure satisfying certain intuitively-reasonable properties can be expressed as a g-leakage.

7.1.2. *Up-To Techniques for Generalized Bisimulation Metrics*

Bisimulation metrics allow us to compute distances between the behaviors of probabilistic systems. In [18] we presented enhancements of the proof method based on bisimulation metrics, by extending the theory of up-to techniques to (pre)metrics on discrete probabilistic concurrent processes.

Up-to techniques have proved to be a powerful proof method for showing that two systems are bisimilar, since they make it possible to build (and thereby check) smaller relations in bisimulation proofs. We defined soundness conditions for up-to techniques on metrics, and studied compatibility properties that allow us to safely compose up-to techniques with each other. As an example, we derived the soundness of the up-to-bisimilarity-metric-and-context technique.

The study was carried out for a generalized version of the bisimulation metrics, in which the Kantorovich lifting is parametrized with respect to a distance function. The standard bisimulation metrics, as well as metrics aimed at capturing multiplicative properties such as differential privacy, are specific instances of this general definition.

7.1.3. *Compositional methods for 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 with the inference of the secret information. In [12] 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 derive a generalization of Chaum's strong anonymity result.

7.1.4. Differential Privacy Models for Location-Based Services

In [13], we considered the adaptation of differential privacy to the context of location-based services (LBSs), which personalize the information provided to a user based on his current position. Assuming that the LBS provider is queried with a perturbed version of the position of the user instead of his exact one, we relied on differential privacy to quantify the level of indistinguishability (i.e., privacy) provided by this perturbation with respect to the user's position. In this setting, the adaptation of differential privacy can lead to various models depending on the precise form of indistinguishability required. We discussed the set of properties that hold for these models in terms of privacy, utility and also implementation issues. More precisely, we first introduced and analyzed one of these models, the (D,eps)-location privacy, which is directly inspired from the standard differential privacy model. In this context, we described a general probabilistic model for obfuscation mechanisms for the locations whose output domain is the Euclidean space E^2. In this model, we characterized the satisfiability conditions of (D,eps)-location privacy for a particular mechanism and also measured its utility with respect to an arbitrary loss function. Afterwards, we presented and analyzed symmetric mechanisms in which all locations are perturbed in a unified manner through a noise function, focusing in particular on circular noise functions. We proved that, under certain assumptions, the circular functions are rich enough to provide the same privacy and utility levels as other more complex (i.e., non-circular) noise functions, while being easier to implement. Finally, we extended our results to a generalized notion for location privacy, called 'l-privacy' capturing both (D,eps)-location privacy and also the notion of geo-indistinguishability recently introduced by Andrès, Bordenabe, Chatzikokolakis and Palamidessi.

7.1.5. Practical Mechanisms for Location Privacy

The continuously increasing use of location-based services poses an important threat to the privacy of users. A natural defense is to employ an obfuscation mechanism, such as those providing geo-indistinguishability, a framework for obtaining formal privacy guarantees that has become popular in recent years.

Ideally, one would like to employ an optimal obfuscation mechanism, providing the best utility among those satisfying the required privacy level. In theory optimal mechanisms can be constructed via linear programming. In practice, however, this is only feasible for a radically small number of locations. As a consequence, all known applications of geo-indistinguishability simply use noise drawn from a planar Laplace distribution.

In [23] we studied methods for substantially improving the utility of location obfuscation, while having practical applicability as a central constraint. We provided such solutions for both infinite (continuous or discrete) as well as large but finite domains of locations, using a Bayesian remapping procedure as a key ingredient. We evaluated our techniques in two real world complete datasets, without any restriction on the evaluation area, and showed important utility improvements wrt the standard planar Laplace approach.

7.1.6. 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. In [14] we focussed on differential privacy, an approach to privacy that emerged from the area of statistical databases and is now widely applied also in other domains. In this approach, privacy is protected by adding noise to the values correlated to the private data. The typical mechanisms used to achieve differential privacy have been proved correct in the ideal case in which computations are made using infinite-precision semantics. We analyzed the situation at the implementation level, where the semantics is necessarily limited by 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, acceptable) variant of the property, under only a minor degradation of the privacy level. Finally, we illustrated our results on two examples: the standard Laplacian mechanism commonly used in differential privacy, and a bivariate version of it recently introduced in the setting of privacy-aware geolocation.

7.1.7. *Quantifying Leakage in the Presence of Unreliable Sources of Information*

Belief and min-entropy leakage are two well-known approaches to quantify information flow in security systems. Both concepts stand as alternatives to the traditional approaches founded on Shannon entropy and mutual information , which were shown to provide inadequate security guarantees. In [16] we unified the two concepts in one model so as to cope with the frequent (potentially inaccurate, misleading or outdated) attackers' side information about individuals on social networks, online forums, blogs and other forms of online communication and information sharing. To this end we proposed a new metric based on min-entropy that takes into account the adversary's beliefs.

7.1.8. *On the Compositionality of Quantitative Information Flow*

In the min-entropy approach to quantitative information flow, the leakage is defined in terms of a minimization problem, which, in the case of large systems, can be computationally rather heavy. The same happens for the recently proposed generalization called g-vulnerability. In [25] we studied the case in which the channel associated to the system can be decomposed into simpler channels, which typically happens when the observables consist of several components. Our main contribution is the derivation of bounds on the g-leakage of the whole system in terms of the g-leakages of its components. We also considered the particular cases of min-entropy leakage and of parallel channels, generalizing and systematizing results from the literature. We demonstrated the effectiveness of our method and evaluate the precision of our bounds using examples.

7.2. 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.

7.2.1. *Belief, Knowledge, Lies and Other Utterances in an Algebra for Space and Extrusion*

Spatial constraint systems are algebraic structures from concurrent constraint programming to specify spatial and epistemic behavior in multi-agent system. In [15], [11] we developed the theory of spatial constraint systems with operators to specify information and processes moving from a space to another. We investigated the properties of this new family of constraint systems and illustrated their applications. From a computational point of view the new operators provide for process/information extrusion, a central concept in formalisms for mobile communication. From an epistemic point of view extrusion corresponds to a notion we called utterance; a piece of information that an agent communicates to others but that may be inconsistent with the agent's beliefs. Utterances can then be used to express instances of epistemic notions such as hoaxes or intentional lies. Spatial constraint system can express the epistemic notion of belief by means of space functions that specify local information. We showed that spatial constraint can also express the epistemic notion of knowledge by means of a derived spatial operator that specifies global information. In [21] we reported on our progress using spatial constraint system as an abstract representation of modal and epistemic behaviour.

7.2.2. *Deriving Inverse Operators for Modal Logic*

In [20] we used spatial constraint systems to give an abstract characterization of the notion of normality in modal logic and to derive right inverse/reverse operators for modal languages. In particular, we identified the weakest condition for the existence of right inverses and showed that the abstract notion of normality corresponds to the preservation of finite suprema. We applied our results to existing modal languages such as the weakest normal modal logic, Hennessy-Milner logic, and linear-time temporal logic. We also discussed our results in the context of modal concepts such as bisimilarity and inconsistency invariance.

7.2.3. D-SPACES: Implementing Declarative Semantics for Spatially Structured Information

In [22] we introduced D-SPACES, an implementation of constraint systems with space and extrusion operators. D-SPACES is coded as a c++11 library providing implementations for constraint systems, space functions and extrusion functions. D-SPACES provides property-checking methods as well as an implementation of a specific type of constraint systems (boolean algebras). We illustrated the implementation with a small social network where users post their beliefs and utter their opinions.

7.2.4. Slicing Concurrent Constraint Programs

Concurrent Constraint Programming (CCP) is a declarative model for concurrency where agents interact by telling and asking constraints (pieces of information) in a shared store. Some previous works have developed (approximated) declarative debuggers for CCP languages. However, the task of debugging concurrent programs remains difficult. In [19] we defined a dynamic slicer for CCP and we showed it to be a useful companion tool for the existing debugging techniques. Our technique starts by considering a partial computation (a trace) that shows the presence of bugs. Often, the quantity of information in such a trace is overwhelming, and the user gets easily lost, since she cannot focus on the sources of the bugs. Our slicer allows for marking part of the state of the computation and assists the user to eliminate most of the redundant information in order to highlight the errors. We showed that this technique can be tailored to timed variants of CCP. We also developed a prototypical implementation freely available for making experiments.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. Contract with Renault

Project title: Protection techniques for location data

Duration: July 2016 - December 2016

Budget: 38K euros, financed by Renault

Coordinator: Catuscia Palamidessi, Inria Saclay, EPI Comète

Abstract: The goal of this project is to produce a survey of the state of the art methods for protecting location data, as well as a protytype showing the application of some of these methods in the context of a "connected car".

Stage: A six month intern (Anna Pazii) was funded by this project.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. Projects funded by Digiteo-DigiCosme

9.1.1.1. OPTIMEC

Project title: Optimal Mechanisms for Privacy Protection

Duration: September 2016 - August 2019

Coordinator: Catuscia Palamidessi, Inria Saclay, EPI Comète

Other PI's: Serge Haddadm ENS Cachan.

Abstract: In this project we plan to investigate classes of utility and privacy measures, and to devise methods to obtain optimal mechanisms with respect to the trade-off between utility and privacy. In order to represent the probabilistic knowledge of the adversary and of the user, and the fact that mechanisms themselves can be randomized, we will consider a probabilistic setting. We will focus, in particular, on measures that are expressible as linear functions of the probabilities.

9.1.1.2. D-SPACES

Project title: D-spaces : Distributed Spaces in Concurrent Epistemic Systems

Duration: Nov 2013 - Oct 2016

Coordinator: Frank Valencia, CNRS-LIX and Inria Saclay, EPI Comète

Other PI's: Stefan Haar ENS Cachan.

Abstract: In this project we developed an innovative and expressive computational model for these systems that coherently combines techniques for the analysis of concurrent systems such as process calculi with epistemic and spatial formalisms.

9.2. National Initiatives

9.2.1. Large-scale initiatives

Project acronym: CAPPRIS

Project title: Collaborative Action on the Protection of Privacy Rights in the Information Society

Duration: September 2013 - December 2016

URL: https://cappris.inria.fr/

Coordinator: Daniel Le Metayer, Inria Grenoble

Other partner institutions: The project involves four Inria research centers (Saclay, Saphia-Antipolis, Rennes and Grenoble), CNRS-LAAS, Eurecom and the university of Namur. Besides computer scientists, the consortium also includes experts in sociology and in law, thus covering the complementary areas of expertise required to reach the objectives.

Abstract: The goal of this project is to study the challenges related to privacy in the modern information society, trying to consider not only the technical, but also the social and legal ones, and to develop methods to enhance the privacy protection.

9.3. International Initiatives

9.3.1. Inria-MSR joint lab

9.3.1.1. Privacy-Friendly Services and Apps

Title: Privacy-Friendly Services and Applications

Inria principal investigator: Catuscia Palamidessi

International Partners:

Cedric Fournet, Microsoft Research Lab, Cambridge, UK

Andy Gordon, Microsoft Research Lab, Cambridge, UK

Duration: 2014 - 2016

URL: http://www.msr-inria.fr/projects/privacy-friendly-services-and-apps/

Abstract: This is a project sponsored by Microsoft Research Lab, on methods to preserve privacy in web services and location-based services.

9.3.2. Inria Associate Teams

9.3.2.1. LOGIS

Title: Logical and Formal Methods for Information Security

Inria principal investigator: Konstantinos Chatzikokolakis

International Partners:

>   Mitsuhiro Okada, Keio University (Japan)

>   Yusuke Kawamoto, AIST (Japan)

>   Tachio Terauchi, JAIST (Japan)

>   Masami Hagiya, University of Tokyo (Japan)

Start year: 2016

URL: http://www.lix.polytechnique.fr/~kostas/projects/logis/

Abstract: The project aims at integrating the logical / formal approaches to verify security protocols with (A) complexity theory and (B) information theory. The first direction aims at establishing the foundations of logical verification for security in the computational sense, with the ultimate goal of automatically finding attacks that probabilistic polynomial-time adversaries can carry out on protocols. The second direction aims at developing frameworks and techniques for evaluating and reducing information leakage caused by adaptive attackers.

9.3.3. Inria International Partners

9.3.3.1. Informal International Partners

>   Geoffrey Smith, Florida International University (United States)

>   Carroll Morgan, NICTA (Australia)

>   Annabelle McIver, Maquarie University (Australia)

>   Moreno Falaschi, Professor, University of Siena, Italy

>   Mario Ferreira Alvim Junior, Assistant Professor, Federal University of Minas Gerais, Brazil

>   Camilo Rueda, Professor, Universidad Javeriana Cali, Colombia

9.3.4. Participation in Other International Programs

9.3.4.1. REPAS

>   Program: ANR Blanc

>   Project title: Reliable and Privacy-Aware Software Systems via Bisimulation Metrics

>   Duration: October 2016 - September 2021

>   Coordinator: Catuscia Palamidessi, Inria Saclay, EPI Comète

>   Other PI's and partner institutions: Ugo del Lago, Inria Sophia Antipolis (EPI Focus) and University of Bologna (Italy). Vincent Danos, ENS Paris. Filippo Bonchi, ENS Lyon.

>   Abstract: In this project, we aim at investigating quantitative notions and tools for proving program correctness and protecting privacy. In particular, we will focus on bisimulation metrics, which are the natural extension of bisimulation on quantitative systems. As a key application, we will develop a mechanism to protect the privacy of users when their location traces are collected.

9.3.4.2. PACE

>   Program: ANR Blanc International

>   Project title: Beyond plain Processes: Analysis techniques, Coinduction and Expressiveness

>   Duration: January 2013 - December 2016

>   URL: http://perso.ens-lyon.fr/daniel.hirschkoff/pace/

>   Coordinator: Daniel Hirschkoff, Ecole Normale Supérieure de Lyon

Other PI's and partner institutions: Catuscia Palamidessi, Inria Saclay, Frank Valencia, CNRS-LIX and Inria Saclay (France). Davide Sangiorgi, University of Bologna (Italy). Yuxi Fu, Shanghai Jiao Tong University (China).

Abstract: This project objective is to enrich and adapt these methods, techniques, and tools to much broader forms of interactive models, well beyond the realm of "traditional" processes.

9.3.4.3. LOCALI

Program: ANR Blanc International

Project title: Logical Approach to Novel Computational Paradigms

Duration: January 2012 - December 2016

URL: http://www.agence-nationale-recherche.fr/?Project=ANR-11-IS02-0002

Coordinator: Gilles Dowek, Inria Rocquencourt

Other PI's and partner institutions: Catuscia Palamidessi, Inria Saclay. Thomas Erhard, Paris VII. Ying Jiang , Chinese Academy of Science in Beijin (China).

Abstract: This project aims at exploring the interplays between logic and sequential/distributed computation in formalisms like the lambda calculus and the π calculus. Going back to the fundamentals of the definitions of these calculi, the project plans to design new programming languages and proof systems via a logical approach.

9.3.4.4. MUSICAL

Program: CNPq Science Without Borders.

Project title: Music and Spatial Interaction with Constraints, Algebra and Logic: Foundations and Applications.

Duration: Oct 2014 - Oct 2016

URL: http://cic.puj.edu.co/~caolarte/musical/Musical/Welcome.html

Coordinator: Elaine Pimentel, Universidade Federal do Rio Grande do Norte (Brazil),

Other PI's and partner institutions: Camilo Rueda, PUJ Cali (Colombia). Carlos Olarte, Universidade Federal do Rio Grande do Norte (Brazil). Frank Valencia, CNRS-LIX and Inria Saclay (France). Gerard Assayag, IRCAM (France).

Abstract: This multi-disciplinary project aims to develop and integrate tools from logic and concurrency theory for the design and analysis of reactive systems and to their application to musical processes and multimedia systems.

9.3.4.5. CLASSIC

Program: Colciencias - Conv. 712.

Project title: Concurrency, Logic and Algebra for Social and Spatial Interactive Computation.

Duration: Oct 2016 - Oct 2019

URL: http://goo.gl/Gv6Lij

Coordinator: Camilo Rueda PUJ Cali (Colombia).

Other PI's and partner institutions: Carlos Olarte, Universidade Federal do Rio Grande do Norte (Brazil). Frank Valencia, CNRS-LIX and Inria Saclay (France).

Abstract:This project will advance the state of the art of domains such as mathematical logic, order theory and concurrency for reasoning about spatial and epistemic behaviour in multi-agent systems..

9.4. International Research Visitors

9.4.1. Visits of International Scientists

Mario Ferreira Alvim Junior, Assistant Professor, Federal University of Minas Gerais, Brazil, Dec 2016

Annabelle McIver, Associate Professor, Macquarie University, Australia, Dec 2016

Carroll Morgan, Professor, University of New South Wales and NICTA, Australia, Dec 2016

Geoffrey Smith, Professor, Florida International University, USA, Dec 2016

Camilo Rueda, Professor, PUJ Cali, Colombia, May 2016 and Nov 2016.

Camilo Rocha, Professor, PUJ Cali, Colombia, Oct 2016.

9.4.2. Visits to International Teams

Catuscia Palamidessi visited the Computer Security team of Roberto Focardi at the University of Venice, Italy, from 4 April to 30 April, 2016.

10. Dissemination

10.1. Promoting Scientific Activities

Note: In this section we include only the activities of the permanent internal members of Comète.

10.1.1. Scientific events organisation

10.1.1.1. Member of the organizing committee

Catuscia Palamidessi is member of:

The Executive Committee of SIGLOG, the ACM Special Interest Group on Logic and Computation. Since 2014.

The Organizing Committee of LICS, the ACM/IEEE Symposium on Logic in Computer Science. Since 2010.

The Council of EATCS, the European Association for Theoretical Computer Science. Since 2005.

The Steering Committee of ETAPS, the European Joint Conferences on Theory and Practice of Software. Since 2006.

The Steering Committee of EACSL, the European Association for Computer Science Logics. Since 2015.

The Steering Committee of CONCUR, the International Conference in Concurrency Theory. Since 2016.

The Steering Committee of FORTE, the International Conference on Formal Techniques for Distributed Objects, Components, and Systems. Since 2014.

The IFIP Technical Committee 1 – Foundations of Computer Science. Since 2007.

The IFIP Working Group 2.2 – Formal Description of Programming Concepts. Since 2001.

The IFIP Working Group 1.7 – Theoretical Foundations of Security Analysis and Design. Since 2010.

Frank D. Valencia is member of:

The steering committee of the International Workshop in Concurrency EXPRESS. Since 2010.

10.1.2. Scientific events selection

10.1.2.1. Member of conference program committees

Catuscia Palamidessi is/has been a member of the program committees of the following conferences and workshops:

ICTAC 2017. The 14th International Colloquium on Theoretical Aspects of Computing. Hanoi, Vietnam, 23-27 October 2017.

TASE 2017. The 11th International Symposium on Theoretical Aspects of Software Engineering. Nice, France, 13-15 September 2017.

CONCUR 2017. The 28th International Conference on Concurrency Theory. Berlin, Germany, 5-8 September 2017.

CSL 2017. The 26th EACSL Annual Conference on Computer Science Logic. Stockholm, Sweden, 20-25 August 2017.

ICSOFT-PT 2017. The 12th International Conference on Software Paradigm Trends. Lisbon, Portugal, 24-26 July 2017.

ICALP 2017 (Track B). The 44th International Colloquium on Automata, Languages, and Programming. Warsaw, Poland, 10–14 July 2017.

FORTE 2017. The 37th IFIP International Conference on Formal Techniques for Distributed Objects, Components, and Systems. Neuchâtel, Switzerland, 19-22 June 2017.

CSR 2017. The 12th International Computer Science Symposium in Russia. Kazan, Russia, 8–12 June 2017.

ICTAC 2016. The 13th International Colloquium on Theoretical Aspects of Computing. Taipei, Taiwan, 24-31 October 2016.

LOPSTR 2016. The 26th International Symposium on Logic-Based Program Synthesis and Transformation, 6-8 September 2016.

CONCUR 2016. The 27th International Conference on Concurrency Theory. Québec City, Canada, 23-26 August 2016.

TASE 2016. The 10th International Symposium on Theoretical Aspects of Software Engineering. Shanghai, China, 17-19 July 2016.

FCS 2016. The Workshop on Foundations of Computer Security. Lisbon, Portugal, 27 June 2016.

MFPS XXXII. The Thirty-second Conference on the Mathematical Foundations of Programming Semantics. Carnegie Mellon University, Pittsburgh, USA, 23-26 May 2016.

PhDs in Logic VIII. Darmstadt, Germany, 9-11 May 2016.

UEOP 2016. The 1st Workshop on Understanding and Enhancing Online Privacy. San Diego, USA, 21 February 2016.

Konstantinos Chatzikokolakis is/has been a member of the program committees of the following conferences and workshops:

ICDE 2017: IEEE International Conference on Data Engineering

CSF 2017: 30th IEEE Computer Security Foundations Symposium

POST 2017: 6th International Conference on Principles of Security and Trust

BIGQP 2017: International Workshop on Big Geo Data Quality and Privacy

PETS 2016: The 16th Privacy Enhancing Technologies Symposium

WWW 2016: 25th World Wide Web conference

APVP 2016: 7ème Atelier sur la Protection de la Vie Privée

Frank D. Valencia is/has been a member of the program committees of the following conferences and workshops:

PPDP 2016. The 18th International Symposium on Principles and Practice of Declarative Programming (PPDP 2016).

ICTAC 2016. The 13th International Colloquium on Theoretical Aspects of Computing (ICTAC 2016).

ICLP DC 2016. 12th ICLP Doctoral Consortium.

10.1.2.2. Reviewer

The members of the team reviewed several papers for international conferences and workshops.

10.1.3. Journals

10.1.3.1. Member of the editorial board

Catuscia Palamidessi is:

Member of the Editorial Board of Mathematical Structures in Computer Science, published by the Cambridge University Press.

Member of the Editorial Board of Acta Informatica, published by Springer.

Member of the Editorial Board of the Electronic Notes of Theoretical Computer Science, published by Elsevier Science.

Member of the Editorial Board of LIPIcs: Leibniz International Proceedings in Informatics, Schloss Dagstuhl –Leibniz Center for Informatics.

Konstantinos Chatzikokolakis is:

Editorial board member of the newly established Proceedings on Privacy Enhancing Technologies (PoPETs), a scholarly journal for timely research papers on privacy.

10.1.3.2. Reviewer

The members of the team reviewed several papers for international journals.

10.1.4. Other Editorial Activities

Frank D. Valencia has been:

Co-editor of the special issue on Mathematical Structures in Computer Science dedicated to the best papers from the 12th International Colloquium on Theoretical Aspects of Computing.

10.1.5. Other Activities

10.1.5.1. Invited talks

Catuscia Palamidessi has given invited talks at the following conferences and workshops:

DISCOTEC 2016 (Keynote speaker). The 11th International Federated Conference on Distributed Computing Techniques. Crete, Greece, 6-9 June 2016.

Journée sur la Securité, la Sureté et la Confidentialité. Organized by Paris VII, Paris XIII and Systematic. Paris, France, 10 May 2016.

10.1.5.2. Participation in other committees

Catuscia Palamidessi has been serving in the following committees:

Member of the Alonzo Church Award Committee. Since 2015. This award is for an outstanding contribution to Logic and Computation within the past 25 years.

President of the selection committee for the EATCS Best Paper Award at the ETAPS conferences. Since 2006.

10.1.5.3. Service

Catuscia Palamidessi has served as:

Reviewer for the projects proposal for the program PRIN, sponsored by the Italian MIUR ("Ministero dell'Istruzione, dell'Università e della Ricerca"). Since 2004.

Member of the comité de selection for a position for Maitre de Conferences at l'Université de Paris VII (Paris Diderot). Spring 2016.

Frank Valencia has served as:

Directeur adjoint de l'UMR 7161, le Laboratoire d'Informatique de l'Ecole Polytechnique (LIX). May 2016 - .

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

PhD : Catuscia Palamidessi has been teaching a course for PhD students, on Protection of sensitive information, at the University of Venice, Italy. April 2016. Total 30 hours.

Master : Frank D. Valencia has been teaching the undergraduate course "Computability", 45 hours, at the Pontificia Universidad Javeriana de Cali, Colombia. July 27 - Nov 1, 2016.

Master : Frank D. Valencia has been teaching the masters course "Foundations of Computer Science", 45 hours, at the Pontificia Universidad Javeriana de Cali, Colombia. Jan 27 - Jun 1, 2016.

Master: Konstantinos Chatzikokolakis and Catuscia Palamidessi have been teaching a course on the Foundations of Privacy at the MPRI, the Master Parisien pour la Recherche en Informatique. University of Paris VII. A.Y. 2016-17. Total: 24 hours plus 6 hours for the exam and the exercise session is preparation to the exam.

10.2.2. Supervision

PhD in progress (2016-) Tymofii Prokopenko. Ecole Polytechnique and ENS Cachan. Grant Digiteo-Digicosme. Co-supervised by Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Serge Haddad.

PhD in progress (2015-) Joris Lamare. Ecole Polytechnique. Grant MSR Center. Co-supervised by Catuscia Palamidessi and Konstantinos Chatzikokolakis.

PhD in progress (2014-) Michel Guzman. Ecole Polytechnique. Grant Inria CORDI-S. Co-supervised by Catuscia Palamidessi and Frank D. Valencia.

PhD completed (2013-16) Salim Percy. Ecole Polytechnique. Grant Digiteo-Digicosme. Co-supervised by Frank D. Valencia and Stefan Haar.

10.2.3. Juries

Catuscia Palamidessi has been reviewer and member of the board at the PhD defense for the thesis of the following PhD student:

Huu-Hiep Nguyen, PhD student supervised by Abdessamad Imine, University of Lorraine, France. November 2016. Title of the thesis: Social Graph Anonymization.

10.2.4. Other didactical duties

Catuscia Palamidessi is:

External member of the scientific council for the PhD in Computer Science at the University of Pisa, Italy. Since 2012.

Member of the Committee d'Encadrement de Thèse of Jun Wang (PhD student supervised by Qiang Tang and Peter Ryan), University of Luxembourg. Since December 2014.

Member of the advising committee for the PhD of Andrea Margheri (PhD student supervised by Rosario Pugliese), University of Florence, Italy. 2014-16.

Konstantinos Chatzikokolakis and Catuscia Palamidessi have designed, and coordinate, a course on the Foundations of Privacy at the MPRI, the Master Parisien pour la Recherche en Informatique. University of Paris VII. A.Y. 2016-17.

11. Bibliography

Major publications by the team in recent years

[1] M. S. ALVIM, M. E. ANDRÉS, K. CHATZIKOKOLAKIS, P. DEGANO, C. PALAMIDESSI.*On the information leakage of differentially-private mechanisms*, in "Journal of Computer Security", 2015, vol. 23, n° 4, p. 427-469 [*DOI : 10.3233/JCS-150528*], https://hal.inria.fr/hal-00940425.

[2] M. S. ALVIM, K. CHATZIKOKOLAKIS, A. MCIVER, C. MORGAN, C. PALAMIDESSI, G. SMITH.*Additive and multiplicative notions of leakage, and their capacities*, in "27th Computer Security Foundations Symposium (CSF 2014)", Vienna, Austria, IEEE, July 2014, p. 308–322 [*DOI : 10.1109/CSF.2014.29*], https://hal.inria.fr/hal-00989462.

[3] M. ALVIM, K. CHATZIKOKOLAKIS, C. PALAMIDESSI, G. SMITH.*Measuring Information Leakage using Generalized Gain Functions*, in "Computer Security Foundations", Cambridge MA, United States, IEEE, 2012, p. 265-279 [*DOI : 10.1109/CSF.2012.26*], http://hal.inria.fr/hal-00734044.

[4] M. ANDRÉS, N. BORDENABE, K. CHATZIKOKOLAKIS, C. PALAMIDESSI.*Geo-Indistinguishability: Differential Privacy for Location-Based Systems*, in "20th ACM Conference on Computer and Communications Security", Berlin, Allemagne, ACM Press, 2013, p. 901-914, DGA, Inria large scale initiative CAPPRIS [*DOI : 10.1145/2508859.2516735*], http://hal.inria.fr/hal-00766821.

[5] A. ARISTIZÁBAL, F. BONCHI, C. PALAMIDESSI, L. PINO, D. VALENCIA.*Deriving Labels and Bisimilarity for Concurrent Constraint Programming*, in "FOSSACS 2011 : 14th International Conference on Foundations of Software Science and Computational Structures", Saarbrücken, Germany, M. HOFMANN (editor), Lecture Notes in Computer Science, Springer, March 2011, vol. 6604, p. 138-152 [*DOI : 10.1007/ISBN 978-3-642-19804-5*], https://hal.archives-ouvertes.fr/hal-00546722.

[6] N. E. BORDENABE, K. CHATZIKOKOLAKIS, C. PALAMIDESSI.*Optimal Geo-Indistinguishable Mechanisms for Location Privacy*, in "CCS - 21st ACM Conference on Computer and Communications Security", Scottsdale, Arizona, United States, G.-J. AHN, M. YUNG, N. LI (editors), Proceedings of the 21st ACM Conference on Computer and Communications Security, ACM, November 2014, p. 251-262 [*DOI : 10.1145/2660267.2660345*], https://hal.inria.fr/hal-00950479.

[7] K. CHATZIKOKOLAKIS, M. ANDRÉS, N. BORDENABE, C. PALAMIDESSI.*Broadening the Scope of Differential Privacy Using Metrics*, in "The 13th Privacy Enhancing Technologies Symposium", Bloomington, Indiana, États-Unis, E. DE CRISTOFARO, M. WRIGHT (editors), Springer, 2013, vol. 7981, p. 82-102, DGA, Inria large scale initiative CAPPRIS [*DOI : 10.1007/978-3-642-39077-7*], http://hal.inria.fr/hal-00767210.

[8] K. CHATZIKOKOLAKIS, C. PALAMIDESSI, M. STRONATI.*Constructing elastic distinguishability metrics for location privacy*, in "Proceedings on Privacy Enhancing Technologies", June 2015, vol. 2015, n° 2, p. 156-170 [*DOI : 10.1515/POPETS-2015-0023*], https://hal.inria.fr/hal-01270197.

[9] M. GUZMAN, S. HAAR, S. PERCHY, C. RUEDA, F. VALENCIA.*Belief, Knowledge, Lies and Other Utterances in an Algebra for Space and Extrusion*, in "Journal of Logical and Algebraic Methods in Programming", September 2016 [*DOI :* 10.1016/J.JLAMP.2016.09.001], https://hal.inria.fr/hal-01257113.

[10] S. KNIGHT, C. PALAMIDESSI, P. PANANGADEN, F. D. VALENCIA.*Spatial and Epistemic Modalities in Constraint-Based Process Calculi*, in "CONCUR 2012 - Concurrency Theory - 23rd International Conference, CONCUR 2012", Newcastle upon Tyne, United Kingdom, September 2012, vol. 7454, p. 317-332 [*DOI :* 10.1007/978-3-642-32940-1], http://hal.inria.fr/hal-00761116.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] S. PERCHY.*Opinions, Lies and Knowledge. An Algebraic Approach to Mobility of Information and Processes*, Université Paris-Saclay, October 2016, https://hal.inria.fr/tel-01413970.

Articles in International Peer-Reviewed Journal

[12] K. CHATZIKOKOLAKIS, C. PALAMIDESSI, C. BRAUN.*Compositional methods for information-hiding*, in "Mathematical Structures in Computer Science", September 2016, vol. 26, n° 6, p. 908-932 [*DOI :* 10.1017/S0960129514000292], https://hal.inria.fr/hal-01006384.

[13] E. ELSALAMOUNY, S. GAMBS.*Differential Privacy Models for Location- Based Services*, in "Transactions on Data Privacy", 2016, vol. 9, n° 1, p. 15 - 48, https://hal.inria.fr/hal-01418136.

[14] I. GAZEAU, D. MILLER, C. PALAMIDESSI.*Preserving differential privacy under finite-precision semantics*, in "Journal of Theoretical Computer Science (TCS)", 2016, https://hal.inria.fr/hal-01390927.

[15] M. GUZMAN, S. HAAR, S. PERCHY, C. RUEDA, F. VALENCIA.*Belief, Knowledge, Lies and Other Utterances in an Algebra for Space and Extrusion*, in "Journal of Logical and Algebraic Methods in Programming", September 2016 [*DOI :* 10.1016/J.JLAMP.2016.09.001], https://hal.inria.fr/hal-01257113.

[16] S. HAMADOU, C. PALAMIDESSI, V. SASSONE.*Quantifying Leakage in the Presence of Unreliable Sources of Information*, in "Journal of Computer and System Sciences", 2017, forthcoming, https://hal.inria.fr/hal-01421417.

International Conferences with Proceedings

[17] M. S. ALVIM, K. CHATZIKOKOLAKIS, A. MCIVER, C. MORGAN, C. PALAMIDESSI, G. SMITH.*Axioms for Information Leakage*, in "29th Computer Security Foundations Symposium (CSF 2016)", Lisbon, Portugal, IEEE, June 2016, 16, https://hal.inria.fr/hal-01330414.

[18] K. CHATZIKOKOLAKIS, C. PALAMIDESSI, V. VIGNUDELLI.*Up-To Techniques for Generalized Bisimulation Metrics*, in "27th International Conference on Concurrency Theory (CONCUR 2016)", Québec City, Canada, Leibniz International Proceedings in Informatics (LIPIcs), August 2016, vol. 59, p. 35:1–35:14 [*DOI :* 10.4230/LIPICS.CONCUR.2016.35], https://hal.inria.fr/hal-01335234.

[19] M. FALASCHI, M. GABBRIELLI, C. OLARTE, C. PALAMIDESSI.*Slicing Concurrent Constraint Programs*, in "Pre-proceedings of the 26th International Symposium on Logic-Based Program Synthesis and Transformation

(LOPSTR 2016)", Edinburgh, United Kingdom, M. V. HERMENEGILDO, P. LOPEZ-GARCIA (editors), 2016, https://hal.inria.fr/hal-01421407.

[20] M. GUZMAN, S. PERCHY, C. RUEDA, F. VALENCIA.*Deriving Inverse Operators for Modal Logic*, in "Theoretical Aspects of Computing – ICTAC 2016", Taipei, Taiwan, A. SAMPAIO, F. WANG (editors), Theoretical Aspects of Computing – ICTAC 2016, Springer, October 2016, vol. 9965, p. 214-232 [*DOI :* 10.1007/978-3-319-46750-4_13], https://hal.inria.fr/hal-01328188.

[21] M. GUZMÁN, F. D. VALENCIA.*On the Expressiveness of Spatial Constraint Systems* , in "ICLP 2016 - Proceedings of the Technical Communications of the 32nd International Conference on Logic Programming", New York, United States, M. CARRO, A. KING, N. SAEEDLOEI, M. D. VOS (editors), Technical Communications of the 32nd International Conference on Logic Programming (ICLP 2016), John Gallagher and Germán Vidal, October 2016, vol. OASICS Vol. 52, p. 16:1 - 16:12, Appears as a Technical Communications of the 32nd International Conference on Logic Programming (ICLP 2016) [*DOI :* 10.4230/OASIcs.ICLP.2016.16], https://hal.inria.fr/hal-01418166.

[22] S. HAAR, S. PERCHY, F. VALENCIA.*D-SPACES: Implementing Declarative Semantics for Spatially Structured Information*, in "11th International Conference on Semantic Computing", San Diego, California, United States, IEEE ICSC 2017, IEEE, January 2017, vol. 11, https://hal.inria.fr/hal-01328189.

Other Publications

[23] K. CHATZIKOKOLAKIS, E. ELSALAMOUNY, C. PALAMIDESSI.*Practical Mechanisms for Location Privacy*, 2016, working paper or preprint, https://hal.inria.fr/hal-01422842.

[24] K. CHATZIKOKOLAKIS, C. PALAMIDESSI, A. PAZII.*Methods for Location Privacy: A comparative overview*, 2016, Submitted for publication, https://hal.inria.fr/hal-01421457.

[25] Y. KAWAMOTO, K. CHATZIKOKOLAKIS, C. PALAMIDESSI.*On the Compositionality of Quantitative Information Flow*, 2016, Submitted for publication to Logical Methods in Computer Science, https://hal.inria.fr/hal-01421424.

References in notes

[26] M. ALVIM, K. CHATZIKOKOLAKIS, C. PALAMIDESSI, G. SMITH.*Measuring Information Leakage using Generalized Gain Functions*, in "Computer Security Foundations", Cambridge MA, United States, IEEE, 2012, p. 265-279 [*DOI :* 10.1109/CSF.2012.26], http://hal.inria.fr/hal-00734044.

[27] M. ANDRÉS, N. BORDENABE, K. CHATZIKOKOLAKIS, C. PALAMIDESSI.*Geo-Indistinguishability: Differential Privacy for Location-Based Systems*, in "20th ACM Conference on Computer and Communications Security", Berlin, Allemagne, ACM Press, 2013, p. 901-914, DGA, Inria large scale initiative CAPPRIS [*DOI :* 10.1145/2508859.2516735], http://hal.inria.fr/hal-00766821.

[28] K. CHATZIKOKOLAKIS, D. GEBLER, C. PALAMIDESSI, L. XU.*Generalized bisimulation metrics*, in "CONCUR - 25th Conference on Concurrency Theory", Rome, Italy, P. BALDAN, D. GORLA (editors), Lecture Notes in Computer Science, Springer, September 2014, vol. 8704, p. 32-46 [*DOI :* 10.1007/978-3-662-44584-6_4], https://hal.inria.fr/hal-01011471.

[29] E. ELSALAMOUNY, K. CHATZIKOKOLAKIS, C. PALAMIDESSI.*Generalized differential privacy: regions of priors that admit robust optimal mechanisms*, in "Horizons of the Mind. A Tribute to Prakash Panangaden", F.

VAN BREUGEL, E. KASHEFI, C. PALAMIDESSI, J. RUTTEN (editors), Lecture Notes in Computer Science, Springer International Publishing, 2014, vol. 8464, p. 292-318 [*DOI :* 10.1007/978-3-319-06880-0_16], https://hal.inria.fr/hal-01006380.

Project-Team COMMANDS

Control, Optimization, Models, Methods and Applications for Nonlinear Dynamical Systems

IN COLLABORATION WITH: Centre de Mathématiques Appliquées (CMAP), Unité de Mathématiques Appliquées (UMA - ENSTA)

IN PARTNERSHIP WITH:
CNRS

Ecole Polytechnique

Ecole nationale supérieure des techniques avancées

RESEARCH CENTER
Saclay - Île-de-France

THEME
Optimization and control of dynamic systems

Table of contents

Project-Team COMMANDS

Creation of the Project-Team: 2009 January 01

Keywords:

Computer Science and Digital Science:
6.2.1. - Numerical analysis of PDE and ODE
6.2.6. - Optimization
6.2.7. - High performance computing
6.3.2. - Data assimilation
6.4.1. - Deterministic control
6.4.2. - Stochastic control

Other Research Topics and Application Domains:
4.4.1. - Smart grids
7.1.2. - Road traffic
7.1.3. - Air traffic
7.2.1. - Smart vehicles

1. Members

Research Scientists
J. Frédéric Bonnans [Team leader, Inria, Senior Researcher, HDR]
Axel Kröner [Inria, Starting Research position]
Pierre Martinon [Inria, Researcher]

Technical Staff
Jinyan Liu [Inria, from Feb 2016]

PhD Students
Benjamin Heymann [Ecole Polytechnique, until Sep 2016]
Arthur Le Rhun [Ifpen, from Apr 2016]
Cédric Rommel [Safety Line, granted by CIFRE]
Justina Gianatti [U. Rosario (Argentina), intern, until Jul 2016]

Administrative Assistant
Jessica Gameiro [Inria]

Other
Luis Alberto Croquevielle [Inria, intern, until Mar 2016]

2. Overall Objectives

2.1. Scientific directions

Commands is a team devoted to dynamic optimization, both for deterministic and stochastic systems. This includes the following approaches: trajectory optimization, deterministic and stochastic optimal control, stochastic programming, dynamic programming and Hamilton-Jacobi-Bellman equation.

Our aim is to derive new and powerful algorithms for solving numerically these problems, with applications in several industrial fields. While the numerical aspects are the core of our approach it happens that the study of convergence of these algorithms and the verification of their well-posedness and accuracy raises interesting and difficult theoretical questions, such as, for trajectory optimization: qualification conditions and second-order optimality condition, well-posedness of the shooting algorithm, estimates for discretization errors; for the Hamilton-Jacobi-Bellman approach: accuracy estimates, strong uniqueness principles when state constraints are present, for stochastic programming problems: sensitivity analysis.

2.2. Industrial impact

For many years the team members have been deeply involved in various industrial applications, often in the framework of PhD theses. The Commands team itself has dealt since its foundation in 2009 with several types of applications:

- Space vehicle trajectories, in collaboration with CNES, the French space agency.
- Aeronautics, in collaboration with the startup Safety Line.
- Production, management, storage and trading of energy resources, in collaboration with EDF, GDF and TOTAL.
- Energy management for hybrid vehicles, in collaboration with Renault and IFPEN.

We give more details in the Bilateral contracts section.

3. Research Program

3.1. Historical aspects

The roots of deterministic optimal control are the "classical" theory of the calculus of variations, illustrated by the work of Newton, Bernoulli, Euler, and Lagrange (whose famous multipliers were introduced in [45]), with improvements due to the "Chicago school", Bliss [32] during the first part of the 20th century, and by the notion of relaxed problem and generalized solution (Young [51]).

Trajectory optimization really started with the spectacular achievement done by Pontryagin's group [50] during the fifties, by stating, for general optimal control problems, nonlocal optimality conditions generalizing those of Weierstrass. This motivated the application to many industrial problems (see the classical books by Bryson and Ho [38], Leitmann [47], Lee and Markus [46], Ioffe and Tihomirov [43]).

Dynamic programming was introduced and systematically studied by R. Bellman during the fifties. The HJB equation, whose solution is the value function of the (parameterized) optimal control problem, is a variant of the classical Hamilton-Jacobi equation of mechanics for the case of dynamics parameterized by a control variable. It may be viewed as a differential form of the dynamic programming principle. This nonlinear first-order PDE appears to be well-posed in the framework of *viscosity solutions* introduced by Crandall and Lions [39]. The theoretical contributions in this direction did not cease growing, see the books by Barles [30] and Bardi and Capuzzo-Dolcetta [29].

3.2. Trajectory optimization

The so-called *direct methods* consist in an optimization of the trajectory, after having discretized time, by a nonlinear programming solver that possibly takes into account the dynamic structure. So the two main problems are the choice of the discretization and the nonlinear programming algorithm. A third problem is the possibility of refinement of the discretization once after solving on a coarser grid.

In the *full discretization approach*, general Runge-Kutta schemes with different values of control for each inner step are used. This allows to obtain and control high orders of precision, see Hager [42], Bonnans [35]. In an interior-point algorithm context, controls can be eliminated and the resulting system of equation is easily solved due to its band structure. Discretization errors due to constraints are discussed in Dontchev et al. [40]. See also Malanowski et al. [48].

In the *indirect* approach, the control is eliminated thanks to Pontryagin's maximum principle. One has then to solve the two-points boundary value problem (with differential variables state and costate) by a single or multiple shooting method. The questions are here the choice of a discretization scheme for the integration of the boundary value problem, of a (possibly globalized) Newton type algorithm for solving the resulting finite dimensional problem in IR^n (n is the number of state variables), and a methodology for finding an initial point.

For state constrained problems or singular arcs, the formulation of the shooting function may be quite elaborate [33], [34], [28]. As initiated in [41], we focus more specifically on the handling of discontinuities, with ongoing work on the geometric integration aspects (Hamiltonian conservation).

3.3. Hamilton-Jacobi-Bellman approach

This approach consists in calculating the value function associated with the optimal control problem, and then synthesizing the feedback control and the optimal trajectory using Pontryagin's principle. The method has the great particular advantage of reaching directly the global optimum, which can be very interesting when the problem is not convex.

Characterization of the value function >From the dynamic programming principle, we derive a characterization of the value function as being a solution (in viscosity sense) of an Hamilton-Jacobi-Bellman equation, which is a nonlinear PDE of dimension equal to the number n of state variables. Since the pioneer works of Crandall and Lions [39], many theoretical contributions were carried out, allowing an understanding of the properties of the value function as well as of the set of admissible trajectories. However, there remains an important effort to provide for the development of effective and adapted numerical tools, mainly because of numerical complexity (complexity is exponential with respect to n).

Optimal stochastic control problems occur when the dynamical system is uncertain. A decision typically has to be taken at each time, while realizations of future events are unknown (but some information is given on their distribution of probabilities). In particular, problems of economic nature deal with large uncertainties (on prices, production and demand). Specific examples are the portfolio selection problems in a market with risky and non-risky assets, super-replication with uncertain volatility, management of power resources (dams, gas). Air traffic control is another example of such problems.

Nonsmoothness of the value function. Sometimes the value function is smooth and the associated HJB equation can be solved explicitly. Still, the value function is not smooth enough to satisfy the HJB equation in the classical sense. As for the deterministic case, the notion of viscosity solution provides a convenient framework for dealing with the lack of smoothness, see Pham [49], that happens also to be well adapted to the study of discretization errors for numerical discretization schemes [44], [31].

For solving stochastic control problems, we studied the so-called Generalized Finite Differences (GFD), that allow to choose at any node, the stencil approximating the diffusion matrix up to a certain threshold [37]. Determining the stencil and the associated coefficients boils down to a quadratic program to be solved at each point of the grid, and for each control. This is definitely expensive, with the exception of special structures where the coefficients can be computed at low cost. For two dimensional systems, we designed a (very) fast algorithm for computing the coefficients of the GFD scheme, based on the Stern-Brocot tree [36].

4. Application Domains

4.1. Fuel saving by optimizing airplanes trajectories

We have a collaboration with the startup Safety Line on the optimization of trajectories for civil aircrafts. Key points include the reliable identification of the plane parameters (aerodynamic and thrust models) using data from the flight recorders, and the robust trajectory optimization of the climbing and cruise phases. We use both local (quasi-Newton interior-point algorithms) and global optimization tools (dynamic programming).

4.2. Hybrid vehicles

We started a collaboration with IFPEN on the energy management for hybrid vehicles. A significant direction is the analysis and classification of traffic data. We have preliminary results on the choice of the routing which amounts to some type of constrained shortest path.

5. Highlights of the Year

5.1. Highlights of the Year

We started at the beginning of 2016 an Innovation Lab (Ilab) 'OSCAR', jointly with the startup Safety Line. The subject of the Ilab is the design of algorithmic tools for the (i) identification of aircraft dynamics, based on flight data recorders, and (ii) the computation of energy efficient flight trajectories.

6. New Software and Platforms

6.1. BOCOP

Boite à Outils pour le Contrôle OPtimal
KEYWORDS: Energy management - Numerical optimization - Biology - Identification - Dynamic Optimization - Transportation
FUNCTIONAL DESCRIPTION

Bocop is an open-source toolbox for solving optimal control problems, with collaborations with industrial and academic partners. Optimal control (optimization of dynamical systems governed by differential equations) has numerous applications in transportation, energy, process optimization, energy and biology. Bocop includes a module for parameter identification and a graphical interface, and runs under Linux / Windows / Mac.

- Participants: Joseph Frédéric Bonnans, Pierre Martinon, Benjamin Heymann and Jinyan Liu
- Contact: Pierre Martinon
- URL: http://bocop.org

6.2. Bocop Avion

KEYWORDS: Optimization - Aeronautics
FUNCTIONAL DESCRIPTION

Optimize the climb speeds and associated fuel consumption for the flight planning of civil airplanes.

- Participants: Joseph Frédéric Bonnans, Pierre Martinon, Stéphan Maindrault, Cindie Andrieu, Pierre Jouniaux and Karim Tekkal
- Contact: Pierre Martinon
- URL: http://www.safety-line.fr

6.3. Bocop HJB

- Participants: Joseph Frédéric Bonnans, Pierre Martinon, Benjamin Heymann and Jinyan Liu
- Contact: Joseph Frédéric Bonnans
- URL: http://bocop.org

7. New Results

7.1. Optimal control of ordinary and partial differential equations

7.1.1. *On the Design of Optimal Health Insurance Contracts under Ex Post Moral Hazard*
Participant: Pierre Martinon.

With Pierre Picard and Anasuya Raj, Ecole Polytechnique.

We analyze in [27] the design of optimal medical insurance under ex post moral hazard, i.e., when illness severity cannot be observed by insurers and policyholders decide on their health expenditures. We characterize the trade-o§ between ex ante risk sharing and ex post incentive compatibility, in an optimal revelation mechanism under hidden information and risk aversion. We establish that the optimal contract provides partial insurance at the margin, with a deductible when insurersí rates are a§ected by a positive loading, and that it may also include an upper limit on coverage. We show that the potential to audit the health state leads to an upper limit on out-of-pocket expenses.

7.1.2. *Optimal control of infinite dimensional bilinear systems: application to the heat and wave equations*
Participants: J. Frédéric Bonnans, Axel Kröner.

With Soledad Aronna, FGV, Rio de Janeiro. In this paper [13] we consider second order optimality conditions for a bilinear optimal control problem governed by a strongly continuous semigroup operator, the control entering linearly in the cost function. We derive first and second order optimality conditions, taking advantage of the Goh transform. We then apply the results to the heat and wave equations.

7.1.3. *Optimal control of PDEs in a complex space setting; application to the Schrödinger equation*
Participants: J. Frédéric Bonnans, Axel Kröner.

With Soledad Aronna, FGV, Rio de Janeiro. This paper [22] presents some optimality conditions for abstract optimization problems over complex spaces. We then apply these results to optimal control problems with a semigroup structure. As an application we detail the case when the state equation is the Schrödinger one, with pointwise constraints on the "bilinear'" control. We derive first and second order optimality conditions and address in particular the case that the control enters the state equation and cost function linearly.

7.1.4. *Approximation and reduction of optimal control problems in infinite dimension*
Participant: Axel Kröner.

With Michael D. Chekroun, UCLA) and H. Liu, Virginia Tech. Nonlinear optimal control problems in infinite dimensions are considered for which we establish approximation theorems and reduction procedures. Approximation theorems and reduction procedures are available in the literature. The originality of our approach relies on a combination of Galerkin approximation techniques with reduction techniques based on finite-horizon parameterizing manifolds. The numerical approximation of the control in a feedback form based on Hamilton-Jacobi-Equation become also affordable within this approach. The approach is applied to optimal control problems of delay differential equations and nonlinear parabolic equations.

7.2. Stochastic control, electricity production and planning

7.2.1. *MIDAS: A Mixed Integer Dynamic Approximation Scheme*
Participant: J. Frédéric Bonnans.

With Andy Philpott and Faisal Wahid, U. Auckland. Mixed Integer Dynamic Approximation Scheme (MI-DAS) [23] is a new sampling-based algorithm for solving finite-horizon stochastic dynamic programs with monotonic Bellman functions. MIDAS approximates these value functions using step functions, leading to stage problems that are mixed integer programs. We provide a general description of MIDAS, and prove its almost-sure convergence to an epsilon-optimal policy when the Bellman functions are known to be continuous, and the sampling process satisfies standard assumptions.

7.2.2. *Long term aging : an adaptative weights dynamic programming algorithm*
Participants: J. Frédéric Bonnans, Benjamin Heymann, Pierre Martinon.

We introduce [26] a class of optimal control problems with periodic data. A state variable that we call the age of the system represents the negative impact of the operations on the system qualities over time: other things being equal, older systems have higher operating costs. Many industrial problems relate to this class. If we envision to perform an optimization over a large number of periods, there is a tradeoff between minimizing repeatedly the one-period criterion in a short sighted way and taking into account the impact of the decision on the aging speed (which modifies the minimal one period criterion). In general, because the aging process is slow, short term optimization strategies-such as one period sliding horizon strategies-either neglect it or use rule-of-thumb penalization terms in the criterion, which leads to suboptimal solutions. On the other hand, for most applications it is unrealistic to envision a brute-force numerical resolution by dynamic programming of the long term problem because of the computation burden. We introduce a two-scale method to reduce this computation burden. The method relies on Lagrangian duality and some monotony properties. We expose the theoretical foundations of the method and discuss some practical aspects: approximation errors, asymptotic estimation, computation burden, possible extensions, etc. Since our initial motivation was the difficulty to take long term battery aging in Energy Management Systems into account, we implement the method on a toy long term microgrid energy management problem.

7.2.3. *Continuous Optimal Control Approaches to Microgrid Energy Management*
Participants: J. Frédéric Bonnans, Benjamin Heymann, Pierre Martinon.

With Francisco Silva XLIM, U. Limoges, Fernando Lanas and Guillermo Jimenez, U. Chile.

We propose in [18] a novel method for the microgrid energy management problem by introducing a continuous-time, rolling horizon formulation. The energy management problem is formulated as a deterministic optimal control problem (OCP). We solve (OCP) with two classical approaches: the direct method [1], and Bellman's Dynamic Programming Principle (DPP) [2]. In both cases we use the optimal control toolbox BOCOP [3] for the numerical simulations. For the DPP approach we implement a semi-Lagrangian scheme [4] adapted to handle the optimization of switching times for the on/off modes of the diesel generator. The DPP approach allows for an accurate modeling and is computationally cheap. It finds the global optimum in less than 3 seconds, a CPU time similar to the Mixed Integer Linear Programming (MILP) approach used in [5]. We achieve this performance by introducing a trick based on the Pontryagin Maximum Principle (PMP). The trick increases the computation speed by several orders and also improves the precision of the solution. For validation purposes, simulation are performed using datasets from an actual isolated microgrid located in northern Chile. Results show that DPP method is very well suited for this type of problem when compared with the MILP approach.

7.2.4. *A Stochastic Continuous Time Model for Microgrid Energy Management*
Participants: J. Frédéric Bonnans, Benjamin Heymann.

With Francisco Silva XLIM U. Limoges, Guillermo Jimenez, U. Chile.

We propose in [20] a novel stochastic control formulation for the microgrid energy management problem and extend previous works on continuous time rolling horizon strategy to uncertain demand. We modelize the demand dynamics with a stochastic differential equation. We decompose this dynamics into three terms: an average drift, a time-dependent mean-reversion term and a Brownian noise. We use BOCOPHJB for the numerical simulations. This optimal control toolbox implements a semi-Lagrangian scheme and handle the optimization of switching times required for the discrete on/off modes of the diesel generator. The scheme allows for an accurate modelling and is computationally cheap as long as the state dimension is small. As described in previous works, we use a trick to reduce the search of the optimal control values to six points. This increases the computation speed by several orders. We compare this new formulation with the deterministic control approach using data from an isolated microgrid located in northern Chile.

7.2.5. *Mechanism Design and Auctions for Electricity Network*
Participant: Benjamin Heymann.

With Alejandro Jofré, CMM - Center for Mathematical Modeling, U. Chile, Santiago. We present in [25] some key aspects of wholesale electricity markets modeling and more specifically focus our attention on auctions and mechanism design. Some of the results arising from those models are the computation of an optimal allocation for the Independent System Operator, the study of the equilibria (existence and unicity in particular) and the design of mechanisms to increase the social surplus. From a more general perspective, this field of research provides clues to discuss how wholesale electricity market should be regulated. We start with a general introduction and then present some results the authors obtained recently. We also briefly expose some undergoing related work. As an illustrative example, a section is devoted to the computation of the Independent System Operator response function for a symmetric binodal setting with piece-wise linear production cost functions.

7.2.6. *Mechanism design and allocation algorithms for network markets with piece-wise linear costs and externalities*
Participant: Benjamin Heymann.

With Alejandro Jofré, CMM - Center for Mathematical Modeling, U. Chile, Santiago. In [24], motivated by market power in electricity market, we introduce a mechanism design for simplified markets of two agents with linear production cost functions. In standard procurement auctions, the market power resulting from the quadratic transmission losses allow the producers to bid above their true value (i.e. production cost). The mechanism proposed in the previous paper reduces the producers margin to the society benefit. We extend those results to a more general market made of a finite number of agents with piecewise linear cost functions, which make the problem more difficult, but at the same time more realistic. We show that the methodology works for a large class of externalities. We also provide two algorithms to solve the principal allocation problem.

7.2.7. *Variational analysis for options with stochastic volatility and multiple factors*
Participants: J. Frédéric Bonnans, Axel Kröner.

In this ongoing work we discuss the variational analysis for stochastic volatility models with correlation and their applications for the pricing equations for European options is discussed. The considered framework is based on weigthed Sobolev spaces. Furthermore, to verify continuity of the rate term in the pricing equation an approach based on commutator analysis is developed.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. *Ifpen*

In the framework of the PhD thesis of Arthur Le Rhun, we study the energy management of hybrid (parallel) vehicles, and more specifically the optimal use of the thermal engine. Before the PhD, a 4-month internship was focused on the eco-routing problem for hybrid vehicles, ie computing the optimal path. We proposed a method based on graphs: the road network is defined by a graph, and to take into account the hybrid aspect of the vehicle, we dicretized the State of Charge on each node. Then a simple shortest path algorithm (A*) applied to this extended graph is able to solve the routing problem. Numerical simulations indicate that the solution of our discrete eco-routing problem converges to the correct solution when a sufficiently fine discretization of SoC is used. We illustrate the method on the Ille-et-Vilaine department, see Fig. 1 and Table 1. The main disadvantage of the method is the increasingly large computation time when the size of the extended graph grows.
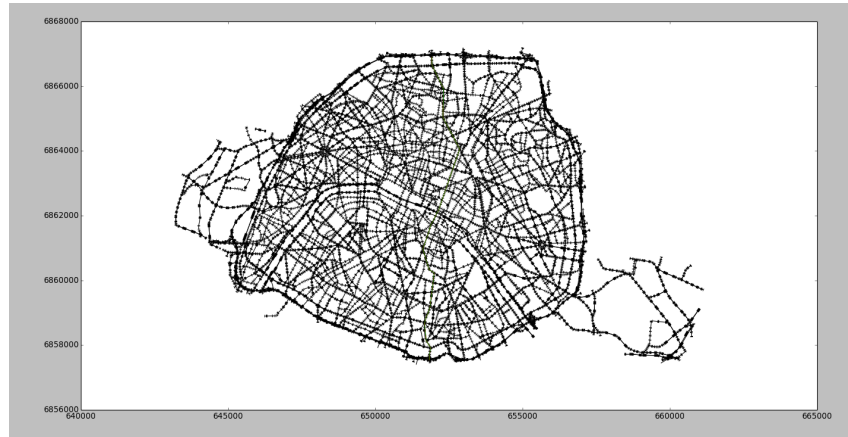
Figure 1.

Table 1. Results on the Ille-et-Vilaine department over 100 simulations

SoC disc.	improved cases	Fuel savings	CPU time (s)
3	19%	0.9753	6.03
5	65%	0.8531	14.64
10	88%	0.5831	52.80
20	88%	0.4222	283.43

8.1.2. *Safety Line*

In the framework of an Ilab with Safety Line (a startup in aeronautics), we design tools for the optimization of fuel consumption for civil planes. A first part is devoted to the identification of the aerodynamic and thrust characteristics of the plane, using recorded data from hundreds of flights. Fig. 2 shows the drag and lift coefficients for a Boeing 737, as functions of Mach and angle of attack. A second part is optimizing the fuel consumption during the climb and cruise phases. Fig. 3 shows a simulated climb phase, along with recorded data from the actual flight. This collaboration relies significantly on the toolboxes BOCOP and BOCOPHJB developed by Commands since 2010.

9. Partnerships and Cooperations

9.1. Regional Initiatives

- Gaspard Monge Program for Optimization and Operational Research (Fondation Jacques Hadamard)

Title	:	Optimal control of partial differential equations using parameterizing manifolds, model reduction, and dynamic programming,
Funding	:	9,000 Euro (for 2015-16), 10,000 Euro (for 2016-17)
PI	:	Axel Kröner
Period	:	$2015 - 2017$
Further members	:	Frédéric Bonnans (Inria Saclay and CMAP, École Polytechnique), Mickaël Chekroun (UCLA, Los Angeles), Martin Gubisch (University of Konstanz), Karl Kunisch (University of Graz), Hasnaa Zidani (ENSTA ParisTech).
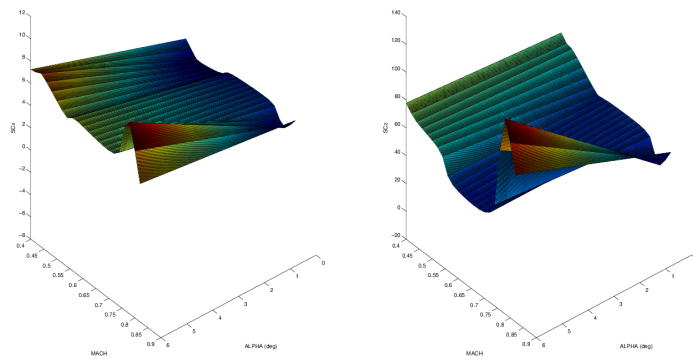
Figure 2.
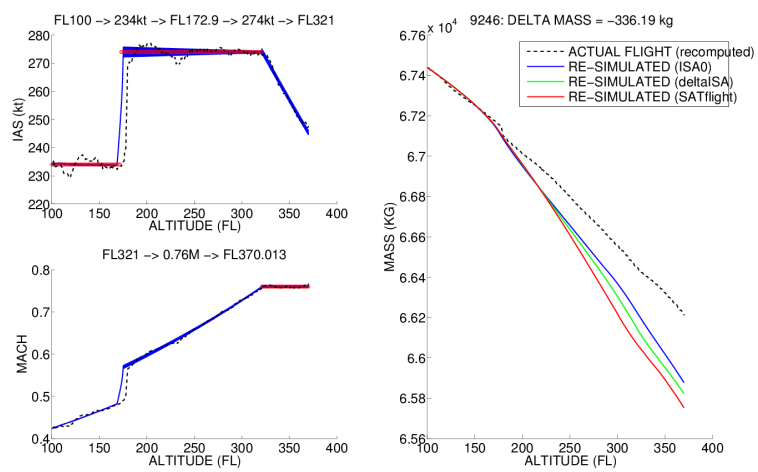
Figure 3.

9.2. International Initiatives

9.2.1. Inria International Partners

9.2.1.1. Informal International Partners

- Michael D. Chekroun, U.C.L.A, collaboration on the approximation and reduction of optimal control problems in infinite dimension.
- Alejandro Jofré, CMM, U. Chile, Santiago de Chile. Cosupervision of B. Heymann's PhD thesis.
- Pablo Lotito, U. Tandil, Argentina, supervision of Justina Gianatti's PhD.

9.3. International Research Visitors

9.3.1. Visits of International Scientists

- M. Chekroun (University of California, Los Angeles), 12.-14.12.2016.
- Johannes Pfefffer (Technische Universität München), 12.-14.12.2016.

9.3.1.1. Internships

- Luis Alberto Croquevielle Rendic: Classification of probability measures based on Optimal Transportation theory. January-March 2016. U. Catolica, Santiago, Chile.
- Justina Gianatti, Discretization of stochastic control problems, U. Rosario (Argentina), May-July 2016.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific events selection

10.1.1.1. Member of the conference program committees

- F. Bonnans, 14th EUROPT Workshop on Advances in Continuous Optimization, Poznań, July 3-6, 2016.

10.1.2. Journal

10.1.2.1. Member of the editorial boards

- F. Bonnans: Corresponding Editor of "ESAIM:COCV" (Control, Optimization and Calculus of Variations), and Associate Editor of "Applied Mathematics and Optimization", "Optimization, Methods and Software", and "Series on Mathematics and its Applications, Annals of The Academy of Romanian Scientists".

10.1.2.2. Reviewer - Reviewing activities

Reviews for major journals in the field such as Applied Mathematics and Optimization, Automatica, J. Diff. Equations, J. of Optimization Theory and Applications the SIAM J. Optimization, SIAM J. Control and Optimization, Inverse problems, Journal of Numerical Mathematics, Operations Research, Optimization, Process Control, Math. Reviews.

10.1.3. Invited talks

- A. Kröner: Seminars in U. Konstanz, U. Hamburg, 2016.

- Minisymposium Numerical methods for time-dependent transportation and optimal control problems. Computational Methods in Applied Mathematics (CMAM), Jyväskyla, Finland, July 31-Aug. 6,2016;

- Minisymposium 'Optimal Control - Theory and Applications', Emerging Trends in Applied Mathematics and Mechanics, Perpignan, May 30-June 3, 2016;

- Minisymposium 'Numerical aspects of controllability of PDEs and inverse problems', CANUM (Congrès d'Analyse Numérique), Obernai, May 9-13, 2016;

- A. Kröner: Workshop on *Numerical methods for Hamilton-Jacobi equations in optimal control and related fields*, Linz, Austria, Nov., 2016;

10.1.4. Leadership within the scientific community

- F. Bonnans: French representative to the IFIP-TC5 committee (International Federation of Information Processing; TC7 devoted to System Modeling and Optimization).

- F. Bonnans: member of the PGMO board and Steering Committee (Gaspard Monge Program for Optimization and Operations Research, EDF-FMJH).

- F. Bonnans: member of the Broyden Prize committee (from the Journal Optimization Methods and Software).

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master :
F. Bonnans: *Numerical analysis of partial differential equations arising in finance and stochastic control*, 36h, M2, Ecole Polytechnique and U. Paris 6, France.
F. Bonnans: *Optimal control*, 15h, M2, Optimization master (U. Paris-Saclay) and Ensta, France.
F. Bonnans: *Stochastic optimization*, 15h, M2, Optimization master (U. Paris-Saclay), France.
A. Kröner : Optimal control of partial differential equations, 20h, M2, Optimization master (U. Paris-Saclay), France.

E-learning F. Bonnans, several lecture notes on the page
http://www.cmap.polytechnique.fr/~bonnans/notes.html

10.2.2. Supervision

- PhD : Benjamin Heymann, Dynamic optimization with uncertainty; application to energy production. Polytechnique fellowship, defense October 2016, F. Bonnans and A. Jofre.

- PhD in progress : Cédric Rommel, Data exploration for the optimization of aircraft trajectories. Started November 2015, F. Bonnans and P. Martinon. CIFRE fellowship (Safety Line).

- PhD in progress : Arthur Le Rhun, Optimal and robust control of hybrid vehicles (IFPEN fellowship), started Sept. 2016, F. Bonnans and P. Martinon.

10.2.3. Juries

- HDR Juries: F. Silva (Limoges), A. Rondepierre (Toulouse, rapporteur).

10.3. Popularization

- J.F. Bonnans: *Comment optimiser la gestion d'un micro-réseau électrique intelligent ?* Cahier de l'Institut Louis Bachelier N. 23 (2016), p. 12-13.

11. Bibliography

Major publications by the team in recent years

[1] A. AFTALION, J. F. BONNANS.*Optimization of running strategies based on anaerobic energy and variations of velocity*, in "SIAM Journal of Applied Mathematics", October 2014, vol. 74, n$^{\text{o}}$ 5, p. 1615-1636, https://hal.inria.fr/hal-00851182.

[2] M. S. ARONNA, J. F. BONNANS, B. S. GOH.*Second order analysis of state-constrained control-affine problems*, in "Mathematical Programming, Series A", January 2016, vol. 160, n$^{\text{o}}$ 1, p. 115-147, https://hal.inria.fr/hal-01081111.

[3] M. S. ARONNA, J. F. BONNANS, A. KRÖNER.*Optimal Control of Infinite Dimensional Bilinear Systems: Application to the Heat and Wave Equations*, in "Mathematical Programming B", November 2016, 32, https://hal.inria.fr/hal-01273496.

[4] M. S. ARONNA, J. F. BONNANS, P. MARTINON.*A Shooting Algorithm for Optimal Control Problems with Singular Arcs*, in "Journal of Optimization Theory and Applications", August 2013, vol. 158, n$^{\text{o}}$ 2, p. 419-459 [*DOI :* 10.1007/S10957-012-0254-8], https://hal.inria.fr/inria-00631332.

[5] J. F. BONNANS, X. DUPUIS, L. PFEIFFER.*Second-order necessary conditions in Pontryagin form for optimal control problems*, in "SIAM Journal on Control and Optimization", 2014, vol. 52, n$^{\text{o}}$ 6, p. 3887-3916 [*DOI :* 10.1137/130923452], https://hal.inria.fr/hal-00825273.

[6] J. F. BONNANS, X. TAN.*A model-free no-arbitrage price bound for variance options*, in "Applied Mathematics and Optimization", July 2013, vol. 68, n$^{\text{o}}$ 1, p. 43-73 [*DOI :* 10.1007/S00245-013-9197-1], https://hal.inria.fr/inria-00634387.

[7] B. BONNARD, M. CLAEYS, O. COTS, P. MARTINON.*Geometric and numerical methods in the contrast imaging problem in nuclear magnetic resonance*, in "Acta Applicandae Mathematicae", February 2015, vol. 135, n$^{\text{o}}$ 1, p. 5-45 [*DOI :* 10.1007/S10440-014-9947-3], https://hal.inria.fr/hal-00867753.

[8] L. GIRALDI, P. MARTINON, M. ZOPPELLO.*Optimal Design for Purcell Three-link Swimmer*, in "Physical Review", February 2015, vol. 91, n$^{\text{o}}$ 2, 023012, https://hal.archives-ouvertes.fr/hal-01098501.

[9] A. KRÖNER, K. KUNISCH, B. VEXLER.*Semismooth Newton methods for optimal control of the wave equation with control constraints*, in "SIAM Journal on Control and Optimization", 2011, vol. 49, n$^{\text{o}}$ 2, p. 830–858, http://dx.doi.org/10.1137/090766541.

[10] A. KRÖNER, K. KUNISCH, H. ZIDANI.*Optimal feedback control of undamped wave equations by solving a HJB equation*, in "ESAIM: Control, Optimisation and Calculus of Variations", 2014, vol. 21, n$^{\text{o}}$ 2, p. 442 - 464 [*DOI :* 10.1051/COCV/2014033], https://hal.archives-ouvertes.fr/hal-00924089.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] B. HEYMANN.*Mathematical contributions for the optimization and regulation of electricity production*, Université Paris Saclay, September 2016, https://hal.archives-ouvertes.fr/tel-01416404.

Articles in International Peer-Reviewed Journal

[12] M. S. ARONNA, J. F. BONNANS, B. S. GOH.*Second order analysis of state-constrained control-affine problems*, in "Mathematical Programming, Series A", January 2016, vol. 160, nº 1, p. 115-147, https://hal.inria.fr/hal-01081111.

[13] M. S. ARONNA, J. F. BONNANS, A. KRÖNER.*Optimal Control of Infinite Dimensional Bilinear Systems: Application to the Heat and Wave Equations*, in "Mathematical Programming B", November 2016, 32, https://hal.inria.fr/hal-01273496.

[14] I. BEN LATIFA, J. F. BONNANS, M. MNIF.*Numerical methods for an optimal multiple stopping problem*, in "Stochastics and Dynamics", January 2016, vol. 16, nº 4, 27 [*DOI :* 10.1142/S0219493716500167], https://hal.inria.fr/hal-01248282.

[15] F. J. BONNANS, J. GIANATTI, F. J. SILVA.*On the convergence of the Sakawa-Shindo algorithm in stochastic control*, in "Mathematical Control and Related Fields", September 2016 [*DOI :* 10.3934/MCRF.2016008], https://hal-unilim.archives-ouvertes.fr/hal-01148272.

[16] J. GARCKE, A. KRÖNER.*Suboptimal feedback control of PDEs by solving HJB equations on adaptive sparse grids*, in "Journal of Scientific Computing", 2016, https://hal.archives-ouvertes.fr/hal-01185912.

[17] C. HERMOSILLA.*Legendre Transform and Applications to Finite and Infinite Optimization*, in "Set-Valued and Variational Analysis", March 2016 [*DOI :* 10.1007/s11228-016-0368-5], https://hal.inria.fr/hal-01055917.

[18] B. HEYMANN, J. F. BONNANS, P. MARTINON, F. SILVA, F. LANAS, G. JIMENEZ.*Continuous Optimal Control Approaches to Microgrid Energy Management*, in "Energy Systems", 2016, https://hal.inria.fr/hal-01129393.

[19] D. KALISE, A. KRÖNER, K. KUNISCH.*Local minimization algorithms for dynamic programming equations*, in "SIAM Journal on Scientific Computing", 2016, vol. 38, nº 3, https://hal.archives-ouvertes.fr/hal-01120450.

International Conferences with Proceedings

[20] B. HEYMANN, J. FRÉDÉRIC BONNANS, F. SILVA, G. JIMENEZ.*A Stochastic Continuous Time Model for Microgrid Energy Management*, in "ECC2016", Aalborg, Denmark, June 2016, https://hal.archives-ouvertes.fr/hal-01314078.

Scientific Books (or Scientific Book chapters)

[21] F. BONNANS, S. GAUBERT.*Recherche opérationnelle. Aspects mathématiques et applications*, Ellipse, March 2016, 391, https://hal.inria.fr/hal-01422645.

Research Reports

[22] M. S. ARONNA, J. F. BONNANS, A. KRÖNER.*Optimal control of PDEs in a complex space setting; application to the Schrödinger equation*, Inria, 2016, https://hal.archives-ouvertes.fr/hal-01311421.

[23] A. PHILPOTT, F. WAHID, F. BONNANS.*MIDAS: A Mixed Integer Dynamic Approximation Scheme*, Inria
Saclay Ile de France, June 2016, 22, https://hal.inria.fr/hal-01401950.

Other Publications

[24] B. HEYMANN, A. JOFRÉ.*Mechanism design and allocation algorithms for network markets with piece-wise
linear costs and externalities*, December 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-
01416411.

[25] B. HEYMANN, A. JOFRÉ.*Mechanism Design and Auctions for Electricity Network*, May 2016, working paper
or preprint, https://hal.archives-ouvertes.fr/hal-01315844.

[26] B. HEYMANN, P. MARTINON, F. BONNANS.*Long term aging : an adaptative weights dynamic programming
algorithm*, July 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01349932.

[27] P. MARTINON, P. PICARD, A. RAJ.*On the Design of Optimal Health Insurance Contracts under Ex
Post Moral Hazard*, July 2016, working paper or preprint, https://hal-polytechnique.archives-ouvertes.fr/hal-
01348551.

References in notes

[28] M. S. ARONNA, J. F. BONNANS, P. MARTINON.*A Shooting Algorithm for Optimal Control Problems with
Singular Arcs*, in "Journal of Optimization Theory and Applications", 2013, Inria Report RR-7763, 2011.

[29] M. BARDI, I. CAPUZZO-DOLCETTA.*Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman
equations*, Systems and Control: Foundations and Applications, Birkhäuser, Boston, 1997.

[30] G. BARLES.*Solutions de viscosité des équations de Hamilton-Jacobi*, Mathématiques et Applications,
Springer, Paris, 1994, vol. 17.

[31] G. BARLES, E. JAKOBSEN.*Error bounds for monotone approximation schemes for Hamilton-Jacobi-Bellman
equations*, in "SIAM J. Numerical Analysis", 2005, vol. 43, n° 2, p. 540–558.

[32] G. BLISS.*Lectures on the Calculus of Variations*, University of Chicago Press, Chicago, Illinois, 1946.

[33] J. F. BONNANS, A. HERMANT.*Well-Posedness of the Shooting Algorithm for State Constrained Optimal
Control Problems with a Single Constraint and Control*, in "SIAM J. Control Optimization", 2007, vol. 46,
n° 4, p. 1398–1430.

[34] J. F. BONNANS, A. HERMANT.*Second-order Analysis for Optimal Control Problems with Pure State
Constraints and Mixed Control-State Constraints*, in "Annales de l'Institut Henri Poincaré. Analyse non
linéaire", 2009, vol. 26, n° 2, p. 561-598.

[35] J. F. BONNANS, J. LAURENT-VARIN.*Computation of order conditions for symplectic partitioned Runge-Kutta
schemes with application to optimal control*, in "Numerische Mathematik", 2006, vol. 103, n° 1, p. 1–10.

[36] J. F. BONNANS, E. OTTENWAELTER, H. ZIDANI.*Numerical schemes for the two dimensional second-order
HJB equation*, in "ESAIM: M2AN", 2004, vol. 38, p. 723-735.

[37] J. F. BONNANS, H. ZIDANI.*Consistency of generalized finite difference schemes for the stochastic HJB equation*, in "SIAM J. Numerical Analysis", 2003, vol. 41, p. 1008-1021.

[38] A. E. BRYSON, Y.-C. HO.*Applied optimal control*, Hemisphere Publishing, New-York, 1975.

[39] M. CRANDALL, P. LIONS.*Viscosity solutions of Hamilton Jacobi equations*, in "Bull. American Mathematical Society", 1983, vol. 277, p. 1–42.

[40] A. DONTCHEV, W. HAGER, V. VELIOV.*Second-order Runge-Kutta approximations in control constrained optimal control*, in "SIAM Journal on Numerical Analysis", 2000, vol. 38, p. 202–226.

[41] J. GERGAUD, P. MARTINON.*Using switching detection and variational equations for the shooting method*, in "Optimal Control Applications and Methods", 2007, vol. 28, n$^{\text{o}}$ 2, p. 95–116.

[42] W. HAGER.*Runge-Kutta methods in optimal control and the transformed adjoint system*, in "Numerische Mathematik", 2000, vol. 87, n$^{\text{o}}$ 2, p. 247–282.

[43] A. IOFFE, V. TIHOMIROV.*Theory of Extremal Problems*, North-Holland Publishing Company, Amsterdam, 1979, Russian Edition: Nauka, Moscow, 1974.

[44] N. KRYLOV.*On the rate of convergence of finite-difference approximations for Bellman's equations with variable coefficients*, in "Probability Theory and Related Fields", 2000, vol. 117, p. 1–16.

[45] J.-L. LAGRANGE.*Mécanique analytique*, Paris, 1788, reprinted by J. Gabay, 1989.

[46] E. LEE, L. MARKUS.*Foundations of optimal control theory*, John Wiley, New York, 1967.

[47] G. LEITMANN.*An introduction to optimal control*, Mc Graw Hill, New York, 1966.

[48] K. MALANOWSKI, C. BÜSKENS, H. MAURER.*Convergence of approximations to nonlinear optimal control problems*, in "Mathematical programming with data perturbations", New York, Lecture Notes in Pure and Appl. Math., Dekker, 1998, vol. 195, p. 253–284.

[49] H. PHAM.*Optimisation et Contrôle Stochastique Appliqués à la Finance*, Springer Verlag, 2007, n$^{\text{o}}$ 61.

[50] L. PONTRYAGIN, V. BOLTYANSKI, R. GAMKRELIDZE, E. MICHTCHENKO.*The Mathematical Theory of Optimal Processes*, Wiley Interscience, New York, 1962.

[51] L. YOUNG.*Lectures on the calculus of variations and optimal control theory*, W. B. Saunders Co., Philadelphia, 1969.

Project-Team DAHU

Verification in databases

IN COLLABORATION WITH: Laboratoire specification et vérification (LSV)

IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Cachan

RESEARCH CENTER
Saclay - Île-de-France

THEME
Data and Knowledge Representation and Processing

Table of contents

Project-Team DAHU

Creation of the Project-Team: 2009 January 01

Keywords:

Computer Science and Digital Science:

3.1.1. - Modeling, representation

3.1.2. - Data management, quering and storage

3.1.3. - Distributed data

3.1.4. - Uncertain data

3.1.5. - Control access, privacy

3.1.9. - Database

4.7. - Access control

7.4. - Logic in Computer Science

Other Research Topics and Application Domains:

9.8. - Privacy

1. Members

Research Scientists

Luc Segoufin [Team leader, Inria, Senior Researcher, HDR]

Serge Abiteboul [Inria, Senior Researcher, HDR]

Faculty Member

Sylvain Schmitz [ENS Cachan, Associate Professor]

PhD Students

Nathan Grosshans [CNRS]

Anthony Lick [CNRS]

David Montoya [Sinovia, granted by CIFRE]

Karima Rafes

Alexandre Vigny [Univ. Paris VII]

Su Yang [ENSM Paris, from Sep 2016]

Post-Doctoral Fellow

Matthias Niewerth [Inria, until Mar 2016]

Administrative Assistant

Thida Iem [Inria]

Others

Thomas Pellissier-Tanon [ENS Lyon, Intern, until Jun 2016]

Pierre Senellart [Telecom Paristech, Professor]

Victor Vianu [UC San Diego, Professor]

2. Overall Objectives

2.1. Overall Objectives

For more information see http://www.lsv.ens-cachan.fr/axes/DAHU/dahu.php.

The need to access and exchange data on the Web has led to database management systems (DBMS) that are increasingly distributed and autonomous. Data extraction and querying on the Web is harder than in classical DBMS, because such data is heterogeneous, redundant, inconsistent and subject to frequent modifications. DBMS thus need to be able to detect errors, to analyze them and to correct them. Moreover, increasingly complex Web applications and services rely on DBMS, and their reliability is crucial. This creates a need for tools for specifying DBMS in a high-level manner that is easier to understand, while also facilitating verification of critical properties.

The study of such specification and verification techniques is the main goal of Dahu.

3. Research Program

3.1. Research Program

Dahu aims at developing mechanisms for high-level specifications of systems built around DBMS, that are easy to understand while also facilitating verification of critical properties. This requires developing tools that are suitable for reasoning about systems that manipulate data. Some tools for specifying and reasoning about data have already been studied independently by the database community and by the verification community, with various motivations. However, this work is still in its infancy and needs to be further developed and unified.

Most current proposals for reasoning about DBMS over XML documents are based on tree automata, taking advantage of the tree structure of XML documents. For this reason, the Dahu team is studying a variety of tree automata. This ranges from restrictions of "classical" tree automata in order to understand their expressive power, to extensions of tree automata in order to understand how to incorporate the manipulation of data.

Moreover, Dahu is also interested in logical frameworks that explicitly refer to data. Such logical frameworks can be used as high level declarative languages for specifying integrity constraints, format change during data exchange, web service functionalities and so on. Moreover, the same logical frameworks can be used to express the critical properties we wish to verify.

In order to achieve its goals, Dahu brings together world-class expertise in both databases and verification.

4. Application Domains

4.1. Application Domains

Databases are pervasive across many application fields. Indeed, most human activities today require some form of data management. In particular, all applications involving the processing of large amounts of data require the use of a database. Increasingly complex Web applications and services also rely on DBMS, and their correctness and robustness is crucial.

We believe that the automated solutions that Dahu aims to develop for verifying such systems will be useful in this context.

5. Highlights of the Year

5.1. Highlights of the Year

Awards

Luc Segoufin together with Mikolaj Bojanczyk, Claire David, Anca Muscholl, and Thomas Schwentick obtained the ACM Alberto O. Mendelzon PODS Test of Time Award in 2016.

6. New Results

6.1. Specification and verification of data-driven systems

Verification of Hierarchical Artifact Systems

Data-driven workflows, of which "business artifacts" are a prime exponent, have been successfully deployed in practice, adopted in industrial standards, and have spawned a rich body of research in academia, focused primarily on static analysis. Over the past few years, we have embarked upon a study of the verification problem for artifact systems. This is a challenging problem because of the presence of unbounded data. In order to deal with the resulting infinite-state system, we developed in earlier work a symbolic approach allowing a reduction to finite-state model checking and yielding a pspace verification algorithm for the simplest variant of the model (no database dependencies and uninterpreted data domain). Subsequently, we extended our approach to allow for database dependencies and numeric data testable by arithmetic constraints. In [19], we make significant progress on several fronts, by considering a much richer and more realistic model than in previous work, incorporating core elements of IBM's successful Guard-Stage-Milestone model. In particular, the model features task hierarchy, concurrency, and richer artifact data. It also allows database key and foreign key dependencies, as well as arithmetic constraints. The results require qualitatively novel techniques, because the reduction to finite-state model checking used in previous work is no longer possible. Instead, the richer model requires the use of a hierarchy of Vector Addition Systems with States. The arithmetic constraints are handled using quantifier elimination techniques, adapted to our setting.

Process-centric views of data-driven workflows.

We also studied the models of *data Petri nets* and *ν-Petri nets*. While these models were introduced in the verification community to analyse protocols and process algebra, they can also be seen as (very limited) data-driven workflows with only unary predicates. Our results this year show that various boundedness problems (e.g. can the database grow unbounded?) are decidable in data Petri nets [22], and pinpoint the exact complexity of safety analysis in *ν*-Petri nets [23].

Complexity in counter systems and in proof systems.

The static analysis of queries on XML trees and data streams relies in a majority of cases on decision procedures expressed in terms of formal systems like counter systems or proof systems. For instance, two-variables first-order data queries on words can be related to reachability in vector addition systems (VAS), and the same queries on trees to reachability in a branching extension of VAS [12]. We are at the forefront on the complexity analysis for such systems [15], [13], [16], [14].

We investigate in the ANR PRODAQ project a different angle on the static analysis of queries, relying on proof systems. Our first results on the subject [18] provide a sequent calculus for a modal data logic with an optimal proof-search algorithm.

6.2. Personal information management.

Thymeflow We developed Thymeflow, a personal knowledge base with spatio-temporal data [24].

The typical Internet user has data spread over several devices and across several online systems. We demonstrate an open-source system for integrating user's data from different sources into a single Knowledge Base. Our system integrates data of different kinds into a coherent whole, starting with email messages, calendar, contacts, and location history. It is able to detect event periods in the user's location data and align them with calendar events. We will demonstrate how to query the system within and across different dimensions, and perform analytics over emails, events, and locations.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

The CIFRE scholarship of David Montoya started in 2014, with Sinovia, Cofely Ineo (group GDF Suez). The topic is on analysis of multimodal itineraries and the integration of itinerary data with other personal data.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

Acronym: PRODAQ

Title: Proof systems for Data Queries

Coordinator: Sylvain Schmitz

Duration: January 2015 – September 2019

Abstract: The project aims at developing proof systems for data logics. It is at the interface between several research communities in database theory, infinite-state system verification and proof theory. The main thrust behind the project is the investigation of proof-theoretic tools for data logic, using in particular insights from substructural logics, and using counter systems as a means to obtain algorithms and complexity results.

8.2. International Research Visitors

8.2.1. Visits of International Scientists

Victor Vianu, June 15 to September 15, UC San diego

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific events organisation

9.1.1.1. General chair, scientific chair

Serge Abiteboul co-organized the workshop "Data, responsibly" at Dagstuhl on ethics in massive data analysis.

9.1.2. Scientific events selection

9.1.2.1. Member of the conference program committees

Serge Abiteboul: Data engineering (ICDE'2016); In 2016, Serge Abiteboul also participated in Awards committees for the French Academy of Sciences and Inria.

Victor Vianu: FOIKS'16.

9.1.2.2. Reviewer

The members of the team reviewed numerous papers for numerous international conferences and journals.

9.1.3. Journal

9.1.3.1. Member of the editorial boards

Victor Vianu: Associate Editor of ACM TOCL, Editor of the Database Theory column in SIGACT News.

9.1.4. Invited talks

Victor Vianu gave invited talks at the Dagstuhl Workshop on Foundations of Data Management, April 2016; POLARIS Colloquium, Univ. of Lille, June 2016; and Ecole Polytechnique, Saclay, July 2016.

Serge Abiteboul gave invited talks in international meetings at Int'l Symp. on DIStributed Computing in Paris, Didapro 6 in Namur, Dagstuhl Workshop on Foundations of Data Management, Dagstuhl workshop on Data, Responsibly, and Paris Open Source Summit.

9.1.5. Leadership within the scientific community

Serge Abiteboul is co-chair of the Parity and Equality committee of Inria.

9.1.6. Scientific expertise

Sylvain Schmitz reviewed grant proposals for the Agence Nationale de la Recherche (ANR).

9.1.7. Research administration

Since 2015 Luc Segoufin is an elected member of the CNHSCT of Inria.

Serge Abiteboul chaired the selection commitee for the call for proposal Dune "Développement d'Universités Numériques Expérimentales"/

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

> Licence : Sylvain Schmitz, Formal Languages, 22.5h, L3, ENS Cachan, France
> Master : Sylvain Schmitz, Tree Automata Techniques and Applications, 22.5h, M1, ENS Cachan, France
> Master : Sylvain Schmitz, Formal Languages, 30h, M2, ENS Cachan, France
> Master : Sylvain Schmitz, Logic, 26.3h, M2, ENS Cachan, France
> Master : Sylvain Schmitz, Logical and Computational Structures for Linguistic Modeling, 18h, M2, ENS Cachan, France
> Master2 : Serge Abiteboul, Web data management, 15h, MPRI
> Master1 : Serge Abiteboul, Initiation to scientific research (lab), 15h
> Licence : Serge Abiteboul, Relational databases, 30h, ENS Cachan
> **E-learning**
>> Mooc Bases de Données Relationnelles - Bador (in French), Inria Fun, Serge Abiteboul, Benjamin Nguyen and Philippe Rigaux, Start January 2016; about 6000 students registered, 6+ weeks. Target audience: students in L3/Master, engineers and scientists using databases.

9.2.2. Supervision

> PhD in progress: Nathan Grosshans, branching program, 01/09/2014, Luc Segoufin and Pierre McKenzie (University of Montreal)
> PhD in progress: David Montoya, Personal information management systems, 01/02/2014, Serge Abiteboul
> PhD in progress: Alexandre Vigny, enumeration, 01/09/2015, Luc Segoufin and Arnaud Durand
> PhD in progress: Karima Rafes, Web semantic and Internet of Thing, 01/02/2015, Serge Abiteboul
> PhD in progress: Simon Halfon, Well quasi orders, 01/09/2015, Sylvain Schmitz and Philippe Schnoebelen
> PhD in progress: Anthony Lick, Proof systems for data queries, 01/09/2016, David Baelde and Sylvain Schmitz
> PhD in progress: Su Yang, Analytics in Personal information management systems, 01/09/2016, Serge Abiteboul et Pierre Senellart

9.2.3. Juries

Luc Segoufin was on the Ph.D committee of Florent Capelli, Université de Paris 7, in June 2016.

Sylvain Schmitz was on the Ph.D. committees of Silva Stella, Università degli studi di Torino, in January 2016, and of Michael Blondin, Université de Montréal, in June 2016.

Serge Abiteboul was on the Ph.D. committee of Damian Bursztyn, University of Paris-Saclay, in December 2016.

9.3. Popularization

For the last couple of years, Serge Abiteboul has initiated and is coordinating an invited blog for Le Monde newspaper, namely Binaire, http://binaire.blog.lemonde.fr. He is the initiator of "The Cabale Informatique de France", a joint action between the Société Informatique de France and Wikipédia France to improve the quality of Wikipédia pages about Computer Science in France. He organized a Workshop Initiation to Wikipedia in Poitiers

Serge Abiteboul is Scientific curator of the exhibit "Terra Data" at Cité des Sciences de La Villette, 2017. He is member of the editorial board of Les big data à découvert, directed Mokrane Bouzeghoub and Rémy Mosseri. He gave invited talks at La Pépinière 4.1 in Nancy, Rencontres du Numérique of ANR, Parole Publique, Sciences à Coeur at Sorbonne Univ., Espace Mendes France at Poitiers,

Serge Abiteboul is Chairman of the Executive Council of the Foundation Blaise Pascal for the promotion of Maths and Computer Science.

10. Bibliography

Major publications by the team in recent years

[1] S. ABITEBOUL, I. MANOLESCU, P. RIGAUX, M.-C. ROUSSET, P. SENELLART.*Web Data Management*, Cambridge University Press, 2012, 456, http://hal.inria.fr/hal-00677720.

[2] S. ABITEBOUL, L. SEGOUFIN, V. VIANU.*Static Analysis of Active XML Systems*, in "ACM Transactions on Database Systems", 2009, vol. 34, n⁰ 4.

[3] V. BARANY, B. T. CATE, L. SEGOUFIN.*Guarded negation*, in "Journal of the ACM", 2015, vol. 62, n⁰ 3, 24, https://hal.inria.fr/hal-01184763.

[4] P. BARCELÓ, L. LIBKIN, A. POGGI, C. SIRANGELO.*XML with incomplete information*, in "J. ACM", 2010, vol. 58, n⁰ 1.

[5] M. BOJAŃCZYK, A. MUSCHOLL, T. SCHWENTICK, L. SEGOUFIN.*Two-variable logic on data trees and applications to XML reasoning*, in "Journal of the ACM", 2009, vol. 56, n⁰ 3.

[6] M. BOJAŃCZYK, L. SEGOUFIN, H. STRAUBING.*Piecewise testable tree languages*, in "Logical Methods in Computer Science (LMCS)", 2012, vol. 8, n⁰ 3.

[7] BALDER TEN. CATE, L. SEGOUFIN.*Transitive Closure Logic, Nested Tree Walking Automata, and XPath*, in "Journal of the ACM", 2010, vol. 57, n⁰ 3.

[8] N. FRANCIS, L. SEGOUFIN, C. SIRANGELO.*Datalog Rewritings of Regular Path Queries using Views*, in "Logical Methods in Computer Science (LMCS)", December 2015, vol. 11, n⁰ 4, https://hal.inria.fr/hal-01248391.

[9] R. LAZIĆ, S. SCHMITZ.*Non-Elementary Complexities for Branching VASS, MELL, and Extensions*, in "ACM Transactions on Computational Logic", May 2015, vol. 16, n⁰ 3:20, p. 1–30 [*DOI : 10.1145/2733375*], https://hal.archives-ouvertes.fr/hal-01168290.

[10] L. LIBKIN, C. SIRANGELO.*Data exchange and schema mappings in open and closed worlds*, in "Journal of Computer System Sciences (JCSS)", 2011.

Publications of the year

Articles in International Peer-Reviewed Journal

[11] P. BEAME, N. GROSSHANS, P. MCKENZIE, L. SEGOUFIN.*Nondeterminism and An Abstract Formulation of Nečiporuk's Lower Bound Method*, in "ACM Transactions on Computation Theory", December 2016, vol. 9, n⁰ 1, p. 1 - 34 [*DOI : 10.1145/3013516*], https://hal.inria.fr/hal-01426213.

[12] F. JACQUEMARD, L. SEGOUFIN, J. DIMINO.*FO2(<,+1,) on data trees, data tree automata and branching vector addition systems*, in "Logical Methods in Computer Science", 2016, vol. 12, n⁰ 2, 32, https://hal.inria.fr/hal-00769249.

[13] S. SCHMITZ.*Complexity Hierarchies Beyond Elementary*, in "ACM Transactions on Computation Theory", February 2016, vol. 8, n⁰ 1 [*DOI : 10.1145/2858784*], https://hal.inria.fr/hal-01267354.

[14] S. SCHMITZ.*Implicational Relevance Logic is 2-ExpTime-Complete*, in "The Journal of Symbolic Logic", June 2016, vol. 81, n⁰ 2, p. 641–661 [*DOI : 10.1017/JSL.2015.7*], https://hal.inria.fr/hal-01340113.

Articles in Non Peer-Reviewed Journal

[15] S. SCHMITZ.*Automata Column: The Complexity of Reachability in Vector Addition Systems*, in "ACM SIGLOG News", January 2016, vol. 3, n⁰ 1, p. 3–21 [*DOI : 10.1145/2893582.2893585*], https://hal.inria.fr/hal-01275972.

Invited Conferences

[16] J. LEROUX, S. SCHMITZ.*Ideal Decompositions for Vector Addition Systems*, in "STACS 2016 - 33rd Symposium on Theoretical Aspects of Computer Science", Orléans, France, N. OLLINGER, H. VOLLMER (editors), Leibniz International Proceedings in Informatics (LIPIcs), Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2016, vol. 47, p. 1–13 [*DOI : 10.4230/LIPIcs.STACS.2016.1*], https://hal.inria.fr/hal-01275991.

International Conferences with Proceedings

[17] S. ABITEBOUL, P. BOURHIS, V. VIANU.*A formal study of collaborative access control in distributed datalog*, in "ICDT 2016 - 19th International Conference on Database Theory", Bordeaux, France, W. MARTENS, T. ZEUME (editors), March 2016, https://hal.inria.fr/hal-01290497.

[18] D. BAELDE, S. LUNEL, S. SCHMITZ.*A Sequent Calculus for a Modal Logic on Finite Data Trees*, in "CSL 2016", Marseille, France, J.-M. TALBOT, L. REGNIER (editors), Leibniz International Proceedings in Informatics, LZI, September 2016, vol. 62, n⁰ 32, p. 1–16 [*DOI : 10.4230/LIPIcs.CSL.2016.32*], https://hal.inria.fr/hal-01191172.

[19] A. DEUTSCH, Y. LI, V. VIANU.*Verification of Hierarchical Artifact Systems*, in "35th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems (PODS 2016)", San Francisco, United States,

ACM (editor), Proceedings of the 35th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems (PODS 2016), June 2016, p. 179 - 194 [*DOI :* 10.1145/2902251.2902275], https://hal.inria.fr/hal-01389845.

[20] J. GOUBAULT-LARRECQ, S. SCHMITZ.*Deciding Piecewise Testable Separability for Regular Tree Languages*, in "ICALP 2016", Rome, Italy, I. CHATZIGIANNAKIS, M. MITZENMACHER, Y. RABANI, D. SANGIORGI (editors), Leibniz International Proceedings in Informatics, July 2016, vol. 55, p. 97:1–97:15 [*DOI :* 10.4230/LIPIcs.ICALP.2016.97], https://hal.inria.fr/hal-01276119.

[21] G. GREFENSTETTE, K. RAFES.*Transforming Wikipedia into an Ontology-based Information Retrieval Search Engine for Local Experts using a Third-Party Taxonomy*, in "Joint Second Workshop on Language and Ontology & Terminology and Knowledge Structures (LangOnto2 + TermiKS) LO2TKS", Portoroz, Slovenia, May 2016, https://hal.inria.fr/hal-01224114.

[22] P. HOFMAN, S. LASOTA, R. LAZIĆ, J. LEROUX, S. SCHMITZ, P. TOTZKE.*Coverability Trees for Petri Nets with Unordered Data*, in "FoSSaCS", Eindhoven, Netherlands, Lecture Notes in Computer Science, Springer, 2016, vol. 9634, p. 445–461 [*DOI :* 10.1007/978-3-662-49630-5_26], https://hal.inria.fr/hal-01252674.

[23] R. LAZIĆ, S. SCHMITZ.*The Complexity of Coverability in ν-Petri Nets*, in "LICS 2016", New York, United States, ACM Press, 2016, p. 467–476 [*DOI :* 10.1145/2933575.2933593], https://hal.inria.fr/hal-01265302.

[24] D. MONTOYA, T. PELLISSIER TANON, S. ABITEBOUL, F. M. SUCHANEK.*Thymeflow, A Personal Knowledge Base with Spatio-Temporal Data*, in "25th ACM International Conference on Information and Knowledge Management", Indianapolis, IN, United States, October 2016 [*DOI :* 10.1145/2983323.2983337], https://hal.inria.fr/hal-01355150.

Conferences without Proceedings

[25] S. ABITEBOUL, G. MIKLAU, J. STOYANOVICH, G. WEIKUM.*Data, Responsibly (Dagstuhl Seminar 16291)*, in "Dagstuhl seminar", Dagstuhl, Germany, 2016, https://hal.inria.fr/hal-01405693.

Scientific Popularization

[26] S. ABITEBOUL.*Analyse des données et choix de société* , in "parole publique", March 2016, https://hal.inria.fr/hal-01273439.

Other Publications

[27] J. STOYANOVICH, S. ABITEBOUL, G. MIKLAU.*Data, Responsibly: Fairness, Neutrality and Transparency in Data Analysis*, March 2016, International Conference on Extending Database Technology, https://hal.inria.fr/hal-01290695.

Team DATASHAPE

Understanding the Shape of Data

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

Table of contents

Team DATASHAPE

Creation of the Team: 2016 January 01

Keywords:

Computer Science and Digital Science:
3.1.1. - Modeling, representation
3.1.4. - Uncertain data
3.3.3. - Big data analysis
3.4. - Machine learning and statistics
5.5.1. - Geometrical modeling
6.2.4. - Statistical methods
6.2.8. - Computational geometry and meshes
6.3.3. - Data processing
7.2. - Discrete mathematics, combinatorics
7.5. - Geometry, Topology
8.2. - Machine learning

Other Research Topics and Application Domains:
5. - Industry of the future
9.4.1. - Computer science
9.4.2. - Mathematics
9.4.5. - Data science

1. Members

Research Scientists
Jean-Daniel Boissonnat [Inria, Senior Researcher, HDR]
Frédéric Chazal [Team leader, Inria, Senior Researcher, HDR]
David Cohen-Steiner [Inria, Researcher]
Marc Glisse [Inria, Researcher]
Alfredo Hubard [Inria, Starting Research position, until Aug 2016, granted by FP7 ERC GUDHI project]
Steve Oudot [Inria, Researcher, HDR]

Faculty Members
Bertrand Michel [Univ. Paris VI, Associate Professor]
Maksim Ovsjanikov [Ecole Polytechnique, Professor]

Technical Staff
Clément Jamin [Inria, granted by FP7 ERC GUDHI project]
Vincent Rouvreau [Inria, SED 40%]

PhD Students
Eddie Aamari [Région]
Bertrand Beaufils [Sysnav, from Apr 2016, industrial grant]
Nicolas Berkouk [Inria, from Nov 2016]
Thomas Bonis [Inria]
Claire Brécheteau [Univ. Paris XI]
Mathieu Carrière [Inria, granted by FP7 ERC GUDHI project]

Jérémy Sylvain Cochoy [Univ. Paris XI]
Alba Chiara de Vitis [Inria, granted by FP7 ERC GUDHI project]
Francois Godi [Univ. Nice]
Ruqi Huang [Inria]
Siargey Kachanovich [Inria, granted by FP7 ERC GUDHI project]
Siddharth Pritam [Inria, from Apr 2016, granted by FP7 ERC GUDHI project]
Mael Rouxel-Labbé [Geometry Factory, granted by CIFRE]

Post-Doctoral Fellows
Paweł Dłotko [Inria, granted by FP7 ERC GUDHI project]
Kunal Dutta [Inria, granted by FP7 ERC GUDHI project]
Mathijs Wintraecken [Inria, granted by FP7 ERC GUDHI project]

Visiting Scientists
Sandip Banerjee [Indian Statistical Institute, Visiting PhD Student, from Mar 2016 until Aug 2016]
Sameer Desai [Indian Statistical Institute, Visiting PhD Student, from Oct 2016 until Dec 2016]
Arijit Ghosh [Indian Statistical Institute, April 2016 and Nov 2016]
Ilaria Giulini [Univ. Paris VI]

Administrative Assistants
Christine Aklouche [Inria]
Florence Barbara [Inria]

Others
Emmanuel Caron-Parte [Inria, intern from Université Pierre et Marie Curie, from Apr 2016 until Sep 2016]
Ramsay Dyer [Freelance Editor, External Collaborator]
Uday Kusupati [Inria, intern from Indian Institute of Technology, from May 2016 until Jul 2016]
Harry Lang [Fullbright fellow, Visiting PhD Student]
Clément Maria [Queensland University, External Collaborator]
Laurent Rineau [Geometry Factory, External Collaborator]
Zongjun Tan [Inria, intern from École Centrale Paris, from Feb 2016 until Aug 2016]
Jane Tournois [Geometry Factory, External Collaborator]
Mariette Yvinec [Retired, Researcher, HDR]

2. Overall Objectives

2.1. Overall Objectives

DataShape is a research project in Topological Data Analysis (TDA), a recent field whose aim is to uncover, understand and exploit the topological and geometric structure underlying complex and possibly high dimensional data. The DATASHAPE project gathers a unique variety of expertise that allows it to embrace the mathematical, statistical, algorithmic and applied aspects of the field in a common framework ranging from fundamental theoretical studies to experimental research and software development.

The expected output of DATASHAPE is two-fold. First, we intend to set-up and develop the mathematical, statistical and algorithmic foundations of Topological and Geometric Data Analysis. Second, we intend to develop the Gudhi platform in order to provide an efficient state-of-the-art toolbox for the understanding of the topology and geometry of data.

3. Research Program

3.1. Algorithmic aspects of topological and geometric data analysis

TDA requires to construct and manipulate appropriate representations of complex and high dimensional shapes. A major difficulty comes from the fact that the complexity of data structures and algorithms used to approximate shapes rapidly grows as the dimensionality increases, which makes them intractable in high dimensions. We focus our research on simplicial complexes which offer a convenient representation of general shapes and generalize graphs and triangulations. Our work includes the study of simplicial complexes with good approximation properties and the design of compact data structures to represent them.

In low dimensions, effective shape reconstruction techniques exist that can provide precise geometric approximations very efficiently and under reasonable sampling conditions. Extending those techniques to higher dimensions as is required in the context of TDA is problematic since almost all methods in low dimensions rely on the computation of a subdivision of the ambient space. A direct extension of those methods would immediately lead to algorithms whose complexities depend exponentially on the ambient dimension, which is prohibitive in most applications. A first direction to by-pass the curse of dimensionality is to develop algorithms whose complexities depend on the intrinsic dimension of the data (which most of the time is small although unknown) rather than on the dimension of the ambient space. Another direction is to resort to cruder approximations that only captures the homotopy type or the homology of the sampled shape. The recent theory of persistent homology provides a powerful and robust tool to study the homology of sampled spaces in a stable way.

3.2. Statistical aspects of topological and geometric data analysis

The wide variety of larger and larger available data - often corrupted by noise and outliers - requires to consider the statistical properties of their topological and geometric features and to propose new relevant statistical models for their study.

There exist various statistical and machine learning methods intending to uncover the geometric structure of data. Beyond manifold learning and dimensionality reduction approaches that generally do not allow to assert the relevance of the inferred topological and geometric features and are not well-suited for the analysis of complex topological structures, set estimation methods intend to estimate, from random samples, a set around which the data is concentrated. In these methods, that include support and manifold estimation, principal curves/manifolds and their various generalizations to name a few, the estimation problems are usually considered under losses, such as Hausdorff distance or symmetric difference, that are not sensitive to the topology of the estimated sets, preventing these tools to directly infer topological or geometric information.

Regarding purely topological features, the statistical estimation of homology or homotopy type of compact subsets of Euclidean spaces, has only been considered recently, most of the time under the quite restrictive assumption that the data are randomly sampled from smooth manifolds.

In a more general setting, with the emergence of new geometric inference tools based on the study of distance functions and algebraic topology tools such as persistent homology, computational topology has recently seen an important development offering a new set of methods to infer relevant topological and geometric features of data sampled in general metric spaces. The use of these tools remains widely heuristic and until recently there were only a few preliminary results establishing connections between geometric inference, persistent homology and statistics. However, this direction has attracted a lot of attention over the last three years. In particular, stability properties and new representations of persistent homology information have led to very promising results to which the DATASHAPE members have significantly contributed. These preliminary results open many perspectives and research directions that need to be explored.

Our goal is to build on our first statistical results in TDA to develop the mathematical foundations of Statistical Topological and Geometric Data Analysis. Combined with the other objectives, our ultimate goal is to provide a well-founded and effective statistical toolbox for the understanding of topology and geometry of data.

3.3. Topological approach for multimodal data processing

Due to their geometric nature, multimodal data (images, video, 3D shapes, etc.) are of particular interest for the techniques we develop. Our goal is to establish a rigorous framework in which data having different representations can all be processed, mapped and exploited jointly. This requires adapting our tools and sometimes developing entirely new or specialized approaches.

The choice of multimedia data is motivated primarily by the fact that the amount of such data is steadily growing (with e.g. video streaming accounting for nearly two thirds of peak North-American Internet traffic, and almost half a billion images being posted on social networks each day), while at the same time it poses significant challenges in designing informative notions of (dis)-similarity as standard metrics (e.g. Euclidean distances between points) are not relevant.

3.4. Experimental research and software development

We develop a high quality open source software platform called GUDHI which is becoming a reference in geometric and topological data analysis in high dimensions. The goal is not to provide code tailored to the numerous potential applications but rather to provide the central data structures and algorithms that underly applications in geometric and topological data analysis.

The development of the GUDHI platform also serves to benchmark and optimize new algorithmic solutions resulting from our theoretical work. Such development necessitates a whole line of research on software architecture and interface design, heuristics and fine-tuning optimization, robustness and arithmetic issues, and visualization. We aim at providing a full programming environment following the same recipes that made up the success story of the CGAL library, the reference library in computational geometry.

Some of the algorithms implemented on the platform will also be interfaced to other software platform, such as the R software [0] for statistical computing, and languages such as Python in order to make them usable in combination with other data analysis and machine learning tools. A first attempt in this direction has been done with the creation of an R package called TDA in collaboration with the group of Larry Wasserman at Carnegie Mellon University (Inria Associated team CATS) that already includes some functionalities of the GUDHI library and implements some joint results between our team and the CMU team. A similar interface with the Python language is also considered a priority. To go even further towards helping users, we will provide utilities that perform the most common tasks without requiring any programming at all.

4. Application Domains

4.1. Main application domains

Our work is mostly of a fundamental mathematical and algorithmic nature but finds applications in a variety of application in data analysis, more precisely in Topological Data Analysis (TDA). Although TDA is a quite recent field, it already founds applications in material science, biology, sensor networks, 3D shapes analysis and processing, to name a few.

More specifically, DATASHAPEhas recently started to work on the analysis of trajectories obtained from inertial sensors (starting PhD thesis of Bertrand Beaufils) and is exploring some possible new applications in material science.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

Jean-Daniel Boissonnat has been elected a professor at the Collège de France, on the Chair Informatics and Computational Sciences for the academic year 2016-2017.

---

[0]https://www.r-project.org/

5.1.2. Books

Publication of a book [29], providing a self-contained presentation of the theory of persistence modules over the real line, the objects that are at the heart of the field of TDA.

6. New Software and Platforms

6.1. GUDHI

Geometric Understanding in Higher Dimensions
SCIENTIFIC DESCRIPTION

The current release of the GUDHI library includes:

- Data structures to represent, construct and manipulate simplicial and cubical complexes.
- Algorithms to compute simplicial complexes from point cloud data.
- Algorithms to compute persistent homology and multi-field persistent homology.
- Simplification methods via implicit representations.

FUNCTIONAL DESCRIPTION

The GUDHI open source library will provide the central data structures and algorithms that underlie applications in geometry understanding in higher dimensions. It is intended to both help the development of new algorithmic solutions inside and outside the project, and to facilitate the transfer of results in applied fields.

- Participants: Jean-Daniel Boissonnat, Marc Glisse, Mariette Yvinec, Clément Maria, David Salinas, Paweł Dłotko, Siargey Kachanovich and Vincent Rouvreau
- Contact: Jean-Daniel Boissonnat
- URL: http://gudhi.gforge.inria.fr/

7. New Results

7.1. Algorithmic aspects of topological and geometric data analysis

7.1.1. An Efficient Representation for Filtrations of Simplicial Complexes
Participant: Jean-Daniel Boissonnat.

In collaboration with Karthik C.S. (Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Israel)

A filtration over a simplicial complex K is an ordering of the simplices of K such that all prefixes in the ordering are subcomplexes of K. Filtrations are at the core of Persistent Homology, a major tool in Topological Data Analysis. In order to represent the filtration of a simplicial complex, the entire filtration can be appended to any data structure that explicitly stores all the simplices of the complex such as the Hasse diagram or the recently introduced Simplex Tree by Boissonnat and Maria [Algorithmica '14]. However, with the popularity of various computational methods that need to handle simplicial complexes, and with the rapidly increasing size of the complexes, the task of finding a compact data structure that can still support efficient queries is of great interest.

This direction has been recently pursued for the case of maintaining simplicial complexes. For instance, Boissonnat et al. [SoCG '15] considered storing the simplices that are maximal for the inclusion and Attali et al. [IJCGA '12] considered storing the simplices that block the expansion of the complex. Nevertheless, so far there has been no data structure that compactly stores the *filtration* of a simplicial complex, while also allowing the efficient implementation of basic operations on the complex.

In this work [22], we propose a new data structure called the Critical Simplex Diagram (CSD) which is a variant of our work on the Simplex Array List (SAL) introduced in [SoCG '15]. Our data structure allows to store in a compact way the filtration of a simplicial complex, and allows for the efficient implementation of a large range of basic operations. Moreover, we prove that our data structure is essentially optimal with respect to the requisite storage space. Next, we show that the CSD representation admits the following construction algorithms.

- A new *edge-deletion* algorithm for the fast construction of Flag complexes, which only depends on the number of critical simplices and the number of vertices.

- A new *matrix-parsing* algorithm to quickly construct the relaxed strong Delaunay complexes, depending only on the number of witnesses and the dimension of the complex.

7.1.2. *Discretized Riemannian Delaunay triangulations*
Participants: Mael Rouxel-Labbé, Mathijs Wintraecken, Jean-Daniel Boissonnat.

Anisotropic meshes are desirable for various applications, such as the numerical solving of partial differential equations and graphics. In [27], we introduce an algorithm to compute discrete approximations of Riemannian Voronoi diagrams on 2-manifolds. This is not straightforward because geodesics, shortest paths between points, and therefore distances cannot in general be computed exactly. Our implementation employs recent developments in the numerical computation of geodesic distances and is accelerated through the use of an underlying anisotropic graph structure. We give conditions that guarantee that our discrete Riemannian Voronoi diagram is combinatorially equivalent to the Riemannian Voronoi diagram and that its dual is an embedded triangulation, using both approximate geodesics and straight edges. Both the theoretical guarantees on the approximation of the Voronoi diagram and the implementation are new and provide a step towards the practical application of Riemannian Delaunay triangulations.

7.1.3. *Efficient and Robust Persistent Homology for Measures*
Participants: Frédéric Chazal, Steve Oudot.

In collaboration with M. Buchet (Tohoku University), D. Sheehy (Univ. Connecticut).

A new paradigm for point cloud data analysis has emerged recently, where point clouds are no longer treated as mere compact sets but rather as empirical measures. A notion of distance to such measures has been defined and shown to be stable with respect to perturbations of the measure. This distance can easily be computed pointwise in the case of a point cloud, but its sublevel-sets, which carry the geometric information about the measure, remain hard to compute or approximate. This makes it challenging to adapt many powerful techniques based on the Euclidean distance to a point cloud to the more general setting of the distance to a measure on a metric space. We propose an efficient and reliable scheme to approximate the topological structure of the family of sublevel-sets of the distance to a measure. We obtain an algorithm for approximating the persistent homology of the distance to an empirical measure that works in arbitrary metric spaces. Precise quality and complexity guarantees are given with a discussion on the behavior of our approach in practice [17].

7.1.4. *Shallow Packings in Geometry*
Participants: Kunal Dutta, Arijit Ghosh.

A merged paper with Ezra, Esther (School of Mathematics, Georgia Institute of Technology, Atlanta, U.S.A.)

We refine the bound on the packing number, originally shown by Haussler, for shallow geometric set systems. Specifically, let V be a finite set system defined over an n-point set X; we view V as a set of indicator vectors over the n-dimensional unit cube. A $delta$-separated set of V is a subcollection W, such that the Hamming distance between each pair $u, v \in W$ is greater than δ, where $\delta > 0$ is an integer parameter. The δ-packing number is then defined as the cardinality of the largest δ-separated subcollection of V. Haussler showed an asymptotically tight bound of $\Theta((n/delta)^d)$ on the δ-packing number if V has VC-dimension (or primal shatter dimension) d. We refine this bound for the scenario where, for any subset, $X' \subset X$ of size $m \leq n$ and for any parameter $1 \leq k \leq m$, the number of vectors of length at most k in the restriction of V to X' is only

$O(m^{d_1} k^{d-d_1})$, for a fixed integer $d > 0$ and a real parameter $1 \leq d_1 \leq d$ (this generalizes the standard notion of bounded primal shatter dimension when $d_1 = d$). In this case when V is "k-shallow" (all vector lengths are at most k), we show that its δ-packing number is $O(n^{d_1} k^{d-d_1}/\delta^d)$, matching Haussler's bound for the special cases where $d_1 = d$ or $k = n$. We present two proofs, the first is an extension of Haussler's approach, and the second extends the proof of Chazelle, originally presented as a simplification for Haussler's proof. [21]

- A new *tight upper bound* for shallow-packings in δ-separated set systems of bounded primal shatter dimension.

7.1.5. *On Subgraphs of Bounded Degeneracy in Hypergraphs*
Participants: Kunal Dutta, Arijit Ghosh.

A k-uniform hypergraph has degeneracy bounded by d if every induced subgraph has a vertex of degree at most d. Given a k-uniform hypergraph $H = (V(H), E(H))$, we show there exists an induced subgraph of size at least

$$\sum_{v \in V(H)} \min 1, ck \left(\frac{d+1}{d_H(v)+1} \right)^{1/(k-1)},$$

where $c_k = 2^{-(1+\frac{1}{k-1})} \left(1 - \frac{1}{k}\right)$ and $d_H(v)$ denotes the degree of vertex v in the hypergraph H. This extends and generalizes a result of Alon-Kahn-Seymour (Graphs and Combinatorics, 1987) for graphs, as well as a result of Dutta-Mubayi-Subramanian (SIAM Journal on Discrete Mathematics, 2012) for linear hypergraphs, to general k-uniform hypergraphs. We also generalize the results of Srinivasan and Shachnai (SIAM Journal on Discrete Mathematics, 2004) from independent sets (0-degenerate subgraphs) to d-degenerate subgraphs. We further give a simple non-probabilistic proof of the Dutta-Mubayi-Subramanian bound for linear k-uniform hypergraphs, which extends the Alon-Kahn-Seymour proof technique to hypergraphs. Our proof combines the random permutation technique of Bopanna-Caro-Wei (see e.g. The Probabilistic Method, N. Alon and J. H. Spencer; Dutta-Mubayi-Subramanian) and also Beame-Luby (SODA, 1990) together with a new local density argument which may be of independent interest. We also provide some applications in discrete geometry, and address some natural algorithmic questions. [28]

- A new algorithmic *lower bound* for largest d-degenerate subgraphs in k-uniform hypergraphs.

7.1.6. *A Simple Proof of Optimal Epsilon Nets*
Participants: Kunal Dutta, Arijit Ghosh.

In collaboration with Nabil Mustafa (Université Paris-Est, Laboratoire d'Informatique Gaspard-Monge, ESIEE Paris, France.)

Showing the existence of ε-nets of small size has been the subject of investigation for almost 30 years, starting from the initial breakthrough of Haussler and Welzl (1987). Following a long line of successive improvements, recent results have settled the question of the size of the smallest ε-nets for set systems as a function of their so-called shallow-cell complexity.

In this paper we give a short proof of this theorem in the space of a few elementary paragraphs, showing that it follows by combining the ε-net bound of Haussler and Welzl (1987) with a variant of Haussler's packing lemma (1991).

This implies all known cases of results on unweighted ε-nets studied for the past 30 years, starting from the result of Matoušek, Seidel and Welzl (1990) to that of Clarkson and Varadajan (2007) to that of Varadarajan (2010) and Chan, Grant, Könemann and Sharpe (2012) for the unweighted case, as well as the technical and intricate paper of Aronov, Ezra and Sharir (2010). [40]

- A new *unified proof* for all known bounds on unweighted ε-nets studied in the last 30 years.

7.1.7. *Combinatorics of Set Systems with Small Shallow Cell Complexity: Optimal Bounds via Packings*

Participants: Kunal Dutta, Arijit Ghosh.

In collaboration with Bruno Jartoux and Nabil Mustafa (Université Paris-Est Marne-la-Vallée, Laboratoire d'Informatique Gaspard-Monge, ESIEE Paris, France.)

The packing lemma of Haussler states that given a set system (X, R) with bounded VC dimension, if every pair of sets in R are 'far apart' (i.e., have large symmetric difference), then R cannot contain too many sets. This has turned out to be the technical foundation for many results in geometric discrepancy using the entropy method as well as recent work on set systems with bounded VC dimension. Recently it was generalized to the shallow packing lemma [Dutta-Ezra-Ghosh SoCG 2015, Mustafa DCG 2016], applying to set systems as a function of their shallow cell complexity. In this paper we present several new results and applications related to packings:

1. an optimal lower bound for shallow packings, thus settling the open question in Ezra (SODA 2014) and Dutta et al. (SoCG 2015),
2. improved bounds on Mnets, providing a combinatorial analogue to Macbeath regions in convex geometry (Annals of Mathematics, 1952),
3. simplifying and generalizing the main technical tool in Fox et al. (J. of the EMS, 2016).

Besides using the packing lemma and a combinatorial construction, our proofs combine tools from polynomial partitioning and the probabilistic method. [37]

- A new *optimal lower bound* for shallow packings.
- New *improved bounds* for M-nets - combinatorial analogs of Macbeath regions in convex geometry.

7.1.8. *A new asymmetric correlation inequality for Gaussian measure*

Participants: Kunal Dutta, Arijit Ghosh.

In collaboration with Nabil Mustafa (Université Paris-Est Marne-la-Vallée, Laboratoire d'Informatique Gaspard-Monge, ESIEE Paris, France.)

The Khatri-Šidák lemma says that for any Gaussian measure μ over $\mathbb{R}^n$, given a convex set K and a slab L, both symmetric about the origin, one has $\mu(K \cap L) \geq \mu(K)\mu(L)$. We state and prove a new asymmetric version of the Khatri-Šidák lemma when K is a symmetric convex body and L is a slab (not necessarily symmetric about the barycenter of K). Our result also extends that of Szarek and Werner (1999), in a special case.

- A new *asymmetric* inequality for gaussian measure. [38].

7.2. Statistical aspects of topological and geometric data analysis

7.2.1. *Stability and Minimax Optimality of Tangential Delaunay Complexes for Manifold Reconstruction*

Participant: Eddie Aamari.

In collaboration with C. Levrard (Univ. Paris Diderot).

we consider the problem of optimality in manifold reconstruction. A random sample $\mathbb{X}_n = \{X_1, \ldots, X_n\} \subset \mathbb{R}^D$ composed of points lying on a d-dimensional submanifold M, with or without outliers drawn in the ambient space, is observed. Based on the tangential Delaunay complex, we construct an estimator $\widehat{M}$ that is ambient isotopic and Hausdorff-close to M with high probability. $\widehat{M}$ is built from existing algorithms. In a model without outliers, we show that this estimator is asymptotically minimax optimal for the Hausdorff distance over a class of submanifolds with reach condition. Therefore, even with no a priori information on the tangent spaces of M, our estimator based on tangential Delaunay complexes is optimal. This shows that the optimal rate of convergence can be achieved through existing algorithms. A similar result is also derived in a model with outliers. A geometric interpolation result is derived, showing that the tangential Delaunay complex is stable with respect to noise and perturbations of the tangent spaces. In the process, a denoising procedure and a tangent space estimator both based on local principal component analysis (PCA) are studied [32].

7.2.2. *Rates in the Central Limit Theorem and diffusion approximation via Stein's Method*
Participant: Thomas Bonis.

We present a way to apply Stein's method in order to bound the Wasserstein distance between a, possibly discrete, measure and another measure assumed to be the invariant measure of a diffusion operator. We apply this construction to obtain convergence rates, in terms of p-Wasserstein distance for $p \geq 2$, in the Central Limit Theorem in dimension 1 under precise moment conditions. We also establish a similar result for the Wasserstein distance of order 2 in the multidimensional setting. In a second time, we study the convergence of stationary distributions of Markov chains in the context of diffusion approximation, with applications to density estimation from geometric random graphs and to sampling using the Langevin Monte Carlo algorithm [33].

7.2.3. *Rates of Convergence for Robust Geometric Inference*
Participants: Frédéric Chazal, Bertrand Michel.

In collaboration with P. Massart (Univ. Paris Sud et Inria Select team).

Distances to compact sets are widely used in the field of Topological Data Analysis for inferring geometric and topological features from point clouds. In this context, the distance to a probability measure (DTM) has been introduced by Chazal et al. as a robust alternative to the distance a compact set. In practice, the DTM can be estimated by its empirical counterpart, that is the distance to the empirical measure (DTEM). In this paper we give a tight control of the deviation of the DTEM. Our analysis relies on a local analysis of empirical processes. In particular, we show that the rate of convergence of the DTEM directly depends on the regularity at zero of a particular quantile function which contains some local information about the geometry of the support. This quantile function is the relevant quantity to describe precisely how difficult is a geometric inference problem. Several numerical experiments illustrate the convergence of the DTEM and also confirm that our bounds are tight [19].

7.2.4. *Data driven estimation of Laplace-Beltrami operator*
Participants: Frédéric Chazal, Bertrand Michel, Ilaria Giulini.

Approximations of Laplace-Beltrami operators on manifolds through graph Laplacians have become popular tools in data analysis and machine learning. These discretized operators usually depend on bandwidth parameters whose tuning remains a theoretical and practical problem. In this paper, we address this problem for the unnormalized graph Laplacian by establishing an oracle inequality that opens the door to a well-founded data-driven procedure for the bandwidth selection. Our approach relies on recent results by Lacour and Massart on the so-called Lepski's method [26].

7.3. Topological approach for multimodal data processing

7.3.1. *Persistence-based Pooling for Shape Pose Recognition*
Participants: Thomas Bonis, Frédéric Chazal, Steve Oudot, Maksim Ovsjanikov.

We propose a novel pooling approach for shape classification and recognition using the bag-of-words pipeline, based on topological persistence, a recent tool from Topological Data Analysis. Our technique extends the standard max-pooling, which summarizes the distribution of a visual feature with a single number, thereby losing any notion of spatiality. Instead, we propose to use topological persistence, and the derived persistence diagrams, to provide significantly more informative and spatially sensitive characterizations of the feature functions, which can lead to better recognition performance. Unfortunately, despite their conceptual appeal, persistence diagrams are difficult to handle, since they are not naturally represented as vectors in Euclidean space and even the standard metric, the bottleneck distance is not easy to compute. Furthermore, classical distances between diagrams, such as the bottleneck and Wasserstein distances, do not allow to build positive definite kernels that can be used for learning. To handle this issue, we provide a novel way to transform persistence diagrams into vectors, in which comparisons are trivial. Finally, we demonstrate the performance of our construction on the Non-Rigid 3D Human Models SHREC 2014 dataset, where we show that topological pooling can provide significant improvements over the standard pooling methods for the shape pose recognition within the bag-of-words pipeline [23].

7.3.2. *Structure and Stability of the 1-Dimensional Mapper*
Participants: Steve Oudot, Mathieu Carrière.

Given a continuous function $f : X \to \mathbb{R}$ and a cover $\mathcal{I}$ of its image by intervals, the Mapper is the nerve of a refinement of the pullback cover $f^{-1}(\mathcal{I})$. Despite its success in applications, little is known about the structure and stability of this construction from a theoretical point of view. As a pixelized version of the Reeb graph of f, it is expected to capture a subset of its features (branches, holes), depending on how the interval cover is positioned with respect to the critical values of the function. Its stability should also depend on this positioning. We propose a theoretical framework that relates the structure of the Mapper to the one of the Reeb graph, making it possible to predict which features will be present and which will be absent in the Mapper given the function and the cover, and for each feature, to quantify its degree of (in-)stability. Using this framework, we can derive guarantees on the structure of the Mapper, on its stability, and on its convergence to the Reeb graph as the granularity of the cover $\mathcal{I}$ goes to zero [25].

7.3.3. *Decomposition of exact pfd persistence bimodules*
Participants: Steve Oudot, Jérémy Cochoy.

We characterize the class of persistence modules indexed over $\mathbb{R}^2$ that are decomposable into summands whose support have the shape of a *block*—i.e. a horizontal band, a vertical band, an upper-right quadrant, or a lower-left quadrant. Assuming the modules are *pointwise finite-dimensional* (pfd), we show that they are decomposable into block summands if and only if they satisfy a certain local property called *exactness*. Our proof follows the same scheme as the proof of decomposition for pfd persistence modules indexed over $\mathbb{R}$, yet it departs from it at key stages due to the product order not being a total order on $\mathbb{R}^2$, which leaves some important gaps open. These gaps are filled in using more direct arguments. Our work is motivated primarily by the stability theory for zigzags and interlevel-sets persistence modules, in which block-decomposable bimodules play a key part. Our results allow us to drop some of the conditions under which that theory holds, in particular the Morse-type conditions [39].

7.4. Experimental research and software development

7.4.1. *Topological Microstructure Analysis Using Persistence Landscapes*
Participant: Paweł Dłotko.

In collaboration with T. Wanner (George Mason University).

Phase separation mechanisms can produce a variety of complicated and intricate microstructures, which often can be difficult to characterize in a quantitative way. In recent years, a number of novel topological metrics for microstructures have been proposed, which measure essential connectivity information and are based on techniques from algebraic topology. Such metrics are inherently computable using computational homology, provided the microstructures are discretized using a thresholding process. However, while in many cases the thresholding is straightforward, noise and measurement errors can lead to misleading metric values. In such situations, persistence landscapes have been proposed as a natural topology metric. Common to all of these approaches is the enormous data reduction, which passes from complicated patterns to discrete information. It is therefore natural to wonder what type of information is actually retained by the topology. In the present paper, we demonstrate that averaged persistence landscapes can be used to recover central system information in the Cahn-Hilliard theory of phase separation. More precisely, we show that topological information of evolving microstructures alone suffices to accurately detect both concentration information and the actual decomposition stage of a data snapshot. Considering that persistent homology only measures discrete connectivity information, regardless of the size of the topological features, these results indicate that the system parameters in a phase separation process affect the topology considerably more than anticipated. We believe that the methods discussed in this paper could provide a valuable tool for relating experimental data to model simulations [36].

7.4.2. *Topological analysis of the connectome of digital reconstructions of neural microcircuits*
Participant: Paweł Dłotko.

In collaboration with K. Hess, L. Ran, H. Markram, E. Muller, M. Nolte, M. Reimann, M. Scolamiero, K. Turner (Univ. of Aberden, EPFL, Brain and Mind Institute).

A first draft digital reconstruction and simulation of a microcircuit of neurons in the neocortex of a two-week-old rat was recently published. Since graph-theoretical methods may not be sufficient to understand the immense complexity of the network formed by the neurons and their connections, we explored whether application of methods from algebraic topology can provide a novel and useful perspective on the structural and functional organization of the microcircuit. Structural topological analysis revealed that directed graphs representing the connectivity between neurons are significantly different from random graphs and that there exist an enormous number of simplicial complexes of different dimensions representing all-to-all connections within different sets of neurons, the most extreme motif of neuronal clustering reported so far in the brain. Functional topological analysis based on data from simulations confirmed the interest of a new approach to studying the relationship between the structure of the connectome and its emergent functions. In particular, functional responses to different stimuli can readily be distinguished by topological methods. This study represents the first algebraic topological analysis of connectomics data from neural microcircuits and shows promise for general applications in network science.

7.4.3. *A persistence landscapes toolbox for topological statistics*
Participant: Paweł Dłotko.

In collaboration with P. Bubenik (University of Florida).

Topological data analysis provides a multiscale description of the geometry and topology of quantitative data. The persistence landscape is a topological summary that can be easily combined with tools from statistics and machine learning. We give efficient algorithms for calculating persistence landscapes, their averages, and distances between such averages. We discuss an implementation of these algorithms and some related procedures. These are intended to facilitate the combination of statistics and machine learning with topological data analysis. We present an experiment showing that the low-dimensional persistence landscapes of points sampled from spheres (and boxes) of varying dimensions differ.

7.5. Miscellaneous

7.5.1. *Monotone Simultaneous Paths Embeddings in $\mathbb{R}^d$*

Participant: Marc Glisse.

In collaboration with O. Devillers and S. Lazard (Inria Nancy), David Bremner (University of New Brunswick, Canada), Giuseppe Liotta (University of Perugia, Italy), Tamara Mchedlidze (KIT, Germany), Sue Whitesides (University of Victoria, Canada), Stephen Wismath (University of Lethbridge, Canada).

We study[24] the following problem: Given k paths that share the same vertex set, is there a simultaneous geometric embedding of these paths such that each individual drawing is monotone in some direction? We prove that for any dimension $d \geq 2$, there is a set of $d + 1$ paths that does not admit a monotone simultaneous geometric embedding.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Research contract with GeometryFactory in the context of Mael Rouxel-Labbé's Ph.D. thesis on anisotropic mesh generation [12].

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

9.1.1.1. ANR TOPDATA

Participants: Jean-Daniel Boissonnat, Frédéric Chazal, David Cohen-Steiner, Mariette Yvinec, Steve Oudot, Marc Glisse.

- Acronym : TopData.

- Type : ANR blanc.

- Title : Topological Data Analysis: Statistical Methods and Inference.

- Coordinator : Frédéric Chazal (DATASHAPE).

- Duration : 4 years starting October 2013.

- Others Partners: Département de Mathématiques (Université Paris Sud), Institut de Mathématiques (Université de Bourgogne), LPMA (Université Paris Diderot), LSTA (Université Pierre et Marie Curie).

- Abstract: TopData aims at designing new mathematical frameworks, models and algorithmic tools to infer and analyze the topological and geometric structure of data in different statistical settings. Its goal is to set up the mathematical and algorithmic foundations of Statistical Topological and Geometric Data Analysis and to provide robust and efficient tools to explore, infer and exploit the underlying geometric structure of various data.

Our conviction, at the root of this project, is that there is a real need to combine statistical and topological/geometric approaches in a common framework, in order to face the challenges raised by the inference and the study of topological and geometric properties of the wide variety of larger and larger available data. We are also convinced that these challenges need to be addressed both from the mathematical side and the algorithmic and application sides. Our project brings together in a unique way experts in Statistics, Geometric Inference and Computational Topology and Geometry. Our common objective is to design new theoretical frameworks and algorithmic tools and thus to contribute to the emergence of a new field at the crossroads of these domains. Beyond the purely scientific aspects we hope this project will help to give birth to an active interdisciplinary community. With these goals in mind we intend to promote, disseminate and make our tools available and useful for a broad audience, including people from other fields.

- See also: http://geometrica.saclay.inria.fr/collaborations/TopData/Home.html

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

9.2.1.1. ERC GUDHI

Title: Algorithmic Foundations of Geometry Understanding in Higher Dimensions.

Program: FP7.

Type: ERC.

Duration: February 2014 - January 2019.

Coordinator: Inria.

PI: Jean-Daniel Boissonnat.

The central goal of this proposal is to settle the algorithmic foundations of geometry understanding in dimensions higher than 3. We coin the term geometry understanding to encompass a collection of tasks including the computer representation and the approximation of geometric structures, and the inference of geometric or topological properties of sampled shapes. The need to understand geometric structures is ubiquitous in science and has become an essential part of scientific computing and data analysis. Geometry understanding is by no means limited to three dimensions. Many applications in physics, biology, and engineering require a keen understanding of the geometry of a variety of higher dimensional spaces to capture concise information from the underlying often highly nonlinear structure of data. Our approach is complementary to manifold learning techniques and aims at developing an effective theory for geometric and topological data analysis. To reach these objectives, the guiding principle will be to foster a symbiotic relationship between theory and practice, and to address fundamental research issues along three parallel advancing fronts. We will simultaneously develop mathematical approaches providing theoretical guarantees, effective algorithms that are amenable to theoretical analysis and rigorous experimental validation, and perennial software development. We will undertake the development of a high-quality open source software platform to implement the most important geometric data structures and algorithms at the heart of geometry understanding in higher dimensions. The platform will be a unique vehicle towards researchers from other fields and will serve as a basis for groundbreaking advances in scientific computing and data analysis.

9.3. International Initiatives

9.3.1. Inria Associate Teams Not Involved in an Inria International Labs

9.3.1.1. CATS

Title: Computations And Topological Statistics

International Partner (Institution - Laboratory - Researcher):

Carnegie Mellon University (United States) - Department of Statistics - Larry Wasserman

Start year: 2015

See also: http://geometrica.saclay.inria.fr/collaborations/CATS/CATS.html

Topological Data Analysis (TDA) is an emergent field attracting interest from various communities, that has recently known academic and industrial successes. Its aim is to identify and infer geometric and topological features of data to develop new methods and tools for data exploration and data analysis. TDA results mostly rely on deterministic assumptions which are not satisfactory from a statistical viewpoint and which lead to a heuristic use of TDA tools in practice. Bringing together the strong expertise of two groups in Statistics (L. Wasserman's group at CMU) and Computational Topology and Geometry (Inria Geometrica), the main objective of CATS is to set-up the mathematical foundations of Statistical TDA to design new TDA methods and to develop efficient and easy-to-use software tools for TDA.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

Ramsay Dyer (April and November 2016)

Arijit Ghosh, Indian Statistical Institute, Kolkata (April and November 2016)

Jose Carlos Gomez Larranaga, CIMAT, Guanajuato, Mexico (September 2016)

Kim Jisu, CMU, Pittsburgh, USA (May and December 2016).

Antony Bak, Palantir company, USA (October 2016)

9.4.1.1. Internships

Uday Kusupati, Indian Institute of Technology, Bombay (May-July 2016)

Sandip Banerjee (bourse Charpak), Indian Statistical Institute, Kolkata (March-August 2016)

Sameer Desai, Indian Statistical Institute, Kolkata (October-December 2016)

9.4.2. Visits to International Teams

9.4.2.1. Research Stays Abroad

Steve Oudot and Jérémy Cochoy spent 3 months (Sept.-Nov.) at the Institute for Computational and Experimental Research in Mathematics (ICERM) at Brown University. They were invited there for the semester program entitled *Topology in Motion* (see https://icerm.brown.edu/programs/sp-f16/).

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

Jean-Daniel Boissonnat and Frédéric Chazal co-organized the joint GUDHI-TOPDATA workshop in Porquerolles, October 17-20.

Frédéric Chazal co-organized the SMAI-SIGMA Conference 2016 at Luminy (CIRM) in November.

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

Maks Ovsjanikov: Paper co-chair of the Symposium on Geometry Processing 2016 (SGP 2016).

10.1.2.2. Member of the Conference Program Committees

Frédéric Chazal: Symposium on Geometry Processing 2016 (SGP 2016).

Steve Oudot: Symposium on Geometry Processing 2016 (SGP 2016).

Marc Glisse: Symposium on Computational Geometry (SoCG 2016).

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Jean-Daniel Boissonnat is a member of the Editorial Board of *Journal of the ACM*, *Discrete and Computational Geometry*, *International Journal on Computational Geometry and Applications*.

Frédéric Chazal is a member of the Editorial Board of *SIAM Journal on Imaging Sciences*, *Discrete and Computational Geometry (Springer)*, *Graphical Models (Elsevier), and Journal of Applied and Computational Topology (Springer)*.

Steve Oudot is a member of the Editorial Board of *Journal of Computational Geometry*.

10.1.4. Invited Talks

Frédéric Chazal, ACCAPT conference, Aalborg, Danmark, April 2016.

Frédéric Chazal, Joint Mathematical Meetings, Seattle, USA, January 2016.

Frédéric Chazal, Séminaire Parisien de Géométrie Algorithmique, Paris, October 2016.

Frédéric Chazal, 9th International Conference of theERCIM WG on Computational and Methodological Statistics, December 2016.

Steve Oudot, ACCAPT conference, Aalborg, Danmark, April 2016.

Steve Oudot, Workshop SIGMA 2016, CIRM, France, November 2016.

Steve Oudot, Applied Topology Seminar, Brown University, USA, November 2016.

Steve Oudot, Topology and Neuroscience Seminar, Princeton University, USA, November 2016.

10.1.5. Scientific Expertise

Frédéric Chazal was a member of the ANR committee, CES 40 (Mathematics and Computer Science).

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master : Frédéric Chazal, Analyse Topologique des Données, 30h eq-TD, Université Paris-Sud, France.

Master : Jean-Daniel Boissonnat and Marc Glisse, Computational Geometry Learning, 36h eq-TD, M2, MPRI, France.

Doctorat : Frédéric Chazal and Bertrand Michel, An introduction to Topological Data Analysis, 18h eq-TD, Universitad Autonoma de Barcelona, Spain.

Master : Steve Oudot, Topological Data Analysis, 45h eq-TD, M1, École Polytechnique, France.

Master : Steve Oudot and Frédéric Cazals, Geometric Methods for Data Analysis, 30h eq-TD, M1, École Centrale Paris, France.

Master : Jean-Daniel Boissonnat, Winter School on Computational geometry and Topology, Inria Sophia Antipolis Méditerrannée, January 2016.

Doctorat : Steve Oudot, École Mathématique en Afrique on *Topologie différentielle, géométrie algébrique et applications*, La Marsa, Tunisia, March-April 2016.

Doctorat : Steve Oudot, Summer School on Mathematical Methods for High-Dimensional Data Analysis, Technical University of Munich, Germany, July 2016.

10.2.2. Supervision

PhD: Thomas Bonis, Statistical Learning Algorithms for Geometric and Topological Data Analysis, December 1st, 2016, Frédéric Chazal.

PhD : Mael Rouxel-Labbé, Génération de maillages anisotropes, december 16, 2016, Jean-Daniel Boissonnat.

PhD: Ruqi Huang, Algorithms for topological inference in metric spaces, December 14, 2016, Frédéric Chazal.

PhD in progress: Eddie Aamari, A Statistical Approach of Topological Data Analysis, started September 1st, 2014, Frédéric Chazal (co-advised by Pascal Massart).

PhD in progress: Claire Brécheteau, Statistical aspects of distance-like functions , started September 1st, 2015, Frédéric Chazal (co-advised by Pascal Massart).

PhD in progress: Bertrand Beaufils, Méthodes topologiques et apprentissage statistique pour l'actimétrie du piéton à partir de données de mouvement, started November 2016, Frédéric Chazal (co-advised by Bertrand Michel).

PhD in progress: Mathieu Carrière, Topological signatures for geometric data, started November 1st, 2014, Steve Oudot.

PhD in progress: Jérémy Cochoy, Decomposition and stability of multidimensional persistence modules, started September 1st, 2015, Steve Oudot.

PhD in progress: Nicolas Berkouk, Categorification of topological graph structures, started November 1st, 2016, Steve Oudot.

PhD in progress: Alba Chiara de Vitis, Concentration of measure and clustering.

PhD in progress: Siargey Kachanovich, Approximate algorithms in higher dimensional geometry.

PhD in progress: François Godi, Data structures and algorithms for topological data analysis and high dimensional geometry.

10.2.3. *Juries*

Frédéric Chazal was a member (and reviewer) of the PhD defense committee of Mariia Fodetenkova (Inria Nancy).

11. Bibliography

Major publications by the team in recent years

[1] D. ATTALI, U. BAUER, O. DEVILLERS, M. GLISSE, A. LIEUTIER.*Homological Reconstruction and Simplification in R3*, in "Computational Geometry", 2014 [*DOI : 10.1016/J.COMGEO.2014.08.010*], https://hal.archives-ouvertes.fr/hal-01132440.

[2] J. BOISSONNAT, R. DYER, A. GHOSH.*Stability of Delaunay-type structures for manifolds: [extended abstract]*, in "Symposium on Computational Geometry 2012, SoCG '12, Chapel Hill, NC, USA, June 17-20, 2012", 2012, p. 229–238, http://doi.acm.org/10.1145/2261250.2261284.

[3] J.-D. BOISSONNAT, A. GHOSH.*Manifold reconstruction using tangential Delaunay complexes*, in "Discrete and Computational Geometry", January 2014, vol. 51, n° 1, p. 221-267 [*DOI : 10.1007/S00454-013-9557-2*], https://hal.inria.fr/hal-00932209.

[4] J. BOISSONNAT, C. MARIA.*The Simplex Tree: An Efficient Data Structure for General Simplicial Complexes*, in "Algorithmica", 2014, vol. 70, n° 3, p. 406–427, http://dx.doi.org/10.1007/s00453-014-9887-3.

[5] F. CHAZAL, D. COHEN-STEINER, M. GLISSE, L. J. GUIBAS, S. OUDOT.*Proximity of Persistence Modules and Their Diagrams*, in "Proc. 25th Annual Symposium on Computational Geometry", 2009, p. 237–246 [*DOI : 10.1145/1542362.1542407*], http://hal.inria.fr/inria-00292566/.

[6] F. CHAZAL, D. COHEN-STEINER, A. LIEUTIER.*A Sampling Theory for Compact Sets in Euclidean Space*, in "Discrete Comput. Geom.", 2009, vol. 41, n° 3, p. 461–479, http://dx.doi.org/10.1007/s00454-009-9144-8.

[7] F. CHAZAL, D. COHEN-STEINER, Q. MÉRIGOT.*Geometric Inference for Measures based on Distance Functions*, in "Foundations of Computational Mathematics", 2011, vol. 11, n° 6, p. 733-751, RR-6930 [*DOI : 10.1007/S10208-011-9098-0*], http://hal.inria.fr/inria-00383685.

[8] F. DE GOES, D. COHEN-STEINER, P. ALLIEZ, M. DESBRUN.*An Optimal Transport Approach to Robust Reconstruction and Simplification of 2D Shapes*, in "Computer Graphics Forum", 2011, vol. 30, n⁰ 5, p. 1593–1602, Special issue for EUROGRAPHICS Symposium on Geometry Processing, http://hal.inria.fr/hal-00758019.

[9] L. J. GUIBAS, S. Y. OUDOT, P. SKRABA, F. CHAZAL.*Persistence-Based Clustering in Riemannian Manifolds*, in "Journal of the ACM", November 2013, vol. 60, n⁰ 6, 38, http://hal.archives-ouvertes.fr/hal-00923563.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[10] T. BONIS.*Statistical learning algorithms for topological and geometric data analysis*, Université Paris-Saclay, December 2016, https://hal.archives-ouvertes.fr/tel-01402801.

[11] R. HUANG.*Two contributions to geometric data analysis: filamentary structures approximations, and stability properties of functional approaches for shape comparison*, University Paris-Saclay, December 2016, https://hal.archives-ouvertes.fr/tel-01419321.

[12] M. ROUXEL-LABBÉ.*Anisotropic Mesh Generation*, Université Côte D'Azur, December 2016, https://hal.archives-ouvertes.fr/tel-01419457.

Articles in International Peer-Reviewed Journal

[13] D. ATTALI, O. DEVILLERS, M. GLISSE, S. LAZARD.*Recognizing Shrinkable Complexes Is NP-Complete*, in "Journal of Computational Geometry", 2016, vol. 7, n⁰ 1, p. 430–443 [*DOI : 10.20382/JOCG.V7I1A18*], https://hal.inria.fr/hal-01384396.

[14] J.-D. BOISSONNAT, D. MAZAURIC.*On the complexity of the representation of simplicial complexes by trees*, in "Theoretical Computer Science", February 2016, vol. 617, 17 [*DOI : 10.1016/J.TCS.2015.12.034*], https://hal.inria.fr/hal-01259806.

[15] J.-D. BOISSONNAT, K. C. SRIKANTA, S. TAVENAS.*Building Efficient and Compact Data Structures for Simplicial Complexe*, in "Algorithmica", September 2016 [*DOI : 10.1007/S00453-016-0207-Y*], https://hal.inria.fr/hal-01364648.

[16] P. BUBENIK, D. PAWEL.*A persistence landscapes toolbox for topological statistics*, in "Journal of Symbolic Computation", 2016, https://hal.inria.fr/hal-01258875.

[17] M. BUCHET, F. CHAZAL, S. OUDOT, D. SHEEHY.*Efficient and Robust Persistent Homology for Measures*, in "Computational Geometry", 2016, vol. 58 [*DOI : 10.1016/J.COMGEO.2016.07.001*], https://hal.inria.fr/hal-01342385.

[18] F. CHAZAL, B. FASY, F. LECCI, B. MICHEL, A. RINALDO, L. WASSERMAN.*Robust Topological Inference: Distance To a Measure and Kernel Distance*, in "Journal of Machine Learning Research", 2017, forthcoming, https://hal.inria.fr/hal-01232217.

[19] F. CHAZAL, P. MASSART, B. MICHEL.*Rates of convergence for robust geometric inference*, in "Electronic journal of statistics ", 2016, vol. 10, n⁰ 2, 44, https://hal.inria.fr/hal-01336913.

[20] O. DEVILLERS, M. GLISSE, X. GOAOC, R. THOMASSE.*Smoothed complexity of convex hulls by witnesses and collectors*, in "Journal of Computational Geometry", 2016, vol. 7, n° 2, p. 101-144 [*DOI :* 10.20382/JOCG.V7I2A6], https://hal.inria.fr/hal-01285120.

[21] K. DUTTA, E. EZRA, A. GHOSH.*Two proofs for Shallow Packings* , in "Discrete and Computational Geometry", December 2016 [*DOI :* 10.1007/s00454-016-9824-0], https://hal.archives-ouvertes.fr/hal-01360460.

International Conferences with Proceedings

[22] J.-D. BOISSONNAT, K. C. SRIKANTA.*An Efficient Representation for Filtrations of Simplicial Complexes*, in "Symposium on Discrete Algorithms SODA 2017", Barcelona, France, January 2017, https://hal.inria.fr/hal-01416683.

[23] T. BONIS, M. OVSJANIKOV, S. OUDOT, F. CHAZAL.*Persistence-based Pooling for Shape Pose Recognition*, in "6th International Workshop on Computational Topology in Image Context (CTIC 2016)", Marseille, France, CTIC 2016 Proceedings of the 6th International Workshop on Computational Topology in Image Context, June 2016, vol. 9667, https://hal.archives-ouvertes.fr/hal-01301937.

[24] D. BREMNER, O. DEVILLERS, M. GLISSE, S. LAZARD, G. LIOTTA, T. MCHEDLIDZE, S. WHITESIDES, S. WISMATH.*Monotone Simultaneous Paths Embeddings in $\mathbb{R}^d$*, in "24th International Symposium on Graph Drawing & Network Visualization", Athens, Greece, Proceedings of 24th International Symposium on Graph Drawing & Network Visualization, Springer, September 2016, vol. 9801, https://hal.inria.fr/hal-01366148.

[25] M. CARRIÈRE, S. OUDOT.*Structure and Stability of the 1-Dimensional Mapper*, in "Proceedings of the International Symposium on Computational Geometry", Boston, United States, June 2016, Conference version, https://hal.inria.fr/hal-01247511.

[26] F. CHAZAL, I. GIULINI, B. MICHEL.*Data driven estimation of Laplace-Beltrami operator*, in "30th Conference on Neural Information Processing Systems (NIPS 2016)", Barcelona, Spain, December 2016, https://hal.inria.fr/hal-01387021.

[27] M. ROUXEL-LABBÉ, M. WINTRAECKEN, J.-D. BOISSONNAT.*Discretized Riemannian Delaunay triangulations*, in "Proceedings 25th International Meshing Roundtable (IMR25)", Washington DC, United States, Elsevier, September 2016, https://hal.inria.fr/hal-01367525.

Conferences without Proceedings

[28] K. DUTTA, A. GHOSH.*On Subgraphs of Bounded Degeneracy in Hypergraphs*, in "42nd International Workshop on Graph-Theoretic Concepts in Computer Science (WG 2016)", Istanbul, Turkey, June 2016, https://hal.archives-ouvertes.fr/hal-01360425.

Scientific Books (or Scientific Book chapters)

[29] F. CHAZAL, S. Y. OUDOT, M. GLISSE, V. DE SILVA.*The Structure and Stability of Persistence Modules*, SpringerBriefs in Mathematics, Springer Verlag, 2016, VII, 116, https://hal.inria.fr/hal-01330678.

[30] F. CHAZAL.*High-Dimensional Topological Data Analysis*, in "3rd Handbook of Discrete and Computational Geometry", CRC Press, 2016, forthcoming, https://hal.inria.fr/hal-01316989.

[31] F. CHAZAL, D. COHEN-STEINER, A. LIEUTIER, Q. MERIGOT, B. THIBERT.*Inference of curvature using tubular neighborhoods*, in "Lecture Notes in Mathematics", 2017, forthcoming, https://hal.archives-ouvertes.fr/hal-01425558.

Other Publications

[32] E. AAMARI, C. LEVRARD.*Stability and Minimax Optimality of Tangential Delaunay Complexes for Manifold Reconstruction*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01245479.

[33] T. BONIS.*Rates in the Central Limit Theorem and diffusion approximation via Stein's Method*, July 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01167372.

[34] C. BRÉCHETEAU.*The DTM-signature for a geometric comparison of metric-measure spaces from samples*, January 2017, working paper or preprint, https://hal.inria.fr/hal-01426331.

[35] P. DLOTKO, K. HESS, L. RAN, H. MARKRAM, E. MULLER, M. NOLTE, M. REIMANN, M. SCOLAMIERO, K. TURNER.*Topological analysis of the connectome of digital reconstructions of neural microcircuits*, January 2016, working paper or preprint, https://hal.inria.fr/hal-01261011.

[36] P. DLOTKO, T. WANNER.*Topological Microstructure Analysis Using Persistence Landscapes*, January 2016, working paper or preprint, https://hal.inria.fr/hal-01259127.

[37] K. DUTTA, A. GHOSH, B. JARTOUX, N. MUSTAFA.*Combinatorics of Set Systems with Small Shallow Cell Complexity: Optimal Bounds via Packings*, September 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01360443.

[38] K. DUTTA, A. GHOSH, N. MUSTAFA.*A new asymmetric correlation inequality for Gaussian measure*, September 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01360457.

[39] C. JÉRÉMY, S. Y. OUDOT.*Decomposition of exact pfd persistence bimodules*, May 2016, working paper or preprint, https://hal.inria.fr/hal-01359312.

[40] N. MUSTAFA, K. DUTTA, A. GHOSH.*A Simple Proof of Optimal Epsilon Nets*, September 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01360452.

Team DEDUCTEAM

Deduction modulo, interopérabilité et démonstration automatique

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

Table of contents

Team DEDUCTEAM

Creation of the Team: 2011 December 01, updated into Project-Team: 2017 January 01

Keywords:

Computer Science and Digital Science:
   2. - Software
   2.1.2. - Object-oriented programming
   2.1.3. - Functional programming
   2.1.11. - Proof languages
   2.4.3. - Proofs
   3.1.1. - Modeling, representation
   7. - Fundamental Algorithmics
   7.4. - Logic in Computer Science
   7.13. - Quantum algorithms

Other Research Topics and Application Domains:
   7. - Transport and logistics

1. Members

Research Scientists
  Gilles Dowek [Team Leader, Inria, Senior Researcher, HDR]
  Frédéric Blanqui [Inria, Researcher, HDR]

PhD Students
  Guillaume Bury [Univ. Paris VII]
  Frédéric Gilbert [Min. Ecologie]
  François Thiré [ENS Cachan, since October 2016]
  Raphaël Cauderlier [CNAM, until August 2016]
  Pierre Halmagrand [CNAM, until November 2016]

Post-Doctoral Fellow
  Simon Martiel [Univ. Paris-Est Créteil]

Visiting Scientists
  Jean-Pierre Jouannaud [Professor Emeritus, Univ. Paris-Saclay, HDR]
  Guillaume Burel [ENSIIE, Associate Professor]
  Catherine Dubois [ENSIIE, Professor, HDR]
  Olivier Hermant [Mines ParisTech, Researcher]

Administrative Assistant
  Thida Iem [Inria]

2. Overall Objectives

2.1. Objectives

The team investigates applications of recent results in proof theory to the design of logical frameworks and automated theorem proving systems. It develops the Dedukti logical framework and the iProver modulo and Zenon modulo automated theorem proving systems.

2.2. History

Deduction modulo is a formulation of predicate logic where deduction is performed modulo an equivalence relation defined on propositions. A typical example is the equivalence relation relating propositions differing only by a re-arrangement of brackets around additions, relating, for instance, the propositions $P((x + y) + z)$ and $P(x + (y + z))$. Reasoning modulo this equivalence relation permits to drop the associativity axiom. Thus, in Deduction modulo, a theory is formed with a set of axioms and an equivalence relation. When the set of axioms is empty the theory is called *purely computational*.

Deduction modulo was proposed at the end of the 20th century as a tool to simplify the completeness proof of equational resolution. Soon, it was noticed that this idea was also present in other areas of logic, such as Martin-Löf's type theory, where the equivalence relation is definitional equality, Prawitz' extended natural deduction, etc. More generally, Deduction modulo gives an account on the way reasoning and computation are articulated in a formal proof, a topic slightly neglected by logic, but of prime importance when proofs are computerized.

The early research on Deduction modulo focused on the design of general proof search methods—Resolution modulo, tableaux modulo, etc.—that could be applied to any theory formulated in Deduction modulo, to general proof normalization and cut elimination results, to the definitions of models taking the difference between reasoning and computation into account, and to the definition of specific theories—simple type theory, arithmetic, some versions of set theory, etc.—as purely computational theories.

3. Research Program

3.1. From proof-checking to Interoperability

A new turn with Deduction modulo was taken when the idea of reasoning modulo an arbitrary equivalence relation was applied to typed λ-calculi with dependent types, that permits to express proofs as algorithms, using the Brouwer-Heyting-Kolmogorov interpretation and the Curry-de Bruijn-Howard correspondence [27]. It was shown in 2007, that extending the simplest λ-calculus with dependent types, the $\lambda\Pi$-calculus, with an equivalence relation (more precisely a coingruence), led to a calculus we called the $\lambda\Pi$-calculus modulo, that permitted to simulate many other λ-calculi, such as the Calculus of Constructions, designed to express proofs in specific theories.

This led to the development of a general proof-checker based on the $\lambda\Pi$-calculus modulo [3], that could be used to verify proofs coming from different proof systems, such as Coq [26], HOL [33], etc. To emphasize this versatility of our proof-system, we called it Dedukti —"to deduce" in Esperanto. This system is currently developed together with companion systems, Coqine, Krajono, Holide, Focalide, and Zenonide, that permits to translate proofs from Coq, HOL, Focalize, and Zenon, to Dedukti. Other tools, such as Zenon Modulo, directly output proofs that can be checked by Dedukti. Dedukti proofs can also be exported to other systems, in particular to the MMT format [37].

A thesis, which is at the root of our research effort, and which was already formulated in [32] is that proof-checkers should be theory independent. This is for instance expressed in the title of our invited talk at Icalp 2012: *A theory independent Curry-De Bruijn-Howard correspondence*. Such a theory independent proof-checker is called a *Logical Framework*.

Using a single prover to check proofs coming from different provers naturally led to investigate how these proofs could interact one with another. This issue is of prime importance because developments in proof systems are getting bigger and, unlike other communities in computer science, the proof-checking community has given little effort in the direction of standardization and interoperability. On a longer term we believe that, for each proof, we should be able to identify the systems in which it can be expressed.

3.2. Automated theorem proving

Deduction modulo has originally been proposed to solve a problem in automated theorem proving and some of the early work in this area focused on the design of an automated theorem proving method called *Resolution modulo*, but this method was so complex that it was never implemented. This method was simplified in 2010 [5] and it could then be implemented. This implementation that builds on the iProver effort [36] is called iProver modulo.

iProver modulo gave surprisingly good results [4], so that we use it now to search for proofs in many areas: in the theory of classes—also known as B set theory—, on finite structures, etc. Similar ideas have also been implemented for the tableau method with in particular several extensions of the Zenon automated theorem prover. More precisely, two extensions have been realized: the first one is called SuperZenon [35] [30] and is an extension to superdeduction (which is a variant of Deduction modulo), and the second one is called ZenonModulo [28], [29] and is an extension to Deduction modulo. Both extensions have been extensively tested over first-order problems (of the TPTP library), and also provide good results in terms of number of proved problems. In particular, these tools provide good performances in set theory, so that SuperZenon has been successfully applied to verify B proof rules of Atelier B (work in collaboration with Siemens). Similarly, we plan to apply ZenonModulo in the framework of the BWare project to verify B proof obligations coming from the modeling of industrial applications.

More generally, we believe that proof-checking and automated theorem proving have a lot to learn from each other, because a proof is both a static linguistic object justifying the truth of a proposition and a dynamic process of proving this proposition.

3.3. Models of computation

The idea of Deduction modulo is that computation plays a major role in the foundations of mathematics. This led us to investigate the role played by computation in other sciences, in particular in physics. Some of this work can be seen as a continuation of Gandy's [31] on the fact that the physical Church-Turing thesis is a consequence of three principles of physics, two well-known: the homogeneity of space and time, and the existence of a bound on the velocity of information, and one more speculative: the existence of a bound on the density of information.

This led us to develop physically oriented models of computations.

4. Application Domains

4.1. Safety of aerospace systems

In parallel with this effort in logic and in the development of proof checkers and automated theorem proving systems, we always have been interested in using such tools. One of our favorite application domain is the safety of aerospace systems. Together with César Muñoz' team in Nasa-Langley, we have proved the correctness of several geometric algorithms used in air traffic control.

This has led us sometimes to develop such algorithms ourselves, and sometimes to develop tools for automating these proofs.

4.2. B-set theory

Set theory appears to be an appropriate theory for automated theorem provers based on Deduction modulo, in particular the several extensions of Zenon (SuperZenon and ZenonModulo). Modeling techniques using set theory are therefore good candidates to assess these tools. This is what we have done with the B method whose formalism relies on set theory. A collaboration with Siemens has been developed to automatically verify the B proof rules of Atelier B [34]. From this work presented in the Doctoral dissertation of Mélanie Jacquel, the SuperZenon tool [35] [30] has been designed in order to be able to reason modulo the B set theory. As a sequel

of this work, we contribute to the BWare project whose aim is to provide a mechanized framework to support the automated verification of B proof obligations coming from the development of industrial applications. In this context, we have recently designed ZenonModulo [28], [29] (Pierre Halmagrand's PhD thesis, which has started on October 2013) to deal with the B set theory. In this work, the idea is to manually transform the B set theory into a theory modulo and provide it to ZenonModulo in order to verify the proof obligations of the BWare project.

4.3. Termination certificate verification

Termination is an important property to verify, especially in critical applications. Automated termination provers use more and more complex theoretical results and external tools (e.g. sophisticated SAT solvers) that make their results not fully trustable and very difficult to check. To overcome this problem, a language for termination certificates, called CPF, has been developed since several years now. Deducteam develops a formally certified tool, Rainbow, based on the Coq library CoLoR, that is able to automatically verify the correctness of such termination certificates.

5. New Software and Platforms

5.1. Software of the team

Deducteam develops several kinds of tools or libraries:

- Proof checkers:
    - Dedukti: proof checker for the $\lambda\Pi$-calculus modulo rewriting
    - Sukerujo: extension of Dedukti with syntactic constructions for records, strings, lists, etc.
    - Rainbow: CPF termination certificate verifier
- Tools for translating into Dedukti's proof format proofs coming from various other provers:
    - Coqine translates Coq proofs
    - Focalide translates Focalize proofs
    - Holide translates OpenTheory proofs (HOL-Light, HOL4, ProofPower)
    - Krajono translates Matita proofs
    - Sigmaid translates ς-calculus
- Automated theorem provers:
    - iProverModulo: theorem prover based on polarized resolution modulo
    - SuperZenon: extension of Zenon using superdeduction
    - ZenonArith: extension of Zenon using the simplex algorithm for arithmetic
    - ZenonModulo: extension of Zenon using deduction modulo and producing Dedukti proofs
    - Zipperposition: superposition prover featuring arithmetic and induction
    - HOT: automated termination prover for higher-order rewrite systems
    - Archsat: theorem prover using tableaux-like rules with a SAT core
- Libraries or generation tools:
    - CoLoR: Coq library on rewriting theory and termination
    - Logtk: library for first-order automated reasoning
    - mSat: modular SAT/SMT solver with proof output
    - Moca: generator of construction functions for types with relations on constructors

5.2. Novelties of the year

The main novelties this year are:

- CoLoR has been ported to Coq 8.5.
- F. Blanqui started to develop a prototype for developing Dedukti proofs interactively.

6. New Results

6.1. Dedukti

A. Assaf, G. Burel, R. Cauderlier, D. Delahaye, G. Dowek, C. Dubois, F. Gilbert, P. Halmagrand, O. Hermant, and R. Saillard, have finished writing a general presentation of the Dedukti system. This paper is submitted for publication.

Under the supervision of P. Halmagrand and G. Burel, D. Pham worked on the conversion of TSTP proof traces, as produced by automated theorem provers such as E, Zipperposition or Vampire, into Dedukti proofs. To that purpose, he modified Zenon modulo so that it reads TSTP files and tries to reprove the proof steps given by the trace.

R. Cauderlier defended his PhD thesis on the translation of programming languages to Dedukti and interoperability of proof systems [11]. He also presented his work on the use of Dedukti for rewriting-based proof transformation [15] and on the translation of FoCaLiZe in Dedukti [16]

6.2. Proof theory

G. Dowek and Y. Jiang have finished a paper on co-inductive and inductive complementation of inference systems. This paper is submitted for publication.

The paper of G. Dowek on the introduction of rules and derivations in a logic course has been published [24].

F. Gilbert has finished a paper on the automated constructivization of proofs, to appear in the proceedings of FOSSACS'17.

F. Thiré is working on the translation of the Fermat little theorem proof written in Matita to a proof written in HOL. A part of this work is developed in its internship report [25]. He is continuing this translation during his PhD thesis.

6.3. B Method

The B Method is a formal method mainly used in the railway industry to specify and develop safety-critical software. To guarantee the consistency of a B project, one decisive challenge is to show correct a large amount of proof obligations, which are mathematical formulas expressed in a classical set theory extended with a specific type system. To improve automated theorem proving in the B Method, Pierre Halmagrand proposes [17], [12] to use a first-order sequent calculus extended with a polymorphic type system, which is in particular the output proof-format of the tableau-based automated theorem prover Zenon. After stating some modifications of the B syntax and defining a sound elimination of comprehension sets, he proposes a translation of B formulas into a polymorphic first-order logic format. Then, he introduces the typed sequent calculus used by Zenon, and shows that Zenon proofs can be translated to proofs of the initial B formulas in the B proof system.

6.4. Termination

F. Blanqui revised his paper on "size-based termination of higher-order rewrite systems" submitted to the Journal of Functional Programming [23]. This paper is concerned with the termination, in Church' simply-typed λ-calculus, of the combination of β-reduction and arbitrary user-defined rewrite rules fired using matching modulo α-congruence only. Several authors have devised termination criteria for fixpoint-based function definitions using deduction rules for bounding the size of terms inhabiting inductively defined types, where the size of a term is (roughly speaking) the set-theoretical height of the tree representation of its normal form. In the present paper, we extend this approach to rewriting-based function definitions and more general notions of size.

G. Dowek has finished writing a paper on the notion of model and its application to termination proofs for the $\lambda\Pi$-calculus modulo theory. This paper is submitted for publication.

6.5. Confluence

In $\lambda\Pi$modulo, congruences are expressed by rewrite rules that must enjoy precise properties, notably confluence, strong normalization, and type preservation. A difficulty is that these properties depend on each other in calculi of dependent types. To break the circularity, confluence is usually proved separately on untyped terms. A another difficulty then arises : computation do not terminate on untyped terms. A result of van Oostrom allows to show confluence of non-terminating left-linear higher-order rules, provided their critical pairs are development closed. This result was used for the encodings of HOL, Matita, and Coq up to version 8.4. Encoding the most recent version of Coq requires rules for universes that are confluent on open terms, while confluence on ground terms sufficed before. The encoding we recently developed for this new version of Coq has higher-order rules which are not left-linear, use pattern matching modulo associativity, commutativity and identity, and whose (joinable) critical pairs are not development closed. We have therefore developed a new powerful result for proving confluence of that sort of rules provided non-linear variables can only be instantiated by first-order expressions [18], [19].

6.6. Physics and computation

The paper of G. Dowek and P. Arrighi Free fall and cellular automata has been published [13]. As a sequel of this paper, G. Dowek and P. Arrighi have written a short note [22].

A. Díaz-Caro and G. Dowek have developed a new typing system for quantum λ-calculus allowing to distinguish between pure states and superpositions.

Under the supervision of S. Martiel and P. Arrighi, C. Chouteau worked on a particular notion of covariance in the model of causal graph dynamics. Causal graph dynamics are graph transformations constrained by Physics-inspired symmetries. The particular object of study of this internship was a restriction of this model to physical transformations of discrete geometrical spaces.

7. Partnerships and Cooperations

7.1. National Initiatives

7.1.1. ANR Locali

We are coordinators of the ANR-NFSC contract Locali with the Chinese Academy of Sciences.

7.1.2. ANR BWare

We are members of the ANR BWare, which started on September 2012 (David Delahaye is the national leader of this project). The aim of this project is to provide a mechanized framework to support the automated verification of proof obligations coming from the development of industrial applications using the B method. The methodology used in this project consists in building a generic platform of verification relying on different theorem provers, such as first-order provers and SMT solvers. We are in particular involved in the introduction of Deduction modulo in the first-order theorem provers of the project, i.e. Zenon and iProver, as well as in the backend for these provers with the use of Dedukti.

7.1.3. ANR Tarmac

We are members of the ANR Tarmac on models of computation, coordinated by Pierre Valarcher.

7.2. European Initiatives

7.2.1. Collaborations in European Programs, Except FP7 & H2020

Program: CA COST Action CA15123

Project acronym: EUTYPES

Project title: European research network on types for programming and verification

Duration: 21/03/16 - 20/03/20

Coordinator: Herman Geuvers

7.3. International Initiatives

7.3.1. Participation in Other International Programs

Login

Title: Logic and Information

International Partner (Institution - Laboratory - Researcher):

Universidad de Buenos Aires (Argentina) - Ricardo Oscar Rodrigues

Duration: 2015 - 2016

This project aims to propose an improvement on a long-term already existing collaboration between Inria, the brazilians and the argentin named team. We already have a CAPES-COFECUB cooperation (n. 690/10, namely "Teorias lógicas contemporâneas e a filosofia da linguagem: questões epistemológicas e semânticas") that leaded to many students interchange and technical visits of Professors, including the organisation of some workshops (the last one was the 2nd Workshop on Logic and Semantics, at UERJ, Ilha Grande-RJ, Brazil. Prof. Gilles Dowek is also a Co-Advisor with Prof. Edward Hermann Haeusler of a brazilian Ph.D. Candidate in this project (and a former one also in this project, these two candidates finalised recently a sandwich doctorate - similar to stage doctorale - at Inria). Prof. Gilles Dowek also collaborates with other members of this team and is supervising a post-doc project of another member. Since 2011 members of the team presents.

FoQCoSS

Title: Foundations of Quantum Computation: Syntax and Semantics

International Partners (Institution - Laboratory - Researcher):

Universidad Nacional de Quilmes (Argentina) - Alejandro Diaz-Caro

CNRS (France) - Simon Perdrix

Duration: 2016 - 2017

The design of quantum programming languages involves the study of many characteristics of languages which can be seen as special cases of classical systems: parallelism, probabilistic systems, non-deterministic systems, type isomorphisms, etc. This project proposes to study some of these characteristics, which are involved in quantum programming languages, but also have a more immediate utility in the study of nowadays systems. In addition, from a more foundational point of view, we are interested in the implications of computer science principles for quantum physics. For example, the consequences of the Church-Turing thesis for Bell-like experiments: if some of the parties in a Bell-like experiment use a computer to decide which measurements to make, then the computational resources of an eavesdropper have to be limited in order to have a proper observation of non-locality. The final aim is to open a new direction in the search for a framework unifying computer science and quantum physics.

7.4. International Research Visitors

7.4.1. Internships

- Clément Chouteau, from May 2016 to July 2016
- David Pham (Univ. Évry) from June 2016 to July 2016

7.4.2. Visits to International Teams

7.4.2.1. Research Stays Abroad

F. Gilbert spent one month in the formal methods team at NASA Langley Research Center, to work with Cesar Munoz on the use of automated theorem provers to verify PVS proofs.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Organisation

8.1.1.1. General Chair, Scientific Chair

G. Dowek has co-organized the meeting Universality of Proofs in Dagstuhl.

8.1.1.2. Member of the Organizing Committees

G. Dowek is a member of the steering committee of FSCD.

8.1.2. Scientific Events Selection

8.1.2.1. Member of the Conference Program Committees

F. Blanqui was member of the program committee of the 2016 Coq Workshop.

8.1.2.2. Reviewer

F. Blanqui reviewed papers for IJCAR 2016 and CSL 2016.

8.1.3. Journal

8.1.3.1. Member of the Editorial Boards

G. Dowek is an editor of TCS-C.

8.1.4. Invited Talks

G. Dowek has been an invited speaker at ISEEP 2016.

G. Dowek has been an invited speaker at Physics and Computation 2016.

8.1.5. *Scientific Expertise*

G. Dowek has been a member of a commitee dedicated to an update of the high school informatics curriculum.

8.1.6. *Research Administration*

G. Dowek is the President of the Scientific Board of the Société informatique de France.

G. Dowek is a member of the Scientific Board of la Main à la Pâte.

G. Dowek is a member of the commission de réflexion sur l'éthique de la recherche en sciences et technologies du numérique d'Allistene.

G. Dowek is a member of the comité national français d'histoire et de philosophie des sciences et des techniques.

F. Blanqui is co-director of the pole 4 (programming: models, algorithms, languages and architectures) of Paris-Saclay University's doctoral school on computer science.

F. Blanqui is referent of LSV PhD students.

8.2. Teaching - Supervision - Juries

8.2.1. *Teaching*

G. Dowek is attached professor at the École normale supérieure de Paris-Saclay. He has given a course at MPRI. He has given a course to the student preparing the teacher's recruiting exam Agrégation. He is responsible for the second year of master.

F. Blanqui gave a course (15h) on rewriting theory at the MPRI.

8.2.2. *Supervision*

PhD : Raphaël Cauderlier, Object-Oriented Mechanisms for Interoperability between Proof Systems, CNAM, 10/10/2016, Catherine Dubois

8.2.3. *Juries*

F. Blanqui was member of the 2016 Inria recruitment committee for young graduate scientists.

F. Blanqui was member of the jury for the best scientific production of the year within Paris-Saclay University's doctoral school on computer science.

9. Bibliography

Major publications by the team in recent years

[1] F. BLANQUI.*Definitions by rewriting in the Calculus of Constructions*, in "Mathematical Structures in Computer Science", 2005, vol. 15, n⁰ 1, p. 37-92 [*DOI :* 10.1017/S0960129504004426], http://hal.inria.fr/inria-00105648/en/.

[2] F. BLANQUI, A. KOPROWSKI.*CoLoR: a Coq library on well-founded rewrite relations and its application to the automated verification of termination certificates*, in "Mathematical Structures in Computer Science", 2011, vol. 21, n⁰ 4, p. 827-859, http://hal.inria.fr/inria-00543157/en/.

[3] M. BOESPFLUG.*Conception d'un noyau de vérification de preuves pour le lambda-Pi-calcul modulo*, École Polytechnique, 2011.

[4] G. BUREL.*Experimenting with Deduction Modulo*, in "CADE 2011", V. SOFRONIE-STOKKERMANS, N. BJØRNER (editors), Lecture Notes in Artificial Intelligence, Springer, 2011, vol. 6803, p. 162–176.

[5] G. DOWEK.*Polarized Resolution Modulo*, in "IFIP Theoretical Computer Science", 2010.

[6] G. DOWEK, T. HARDIN, C. KIRCHNER.*Theorem proving modulo*, in "Journal of Automated Reasoning", 2003, vol. 31, p. 33-73.

[7] C. DUBOIS, T. HARDIN, V. DONZEAU-GOUGE.*Building certified components within FOCAL*, in "Revised Selected Papers from the Fifth Symposium on Trends in Functional Programming, TFP 2004, München, Germany, 25-26 November 2004", H.-W. LOIDL (editor), Trends in Functional Programming, Intellect, 2006, vol. 5, p. 33-48.

[8] O. HERMANT.*Resolution is Cut-Free*, in "Journal of Automated Reasoning", March 2010, vol. 44, n° 3, p. 245-276.

[9] M. JACQUEL, K. BERKANI, D. DELAHAYE, C. DUBOIS.*Verifying B Proof Rules using Deep Embedding and Automated Theorem Proving*, in "Software and Systems Modeling (SoSyM)", June 2013.

[10] M. JACQUEL, K. BERKANI, D. DELAHAYE, C. DUBOIS.*Tableaux Modulo Theories Using Superdeduction*, in "Global Journal of Advanced Software Engineering (GJASE)", December 2014, vol. 1, p. 1 - 13 [*DOI :* 10.1007/978-3-642-31365-3_26], https://hal.archives-ouvertes.fr/hal-01099338.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] R. CAUDERLIER.*Object-Oriented Mechanisms for Interoperability between Proof Systems*, Conservatoire National Des Arts et Métiers, Paris, October 2016, https://hal.inria.fr/tel-01415945.

[12] P. HALMAGRAND.*Automated Deduction and Proof Certification for the B Method*, Conservatoire National Des Arts et Métiers, Paris, December 2016, https://hal.inria.fr/tel-01420460.

Articles in International Peer-Reviewed Journal

[13] P. ARRIGHI, G. DOWEK.*Free fall and cellular automata*, in "Electronic Proceedings in Theoretical Computer Science", 2016, vol. 204, p. 1 - 10 [*DOI :* 10.4204/EPTCS.204.1], https://hal.inria.fr/hal-01421712.

International Conferences with Proceedings

[14] P. ARRIGHI, S. MARTIEL, S. PERDRIX.*Reversible Causal Graph Dynamics*, in "Reversible Computation", Bologna, Italy, Lecture Notes in Computer Science, July 2016, vol. 9720, p. 73-88 [*DOI :* 10.1007/978-3-319-40578-0_5], https://hal.archives-ouvertes.fr/hal-01361427.

[15] R. CAUDERLIER.*A Rewrite System for Proof Constructivization*, in "Workshop on Logical Frameworks and Meta-Languages: Theory and Practice 2016", Porto, Portugal, June 2016, p. 1 - 7 [*DOI :* 10.1145/2966268.2966270], https://hal.inria.fr/hal-01420634.

[16] R. CAUDERLIER, C. DUBOIS.*ML Pattern-Matching, Recursion, and Rewriting: From FoCaLiZe to Dedukti*, in "13th ICTAC International Colloquium on Theoretical Aspects of Computing", Taipei, Taiwan, October 2016, p. 459 - 468 [*DOI :* 10.1007/978-3-319-46750-4_26], https://hal.inria.fr/hal-01420638.

[17] P. HALMAGRAND.*Soundly Proving B Method Formulae Using Typed Sequent Calculus*, in "13th International Colloquium on Theoretical Aspects of Computing (ICTAC)", Taipei, Taiwan, A. SAMPAIO, F. WANG (editors), Lecture Notes in Computer Science, Springer International Publishing, October 2016, vol. 9965, p. 196-213 [*DOI :* 10.1007/978-3-319-46750-4_12], https://hal.archives-ouvertes.fr/hal-01342849.

Conferences without Proceedings

[18] A. ASSAF, G. DOWEK, J.-P. JOUANNAUD, J. LIU.*Encoding Proofs in Dedukti: the case of Coq proofs*, in "Proceedings Hammers for Type Theories", Coimbra, Portugal, Proc. Higher-Order rewriting Workshop, Easy Chair, July 2016, https://hal.inria.fr/hal-01330980.

[19] A. ASSAF, G. DOWEK, J.-P. JOUANNAUD, J. LIU.*Untyped Confluence in Dependent Type Theories*, in "Proceedings Higher-Order Rewriting Workshop", Porto, Portugal, Proc. Higher-Order rewriting Workshop, Easy-Chair, June 2016, https://hal.inria.fr/hal-01330955.

[20] K. JI.*Resolution in Solving Graph Problems*, in "8th Working Conference on Verified Software: Theories, Tools, and Experiments (VSTTE 2016)", Toronto, Canada, July 2016, https://hal.archives-ouvertes.fr/hal-01245138.

[21] S. WANG.*Higher Order Proof Engineering: Proof Collaboration, Transformation, Checking and Retrieval*, in "AITP 2016 - Conference on Artificial Intelligence and Theorem Proving", Obergurgl, Austria, April 2016, https://hal.inria.fr/hal-01250197.

Other Publications

[22] P. ARRIGHI, G. DOWEK.*What is the Planck constant the magnitude of?*, December 2016, working paper or preprint, https://hal.inria.fr/hal-01421711.

[23] F. BLANQUI.*Size-based termination of higher-order rewrite systems*, January 2017, working paper or preprint, https://hal.inria.fr/hal-01424921.

[24] G. DOWEK.*Rules and derivations in an elementary logic course*, January 2016, working paper or preprint, https://hal.inria.fr/hal-01252124.

[25] F. THIRÉ.*Internship report MPRI 2 Reverse engineering on arithmetic proofs*, ENS Cachan ; Paris Diderot University, September 2016, 26, https://hal.inria.fr/hal-01424816.

References in notes

[26] Y. BERTOT, P. CASTÉRAN.*Interactive Theorem Proving and Program Development Coq'Art: The Calculus of Inductive Constructions*, Springer-Verlag, 2004.

[27] D. COUSINEAU, G. DOWEK.*Embedding Pure Type Systems in the lambda-Pi-calculus modulo*, in "Typed lambda calculi and applications", S. RONCHI DELLA ROCCA (editor), Lecture Notes in Computer Science, Springer-Verlag, 2007, vol. 4583, p. 102-117.

[28] D. DELAHAYE, D. DOLIGEZ, F. GILBERT, P. HALMAGRAND, O. HERMANT.*Proof Certification in Zenon Modulo: When Achilles Uses Deduction Modulo to Outrun the Tortoise with Shorter Steps*, in "IWIL - 10th International Workshop on the Implementation of Logics - 2013", Stellenbosch, South Africa, S. SCHULZ, G. SUTCLIFFE, B. KONEV (editors), EasyChair, December 2013, https://hal.inria.fr/hal-00909688.

[29] D. DELAHAYE, D. DOLIGEZ, F. GILBERT, P. HALMAGRAND, O. HERMANT.*Zenon Modulo: When Achilles Outruns the Tortoise using Deduction Modulo*, in "LPAR - Logic for Programming Artificial Intelligence and Reasoning - 2013", Stellenbosch, South Africa, K. MCMILLAN, A. MIDDELDORP, A. VORONKOV (editors), Springer, December 2013, vol. 8312, p. 274-290 [*DOI :* 10.1007/978-3-642-45221-5_20], https://hal.inria.fr/hal-00909784.

[30] D. DELAHAYE, M. JACQUEL.*Recovering Intuition from Automated Formal Proofs using Tableaux with Superdeduction*, in "electronic Journal of Mathematics and Technology", February 2013, vol. 7, n° 2, p. 1 - 20, https://hal.archives-ouvertes.fr/hal-01099371.

[31] R. GANDY.*Church's Thesis and Principles for Mechanisms*, in "The Kleene Symposium", North-Holland, 1980.

[32] R. HARPER, F. HONSELL, G. PLOTKIN.*A Framework for Defining Logics*, in "Journal of the association for computing machinery", 1993, p. 194–204.

[33] J. HARRISON.*HOL Light: An Overview*, in "Theorem Proving in Higher Order Logics", S. BERGHOFER, T. NIPKOW, C. URBAN, M. WENZEL (editors), Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2009, vol. 5674, p. 60-66, http://dx.doi.org/10.1007/978-3-642-03359-9_4.

[34] M. JACQUEL, K. BERKANI, D. DELAHAYE, C. DUBOIS.*Verifying B Proof Rules using Deep Embedding and Automated Theorem Proving*, in "Software Engineering and Formal Methods", November 2011, vol. 7041, p. 253-268 [*DOI :* 10.1007/978-3-642-24690-6_18], https://hal.archives-ouvertes.fr/hal-00722373.

[35] M. JACQUEL, K. BERKANI, D. DELAHAYE, C. DUBOIS.*Tableaux Modulo Theories Using Superdeduction*, in "Global Journal of Advanced Software Engineering (GJASE)", December 2014, vol. 1, p. 1 - 13 [*DOI :* 10.1007/978-3-642-31365-3_26], https://hal.archives-ouvertes.fr/hal-01099338.

[36] K. KOROVIN.*iProver – An Instantiation-Based Theorem Prover for First-Order Logic (System Description)*, in "IJCAR", A. ARMANDO, P. BAUMGARTNER (editors), Lecture Notes in Artificial Intelligence, Springer, 2008, vol. 5195, p. 292-298.

[37] F. RABE, M. KOHLHASE.*A Scalable Module System*, in "Inf. Comput.", September 2013, vol. 230, p. 1–54, http://dx.doi.org/10.1016/j.ic.2013.06.001.

Project-Team DEFI

Shape reconstruction and identification

IN COLLABORATION WITH: Centre de Mathématiques Appliquées (CMAP)

IN PARTNERSHIP WITH:
CNRS

Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Numerical schemes and simulations

Table of contents

Project-Team DEFI

Creation of the Project-Team: 2009 January 01

Keywords:

Computer Science and Digital Science:

6. - Modeling, simulation and control
6.1. - Mathematical Modeling
6.1.1. - Continuous Modeling (PDE, ODE)
6.2. - Scientific Computing, Numerical Analysis & Optimization
6.2.1. - Numerical analysis of PDE and ODE
6.2.6. - Optimization
6.3. - Computation-data interaction
6.3.1. - Inverse problems

Other Research Topics and Application Domains:

1.3.1. - Understanding and simulation of the brain and the nervous system
2.6.1. - Brain imaging
3.3.1. - Earth and subsoil
5.3. - Nanotechnology

1. Members

Research Scientists
Houssem Haddar [Team leader, Inria, Senior Researcher]
Lucas Chesnel [Inria, Researcher]
Jing-Rebecca Li [Inria, Researcher, HDR]

Faculty Member
Grégoire Allaire [Ecole Polytechnique, Professor, HDR]

Technical Staff
Zixian Jiang [Inria, until Sep 2016, granted by SAXSIE- project]

PhD Students
Mathieu Chamaillard [Inria, until May 2016]
Bilel Charfi [Inria, until Mar 2016]
Gabrielle Fournet [Inria, until Nov 2016, granted by ANR CIACM project]
Matteo Giacomini [Ecole Polytechnique, until Oct 2016]
Mohamed Lakhal [Ecole Polytechnique]
Thi Phong Nguyen [Inria, granted by ANR METAMATH- project]
Van Khieu Nguyen [Univ. Paris XI]
Simona Schiavi [Inria]

Post-Doctoral Fellows
Marc Bakry [Inria]
Camille Carvalho [Inria, until Jun 2016]

Visiting Scientists
Fioralba Cakoni [Rutgers University]
Irene de Teresa-Trueba [University of Delaware]
Helle Majander [Aalto University]

Jacob Rezac [University of Delaware]
Semra Ahmetolan [Istanbul Technical University, from Feb 2016]
Administrative Assistants
    Marie Enee [Inria, until Nov 2016]
    Jessica Gameiro [Inria]
    Maeva Jeannot [Inria]
Others
    Lorenzo Audibert [EDF, Associate Member]
    Bumsu Kim [Ecole Polytechnique, from Jul 2016 until Nov 2016]
    Kevish Napal [Inria, from May 2016 until Oct 2016]
    Hoang An Tran [Ecole Polytechnique, from Apr 2016 until Jun 2016]

2. Overall Objectives

2.1. Overall Objectives

The research activity of our team is dedicated to the design, analysis and implementation of efficient numerical methods to solve inverse and shape/topological optimization problems in connection with acoustics, electromagnetism, elastodynamics, and diffusion.

Sought practical applications include radar and sonar applications, bio-medical imaging techniques, non-destructive testing, structural design, composite materials, and diffusion magnetic resonance imaging.

Roughly speaking, the model problem consists in determining information on, or optimizing the geometry (topology) and the physical properties of unknown targets from given constraints or measurements, for instance, measurements of diffracted waves or induced magnetic fields.

In general this kind of problems is non-linear. The inverse ones are also severely ill-posed and therefore require special attention from regularization point of view, and non-trivial adaptations of classical optimization methods.

Our scientific research interests are the following:

- Theoretical understanding and analysis of the forward and inverse mathematical models, including in particular the development of simplified models for adequate asymptotic configurations.

- The design of efficient numerical optimization/inversion methods which are quick and robust with respect to noise. Special attention will be paid to algorithms capable of treating large scale problems (e.g. 3-D problems) and/or suited for real-time imaging.

- Development of prototype softwares for specific applications or tutorial toolboxes.

During the last four years we were particularly interested in the development of the following themes that will be presented in details later.

- Qualitative methods for inverse scattering problems

- Iterative and Hybrid inversion methods

- Topological optimization methods

- Direct and inverse models for Diffusion MRI

- Asymptotic models and methods for waves and diffusion.

3. Research Program

3.1. Research Program

The research activity of our team is dedicated to the design, analysis and implementation of efficient numerical methods to solve inverse and shape/topological optimization problems in connection with wave imaging, structural design, non-destructive testing and medical imaging modalities. We are particularly interested in the development of fast methods that are suited for real-time applications and/or large scale problems. These goals require to work on both the physical and the mathematical models involved and indeed a solid expertise in related numerical algorithms.

This section intends to give a general overview of our research interests and themes. We choose to present them through the specific academic example of inverse scattering problems (from inhomogeneities), which is representative of foreseen developments on both inversion and (topological) optimization methods. The practical problem would be to identify an inclusion from measurements of diffracted waves that result from the interaction of the sought inclusion with some (incident) waves sent into the probed medium. Typical applications include biomedical imaging where using micro-waves one would like to probe the presence of pathological cells, or imaging of urban infrastructures where using ground penetrating radars (GPR) one is interested in finding the location of buried facilities such as pipelines or waste deposits. This kind of applications requires in particular fast and reliable algorithms.

By "imaging" we shall refer to the inverse problem where the concern is only the location and the shape of the inclusion, while "identification" may also indicate getting informations on the inclusion physical parameters.

Both problems (imaging and identification) are non linear and ill-posed (lack of stability with respect to measurements errors if some careful constrains are not added). Moreover, the unique determination of the geometry or the coefficients is not guaranteed in general if sufficient measurements are not available. As an example, in the case of anisotropic inclusions, one can show that an appropriate set of data uniquely determine the geometry but not the material properties.

These theoretical considerations (uniqueness, stability) are not only important in understanding the mathematical properties of the inverse problem, but also guide the choice of appropriate numerical strategies (which information can be stably reconstructed) and also the design of appropriate regularization techniques. Moreover, uniqueness proofs are in general constructive proofs, i.e. they implicitly contain a numerical algorithm to solve the inverse problem, hence their importance for practical applications. The sampling methods introduced below are one example of such algorithms.

A large part of our research activity is dedicated to numerical methods applied to the first type of inverse problems, where only the geometrical information is sought. In its general setting the inverse problem is very challenging and no method can provide a universal satisfactory solution to it (regarding the balance cost-precision-stability). This is why in the majority of the practically employed algorithms, some simplification of the underlying mathematical model is used, according to the specific configuration of the imaging experiment. The most popular ones are geometric optics (the Kirchhoff approximation) for high frequencies and weak scattering (the Born approximation) for small contrasts or small obstacles. They actually give full satisfaction for a wide range of applications as attested by the large success of existing imaging devices (radar, sonar, ultrasound, X-ray tomography, etc.), that rely on one of these approximations.

Generally speaking, the used simplifications result in a linearization of the inverse problem and therefore are usually valid only if the latter is weakly non-linear. The development of these simplified models and the improvement of their efficiency is still a very active research area. With that perspective we are particularly interested in deriving and studying higher order asymptotic models associated with small geometrical parameters such as: small obstacles, thin coatings, wires, periodic media, ... Higher order models usually introduce some non linearity in the inverse problem, but are in principle easier to handle from the numerical point of view than in the case of the exact model.

A larger part of our research activity is dedicated to algorithms that avoid the use of such approximations and that are efficient where classical approaches fail: i.e. roughly speaking when the non linearity of the inverse problem is sufficiently strong. This type of configuration is motivated by the applications mentioned below, and occurs as soon as the geometry of the unknown media generates non negligible multiple scattering effects (multiply-connected and closely spaces obstacles) or when the used frequency is in the so-called resonant region (wave-length comparable to the size of the sought medium). It is therefore much more difficult to deal with and requires new approaches. Our ideas to tackle this problem will be motivated and inspired by recent advances in shape and topological optimization methods and also the introduction of novel classes of imaging algorithms, so-called sampling methods.

The sampling methods are fast imaging solvers adapted to multi-static data (multiple receiver-transmitter pairs) at a fixed frequency. Even if they do not use any linearization the forward model, they rely on computing the solutions to a set of linear problems of small size, that can be performed in a completely parallel procedure. Our team has already a solid expertise in these methods applied to electromagnetic 3-D problems. The success of such approaches was their ability to provide a relatively quick algorithm for solving 3-D problems without any need for a priori knowledge on the physical parameters of the targets. These algorithms solve only the imaging problem, in the sense that only the geometrical information is provided.

Despite the large efforts already spent in the development of this type of methods, either from the algorithmic point of view or the theoretical one, numerous questions are still open. These attractive new algorithms also suffer from the lack of experimental validations, due to their relatively recent introduction. We also would like to invest on this side by developing collaborations with engineering research groups that have experimental facilities. From the practical point of view, the most potential limitation of sampling methods would be the need of a large amount of data to achieve a reasonable accuracy. On the other hand, optimization methods do not suffer from this constrain but they require good initial guess to ensure convergence and reduce the number of iterations. Therefore it seems natural to try to combine the two class of methods in order to calibrate the balance between cost and precision.

Among various shape optimization methods, the Level Set method seems to be particularly suited for such a coupling. First, because it shares similar mechanism as sampling methods: the geometry is captured as a level set of an "indicator function" computed on a cartesian grid. Second, because the two methods do not require any a priori knowledge on the topology of the sought geometry. Beyond the choice of a particular method, the main question would be to define in which way the coupling can be achieved. Obvious strategies consist in using one method to pre-process (initialization) or post-process (find the level set) the other. But one can also think of more elaborate ones, where for instance a sampling method can be used to optimize the choice of the incident wave at each iteration step.The latter point is closely related to the design of so called "focusing incident waves" (which are for instance the basis of applications of the time-reversal principle). In the frequency regime, these incident waves can be constructed from the eigenvalue decomposition of the data operator used by sampling methods. The theoretical and numerical investigations of these aspects are still not completely understood for electromagnetic or elastodynamic problems.

Other topological optimization methods, like the homogenization method or the topological gradient method, can also be used, each one provides particular advantages in specific configurations. It is evident that the development of these methods is very suited to inverse problems and provide substantial advantage compared to classical shape optimization methods based on boundary variation. Their applications to inverse problems has not been fully investigated. The efficiency of these optimization methods can also be increased for adequate asymptotic configurations. For instance small amplitude homogenization method can be used as an efficient relaxation method for the inverse problem in the presence of small contrasts. On the other hand, the topological gradient method has shown to perform well in localizing small inclusions with only one iteration.

A broader perspective would be the extension of the above mentioned techniques to time-dependent cases. Taking into account data in time domain is important for many practical applications, such as imaging in cluttered media, the design of absorbing coatings or also crash worthiness in the case of structural design.

For the identification problem, one would like to also have information on the physical properties of the targets. Of course optimization methods is a tool of choice for these problems. However, in some applications

only a qualitative information is needed and obtaining it in a cheaper way can be performed using asymptotic theories combined with sampling methods. We also refer here to the use of so called transmission eigenvalues as qualitative indicators for non destructive testing of dielectrics.

We are also interested in parameter identification problems arising in diffusion-type problems. Our research here is mostly motivated by applications to the imaging of biological tissues with the technique of Diffusion Magnetic Resonance Imaging (DMRI). Roughly speaking DMRI gives a measure of the average distance travelled by water molecules in a certain medium and can give useful information on cellular structure and structural change when the medium is biological tissue. In particular, we would like to infer from DMRI measurements changes in the cellular volume fraction occurring upon various physiological or pathological conditions as well as the average cell size in the case of tumor imaging. The main challenges here are 1) correctly model measured signals using diffusive-type time-dependent PDEs 2) numerically handle the complexity of the tissues 3) use the first two to identify physically relevant parameters from measurements. For the last point we are particularly interested in constructing reduced models of the multiple-compartment Bloch-Torrey partial differential equation using homogenization methods.

4. Application Domains

4.1. Radar and GPR applications

Conventional radar imaging techniques (ISAR, GPR, etc.) use backscattering data to image targets. The commonly used inversion algorithms are mainly based on the use of weak scattering approximations such as the Born or Kirchhoff approximation leading to very simple linear models, but at the expense of ignoring multiple scattering and polarization effects. The success of such an approach is evident in the wide use of synthetic aperture radar techniques.

However, the use of backscattering data makes 3-D imaging a very challenging problem (it is not even well understood theoretically) and as pointed out by Brett Borden in the context of airborne radar: "In recent years it has become quite apparent that the problems associated with radar target identification efforts will not vanish with the development of more sensitive radar receivers or increased signal-to-noise levels. In addition it has (slowly) been realized that greater amounts of data - or even additional "kinds" of radar data, such as added polarization or greatly extended bandwidth - will all suffer from the same basic limitations affiliated with incorrect model assumptions. Moreover, in the face of these problems it is important to ask how (and if) the complications associated with radar based automatic target recognition can be surmounted." This comment also applies to the more complex GPR problem.

Our research themes will incorporate the development, analysis and testing of several novel methods, such as sampling methods, level set methods or topological gradient methods, for ground penetrating radar application (imaging of urban infrastructures, landmines detection, underground waste deposits monitoring,) using multistatic data.

4.2. Biomedical imaging

Among emerging medical imaging techniques we are particularly interested in those using low to moderate frequency regimes. These include Microwave Tomography, Electrical Impedance Tomography and also the closely related Optical Tomography technique. They all have the advantage of being potentially safe and relatively cheap modalities and can also be used in complementarity with well established techniques such as X-ray computed tomography or Magnetic Resonance Imaging.

With these modalities tissues are differentiated and, consequentially can be imaged, based on differences in dielectric properties (some recent studies have proved that dielectric properties of biological tissues can be a strong indicator of the tissues functional and pathological conditions, for instance, tissue blood content, ischemia, infarction, hypoxia, malignancies, edema and others). The main challenge for these functionalities is to built a 3-D imaging algorithm capable of treating multi-static measurements to provide real-time images with highest (reasonably) expected resolutions and in a sufficiently robust way.

Another important biomedical application is brain imaging. We are for instance interested in the use of EEG and MEG techniques as complementary tools to MRI. They are applied for instance to localize epileptic centers or active zones (functional imaging). Here the problem is different and consists into performing passive imaging: the epileptic centers act as electrical sources and imaging is performed from measurements of induced currents. Incorporating the structure of the skull is primordial in improving the resolution of the imaging procedure. Doing this in a reasonably quick manner is still an active research area, and the use of asymptotic models would offer a promising solution to fix this issue.

4.3. Non destructive testing and parameter identification

One challenging problem in this vast area is the identification and imaging of defaults in anisotropic media. For instance this problem is of great importance in aeronautic constructions due to the growing use of composite materials. It also arises in applications linked with the evaluation of wood quality, like locating knots in timber in order to optimize timber-cutting in sawmills, or evaluating wood integrity before cutting trees. The anisotropy of the propagative media renders the analysis of diffracted waves more complex since one cannot only relies on the use of backscattered waves. Another difficulty comes from the fact that the micro-structure of the media is generally not well known a priori.

Our concern will be focused on the determination of qualitative information on the size of defaults and their physical properties rather than a complete imaging which for anisotropic media is in general impossible. For instance, in the case of homogeneous background, one can link the size of the inclusion and the index of refraction to the first eigenvalue of so-called interior transmission problem. These eigenvalues can be determined form the measured data and a rough localization of the default. Our goal is to extend this kind of idea to the cases where both the propagative media and the inclusion are anisotropic. The generalization to the case of cracks or screens has also to be investigated.

In the context of nuclear waste management many studies are conducted on the possibility of storing waste in a deep geological clay layer. To assess the reliability of such a storage without leakage it is necessary to have a precise knowledge of the porous media parameters (porosity, tortuosity, permeability, etc.). The large range of space and time scales involved in this process requires a high degree of precision as well as tight bounds on the uncertainties. Many physical experiments are conducted in situ which are designed for providing data for parameters identification. For example, the determination of the damaged zone (caused by excavation) around the repository area is of paramount importance since microcracks yield drastic changes in the permeability. Level set methods are a tool of choice for characterizing this damaged zone.

4.4. Diffusion MRI

In biological tissues, water is abundant and magnetic resonance imaging (MRI) exploits the magnetic property of the nucleus of the water proton. The imaging contrast (the variations in the grayscale in an image) in standard MRI can be from either proton density, T1 (spin-lattice) relaxation, or T2 (spin-spin) relaxation and the contrast in the image gives some information on the physiological properties of the biological tissue at different physical locations of the sample. The resolution of MRI is on the order of millimeters: the greyscale value shown in the imaging pixel represents the volume-averaged value taken over all the physical locations contained that pixel.

In diffusion MRI, the image contrast comes from a measure of the average distance the water molecules have moved (diffused) during a certain amount of time. The Pulsed Gradient Spin Echo (PGSE) sequence is a commonly used sequence of applied magnetic fields to encode the diffusion of water protons. The term 'pulsed' means that the magnetic fields are short in duration, an the term gradient means that the magnetic fields vary linearly in space along a particular direction. First, the water protons in tissue are labelled with nuclear spin at a precession frequency that varies as a function of the physical positions of the water molecules via the application of a pulsed (short in duration, lasting on the order of ten milliseconds) magnetic field. Because the precessing frequencies of the water molecules vary, the signal, which measures the aggregate phase of the water molecules, will be reduced due to phase cancellations. Some time (usually tens of

milliseconds) after the first pulsed magnetic field, another pulsed magnetic field is applied to reverse the spins of the water molecules. The time between the applications of two pulsed magnetic fields is called the 'diffusion time'. If the water molecules have not moved during the diffusion time, the phase dispersion will be reversed, hence the signal loss will also be reversed, the signal is called refocused. However, if the molecules have moved during the diffusion time, the refocusing will be incomplete and the signal detected by the MRI scanner if weaker than if the water molecules have not moved. This lack of complete refocusing is called the signal attenuation and is the basis of the image contrast in DMRI. the pixels showning more signal attenuation is associated with further water displacement during the diffusion time, which may be linked to physiological factors, such as higher cell membrane permeability, larger cell sizes, higher extra-cellular volume fraction.

We model the nuclear magnetization of water protons in a sample due to diffusion-encoding magnetic fields by a multiple compartment Bloch-Torrey partial differential equation, which is a diffusive-type time-dependent PDE. The DMRI signal is the integral of the solution of the Bloch-Torrey PDE. In a homogeneous medium, the intrinsic diffusion coeffcient D will appear as the slope of the semi-log plot of the signal (in appropriate units). However, because during typical scanning times, 50-100ms, water molecules have had time to travel a diffusion distance which is long compared to the average size of the cells, the slope of the semi-log plot of the signal is in fact a measure of an 'effective' diffusion coefficient. In DMRI applications, this measured quantity is called the 'apparent diffusion coefficient' (ADC) and provides the most commonly used form the image contrast for DMRI. This ADC is closely related to the effective diffusion coefficient obtainable from mathematical homogenization theory.

5. Highlights of the Year

5.1. Highlights of the Year

- L. Audibert obtained the PhD prize Paul CASEAU of EDF.
- Grégoire Allaire was appointed as president of the scientifique board of IFP Energies Nouvelles.
- Grégoire Allaire broke 15 bones in a climbing accident on the 19th of July, 2016. It takes a long time to fully recover...

6. New Software and Platforms

6.1. FVforBlochTorrey

FUNCTIONAL DESCRIPTION

We developed two numerical codes to solve the multiple-compartments Bloch-Torrey partial differential equation in 2D and 3D to simulate the water proton magnetization of a sample under the influence of diffusion-encoding magnetic field gradient pulses.

We coupled the spatial discretization with an efficient time discretization adapted to diffusive problems called the (explicit) Runge-Kutta-Chebyshev method.

The version of the code using Finite Volume discretization on a Cartesian grid is complete (written by Jing-Rebecca Li). The version of the code using linear Finite Elements discretization is complete (written by Dang Van Nguyen and Jing-Rebecca Li).

- Contact: Jing Rebecca Li
- URL: http://www.cmap.polytechnique.fr/~jingrebeccali/

6.2. InvGIBC

A FreeFem++ routines for solving inverse Maxwell's problem for 3D shape identification using a gradient descent method.

- Contact: Houssem Haddar
- URL: http://www.cmap.polytechnique.fr/~haddar/

6.3. RODIN

FUNCTIONAL DESCRIPTION

In the framework of the RODIN project we continue to develop with our software partner ESI the codes Topolev and Geolev for topology and geometry shape optimization of mechanical structures using the level set method.

- Contact: Grégoire Allaire
- URL: http://www.cmap.polytechnique.fr/~allaire/

6.4. Samplings-3d

FUNCTIONAL DESCRIPTION

This software is written in Fortran 90 and is related to forward and inverse problems for the Helmholtz equation in 3-D. It contains equivalent functionalities to samplings-2d in a 3-D setting.

- Contact: Houssem Haddar
- URL: http://www.cmap.polytechnique.fr/~haddar/

6.5. samplings-2d

This software solves forward and inverse problems for the Helmholtz equation in 2-D.
FUNCTIONAL DESCRIPTION

This software is written in Fortran 90 and is related to forward and inverse problems for the Helmholtz equation in 2-D. It includes three independent components. The first one solves to scattering problem using integral equation approach and supports piecewise-constant dielectrics and obstacles with impedance boundary conditions. The second one contains various samplings methods to solve the inverse scattering problem (LSM, RGLSM(s), Factorization, MuSiC) for near-field or far-field setting. The third component is a set of post processing functionalities to visualize the results

- Contact: Houssem Haddar
- URL: http://sourceforge.net/projects/samplings-2d/

6.6. SAXS-EM

This software solves the inverse problem of determining nono-paticles size distrubutions from SAXS measurements.
FUNCTIONAL DESCRIPTION

This software is written in matlab and determine size distributions is isotropic samples from measurements of X-ray diffraction at small angles. It treats the case of diluted and dense particle distributions.

- Contact: Houssem Haddar, Zixian Jiang
- URL: http://www.cmap.polytechnique.fr/~haddar/

7. New Results

7.1. Methods for inverse problems

7.1.1. *Identifying defects in an unknown background using differential measurements*

L. Audibert and H. Haddar

In the framework of the PhD thesis of Lorenzo Audibert we studied non destructive testing of concrete using ultrasonic waves, and more generallly imaging in complex heterogeneous media. We assume that measurements are multistatic, which means that we record the scattered field on different points by using several sources. For this type of data we wish to build methods that are able to image the obstacle that created the scattered field. We use qualitative methods in this work, which only provide the support of the object independently from its physical property. The first part of this thesis consists of a theoretical analysis of the Linear Sampling Method. Such analysis is done in the framework of regularization theory, and our main contribution is to provide and analyze a regularization term that ensures good theoretical properties. Among those properties we were able to demonstrate that when the regularization parameter goes to zero, we actually construct a sequence of functions that strongly converges to the solution of the interior transmission problem. This behavior gives a central place to the interior transmission problem as it allows describing the asymptotic solution of our regularized problem. Using this characterization of our solution, we are able to give the optimal reconstruction we can get from our method. More importantly this description of the solution allows us to compare the solution coming from two different datasets. Based on the result of this comparison, we manage to produce an image of the connected component that contains the defect which appears between two measurement campaigns and this regardless of the medium. This method is well suited for the characteristics of the microstructure of concrete as shown on several numerical examples with realistic concrete-like microstructure. Finally, we extend our theoretical results to the case of limited aperture, anisotropic medium and elastic waves, which correspond to the real physics of the ultrasounds

7.1.2. *Generalized linear sampling method for elastic-wave sensing of heterogeneous fractures*

B. Guzina, H. Haddar and F. Pourahmadian

A theoretical foundation is developed for active seismic reconstruction of fractures endowed with spatially-varying interfacial condition (e.g. partially-closed fractures, hydraulic fractures). The proposed indicator functional carries a superior localization property with no significant sensitivity to the fracture's contact condition, measurement errors, and illumination frequency. This is accomplished through the paradigm of the F-factorization technique and the recently developed Generalized Linear Sampling Method (GLSM) applied to elastodynamics. The direct scattering problem is formulated in the frequency domain where the fracture surface is illuminated by a set of incident plane waves, while monitoring the induced scattered field in the form of (elastic) far-field patterns. The analysis of the well-posedness of the forward problem leads to an admissibility condition on the fracture's (linearized) contact parameters. This in turn contributes toward establishing the applicability of the F-factorization method, and consequently aids the formulation of a convex GLSM cost functional whose minimizer can be computed without iterations. Such minimizer is then used to construct a robust fracture indicator function, whose performance is illustrated through a set of numerical experiments. For completeness, the results of the GLSM reconstruction are compared to those obtained by the classical linear sampling method.

7.1.3. *Invisibility in scattering theory*

L. Chesnel, A.-S. Bonnet-Ben Dhia and S.A. Nazarov

We are interested in a time harmonic acoustic problem in a waveguide with locally perturbed sound hard walls. We consider a setting where an observer generates incident plane waves at $-\infty$ and probes the resulting scattered field at $-\infty$ and $+\infty$. Practically, this is equivalent to measure the reflection and transmission coefficients respectively denoted R and T. In a recent work, a technique has been proposed to construct waveguides with smooth walls such that $R = 0$ and $|T| = 1$ (non reflection). However the approach fails to

ensure $T = 1$ (perfect transmission without phase shift). First we establish a result explaining this observation. More precisely, we prove that for wavenumbers smaller than a given bound $k_{\star}$ depending on the geometry, we cannot have $T = 1$ so that the observer can detect the presence of the defect if he/she is able to measure the phase at $+\infty$. In particular, if the perturbation is smooth and small (in amplitude and in width), $k_{\star}$ is very close to the threshold wavenumber. Then, in a second step, we change the point of view and, for a given wavenumber, working with singular perturbations of the domain, we show how to obtain $T = 1$. In this case, the scattered field is exponentially decaying both at $-\infty$ and $+\infty$. We implement numerically the method to provide examples of such undetectable defects.

7.1.4. *Nanoparticles volume determination from SAXS measurements*

H. Haddar and Z. Jiang

The aim of this work is to develop a fully automatic method for the reconstruction of the volume distribution of polydisperse non-interacting nanoparticles with identical shapes from Small Angle X-ray Scattering measurements. In the case of diluted systems we proposed a method that solves a maximum likelihood problem with a positivity constraint on the solution by means of an Expectation Maximization iterative scheme coupled with a robust stopping criterion. We prove that the stopping rule provides a regularization method according to an innovative notion of regularization specifically defined for inverse problems with Poisson data. Such a regularization, together with the positivity constraint results in high fidelity quantitative reconstructions of particle volume distributions making the method particularly effective in real applications. We tested the performance of the method on synthetic data in the case of uni- and bi-modal particle volume distributions. We extended the method to the case of dense solutions where the inverse problem becomes non linear. A specific fix-point algorithm has been proposed and convergence has been tested against synthetic data. The developement of this research topic is ongoing under the framework of Saxsize.

7.1.5. *Identifying defects in unknown periodic layers*

H. Haddar and T.P. Nguyen

We investigate the inverse problem where one is interested in reconstructing the support of a perturbation in a periodic media from measurements of scattered waves. We are concerned with the design of a sampling method that would reconstruct the support of inhomogeneities without reconstructing the index of refraction. The development of sampling methods has gained a large interest in recent years and many methods have been introduced in the literature to deal with a variety of problems and we refer to [1] for an account of recent developments of these methods. Up to our knowledge, the sampling methods for locally perturbed infinite periodic layers has not been treated in the literature. Even thought this problem is the one that motivates our study, we considered a slightly different problem that will be referred to as the ML-periodic problem: it corresponds with a locally perturbed infinite periodic layer with period L that has been reduced to a domain of size ML (with M a sufficiently large parameter) with periodic boundary conditions. This is mainly for technical reasons since our analysis for the newly introduced differential imaging functional heavily rely on the discrete Floquet-Bloch transform.

The main contribution of our work is the design of a new sampling method that enable the imaging of the defect location without reconstructing the L periodic background. This method is in the spirit of the Differential LSM introduced above for the imaging of defects in complex backgrounds using differential measurements. However, in the present case we propose a method that does not require the measurement operator for the background media. We exploit the L periodicity of the background and the Floquet-Bloch transform to design a differential criterion between different periods. This criterion is based on the study of sampling methods for the ML-periodic media where a single Floquet-Bloch mode is used. This study constitutes the main theoretical ingredient for our method. The sampling operator for a single Floquet-Bloch mode somehow plays the role of the measurement operator for the background media. Indeed the main interest for this new sampling method is that it is capable of identifying the defect even thought classical sampling methods fail in obtaining high fidelity reconstructions of the (complex) background media.

7.1.6. *Identification of small objects with near-field data in quasi-backscattering configurations*

H. Haddar and M. Lakhal

We present a new sampling method for detecting targets (small inclusions or defects) immersed in a homogeneous medium in three-dimensional space, from measurements of acoustic scattered fields created by point source incident waves. We consider the harmonic regime and a data setting that corresponds with quasi-backscattering configuration: the data is collected by a set a receivers that are distributed on a segment centered at the source position and the device is swept along a path orthogonal to the receiver line. We assume that the aperture of the receivers is small compared with the distance to the targets. Considering the asymptotic form of the scattered field as the size of the targets goes to zero and the small aperture approximation, one is able to derive a special expression for the scattered field. In this expression a separation of the dependence of scattered field on the source location and the distance source-target is performed. This allows us to propose a sampling procedure that characterizes the targets location in terms of the range of a near-field operator constructed from available data. Our procedure is similar to the one proposed by Haddar-Rezac for far-field configurations. The reconstruction algorithm is based on the MUSIC (Multiple SIgnal Classification) algorithm.

7.1.7. *Nondestructive testing of the delaminated interface between two materials*

F. Cakoni, I. De Teresa, H. Haddar and P. Monk

We consider the problem of detecting if two materials that should be in contact have separated or delaminated. The goal is to find an acoustic technique to detect the delamination. We model the delamination as a thin opening between two materials of different acoustic properties, and using asymptotic techniques we derive a asymptotic model where the delaminated region is replaced by jump conditions on the acoustic field and flux. The asymptotic model has potential singularities due to the edges of the delaminated region, and we show that the forward problem is well posed for a large class of possible delaminations. We then design a special Linear Sampling Method (LSM) for detecting the shape of the delamination assuming that the background, undamaged, state is known. Finally we show, by numerical experiments, that our LSM can indeed determine the shape of delaminated regions.

7.2. Shape and topology optimization

7.2.1. *Second-order shape derivatives along normal trajectories, governed by Hamilton-Jacobi equations*

G. Allaire, E. Cancès and J.-L. Vié

In this work we introduce a new variant of shape differentiation which is adapted to the deformation of shapes along their normal direction. This is typically the case in the level-set method for shape optimization where the shape evolves with a normal velocity. As all other variants of the orginal Hadamard method of shape differentiation, our approach yields the same first order derivative. However, the Hessian or second-order derivative is different and somehow simpler since only normal movements are allowed. The applications of this new Hessian formula are twofold. First, it leads to a novel extension method for the normal velocity, used in the Hamilton-Jacobi equation of front propagation. Second, as could be expected, it is at the basis of a Newton optimization algorithm which is conceptually simpler since no tangential displacements have to be considered. Numerical examples are given to illustrate the potentiality of these two applications. The key technical tool for our approach is the method of bicharacteristics for solving Hamilton-Jacobi equations. Our new idea is to differentiate the shape along these bicharacteristics (a system of two ordinary differential equations).

7.2.2. *Introducing a level-set based shape and topology optimization method for the wear of composite materials with geometric constraints*

G. Allaire, F. Feppon, G. Michailidis, M.S. Sidebottom, B.A. Krick and N. Vermaak

The wear of materials continues to be a limiting factor in the lifetime and performance of mechanical systems with sliding surfaces. As the demand for low wear materials grows so does the need for models and methods to systematically optimize tribological systems. Elastic foundation models offer a simplified framework to study the wear of multimaterial composites subject to abrasive sliding. Previously, the evolving wear profile has been shown to converge to a steady-state that is characterized by a time-independent elliptic equation. In this article, the steady-state formulation is generalized and integrated with shape optimization to improve the wear performance of bi-material composites. Both macroscopic structures and periodic material microstructures are considered. Several common tribological objectives for systems undergoing wear are identified and mathematically formalized with shape derivatives. These include (i) achieving a planar wear surface from multimaterial composites and (ii) minimizing the run-in volume of material lost before steady-state wear is achieved. A level-set based topology optimization algorithm that incorporates a novel constraint on the level-set function is presented. In particular, a new scheme is developed to update material interfaces; the scheme (i) conveniently enforces volume constraints at each iteration, (ii) controls the complexity of design features using perimeter penalization, and (iii) nucleates holes or inclusions with the topological gradient. The broad applicability of the proposed formulation for problems beyond wear is discussed, especially for problems where convenient control of the complexity of geometric features is desired.

7.2.3. *Geometric constraints for shape and topology optimization in architectural design*

G. Allaire, C. Dapogny, A. Faure, G. Michailidis, A. Couvelas and R. Estevez

This work proposes a shape and topology optimization framework oriented towards conceptual architectural design. A particular emphasis is put on the possibility for the user to interfere on the optimization process by supplying information about his personal taste. More precisely, we formulate three novel constraints on the geometry of shapes; while the first two are mainly related to aesthetics, the third one may also be used to handle several fabrication issues that are of special interest in the device of civil structures. The common mathematical ingredient to all three models is the signed distance function to a domain, and its sensitivity analysis with respect to perturbations of this domain; in the present work, this material is extended to the case where the ambient space is equipped with an anisotropic metric tensor. Numerical examples are discussed in two and three space dimensions.

7.2.4. *Modal basis approaches in shape and topology optimization of frequency response problems*

G. Allaire and G. Michailidis

The optimal design of mechanical structures subject to periodic excitations within a large frequency interval is quite challenging. In order to avoid bad performances for non-discretized frequencies, it is necessary to finely discretize the frequency interval, leading to a very large number of state equations. Then, if a standard adjoint-based approach is used for optimization, the computational cost (both in terms of CPU and memory storage) may be prohibitive for large problems, especially in three space dimensions. The goal of the present work is to introduce two new non-adjoint approaches for dealing with frequency response problems in shape and topology optimization. In both cases, we rely on a classical modal basis approach to compute the states, solutions of the direct problems. In the first method, we do not use any adjoint but rather directly compute the shape derivatives of the eigenmodes in the modal basis. In the second method, we compute the adjoints of the standard approach by using again the modal basis. The numerical cost of these two new strategies are much smaller than the usual ones if the number of modes in the modal basis is much smaller than the number of discretized excitation frequencies. We present numerical examples for the minimization of the dynamic compliance in two and three space dimensions.

7.3. Direct scattering problems

7.3.1. *Finite element methods for eigenvalue problems with sign-changing coefficients*

C. Carvalho, P. Ciarlet and L. Chesnel

We consider a class of eigenvalue problems involving coefficients changing sign on the domain of interest. We analyse the main spectral properties of these problems according to the features of the coefficients. Under some assumptions on the mesh, we study how one can use classical finite element methods to approximate the spectrum as well as the eigenfunctions while avoiding spurious modes. We also prove localisation results of the eigenfunctions for certain sets of coefficients.

7.3.2. A Volume integral method for solving scattering problems from locally perturbed periodic layers

H. Haddar and T.P. Nguyen

We investigate the scattering problem for the case of locally perturbed periodic layers in R^d, $d = 2, 3$. Using the Floquet-Bloch transform in the periodicity direction we reformulate this scattering problem as an equivalent system of coupled volume integral equations. We then apply a spectral method to discretize the obtained system after periodization in the direction orthogonal to the periodicity directions of the medium. The convergence of this method is established and validating numerical results are provided.

7.4. Asymptotic Analysis

7.4.1. Small obstacle asymptotics for a non linear problem

L. Chesnel, X. Claeys and S.A. Nazarov

We study a 2D semi-linear equation in a domain with a small Dirichlet obstacle of size δ. Using the method of matched asymptotic expansions, we compute an asymptotic expansion of the solution as δ tends to zero. Its relevance is justified by proving a rigorous error estimate. We also construct an approximate model, based on an equation set in the limit domain without the small obstacle, which provides a good approximation of the far field of the solution of the original problem. The interest of this approximate model lies in the fact that it leads to a variational formulation which is very simple to discretize. We present numerical experiments to illustrate the analysis.

7.4.2. Influence of the geometry on plasmonic waves

L. Chesnel X. Claeys and S.A. Nazarov

In the modeling of plasmonic technologies in time harmonic regime, one is led to study the eigenvalue problem $- \div (\sigma \nabla u) = \lambda u$ (P), where σ is a physical coefficient positive in some region Ω_+ and negative in some other region Ω_-. We highlight an unusual instability phenomenon for the source term problem associated with (P): for certain configurations, when the interface between Ω_+ and Ω_- presents a rounded corner, the solution may depend critically on the value of the rounding parameter. We explain this property studying the eigenvalue problem (P). We provide an asymptotic expansion of the eigenvalues and prove error estimates. We establish an oscillatory behaviour of the eigenvalues as the rounding parameter of the corner tends to zero. These theoretical results are illustrated by numerical experiments.

7.4.3. Instability of dielectrics and conductors in electrostatic fields

G. Allaire and J. Rauch

This work proves most of the assertions in section 116 of Maxwell's treatise on electromagnetism. The results go under the name Earnshaw's Theorem and assert the absence of stable equilibrium configurations of conductors and dielectrics in an external electrostatic field.

7.4.4. Optimization of dispersive coefficients in the homogenization of the wave equation in periodic structures

G. Allaire and T. Yamada

We study dispersive effects of wave propagation in periodic media, which can be modelled by adding a fourth-order term in the homogenized equation. The corresponding fourth-order dispersive tensor is called Burnett tensor and we numerically optimize its values in order to minimize or maximize dispersion. More precisely, we consider the case of a two-phase composite medium with an 8-fold symmetry assumption of the periodicity cell in two space dimensions. We obtain upper and lower bound for the dispersive properties, along with optimal microgeometries.

7.4.5. *Homogenization of Stokes System using Bloch Waves*

G. Allaire, T. Ghosh and M. Vanninathan

In this work, we study the Bloch wave homogenization for the Stokes system with periodic viscosity coefficient. In particular, we obtain the spectral interpretation of the homogenized tensor. The presence of the incompressibility constraint in the model raises new issues linking the homogenized tensor and the Bloch spectral data. The main difficulty is a lack of smoothness for the bottom of the Bloch spectrum, a phenomenon which is not present in the case of the elasticity system. This issue is solved in the present work, completing the homogenization process of the Stokes system via the Bloch wave method.

7.5. Diffusion MRI

7.5.1. *Adapting the Kärger model to account for finite diffusion-encoding pulses in diffusion MRI*

H. Haddar, J.R. Li and S. Schiavi

Diffusion magnetic resonance imaging (dMRI) is an imaging modality that probes the diffusion characteristics of a sample via the application of magnetic field gradient pulses. If the imaging voxel can be divided into different Gaussian diffusion compartments with inter-compartment exchange governed by linear kinetics, then the dMRI signal can be described by the Kärger model, which is a well-known model in NMR. However, the Kärger model is limited to the case when the duration of the diffusion-encoding gradient pulses is short compared to the time delay between the start of the pulses. Under this assumption, the time at which to evaluate the Kärger model to obtain the dMRI signal is unambiguously the delay between the pulses. Recently, a new model of the dMRI signal, the Finite-Pulse Kärger (FPK) model, was derived for arbitrary diffusion gradient profiles. Relying on the FPK model, we show that when the duration of the gradient pulses is not short, the time at which to evaluate the Kärger model should be the time delay between the start of the pulses, shortened by one third of the pulse duration. With this choice, we show the sixth order convergence of the Kärger model to the FPK model in the non-dimensionalized pulse duration.

7.5.2. *A macroscopic model for the diffusion MRI signal accounting for time-dependent diffusivity*

H. Haddar, J.R. Li and S. Schiavi

An important quantity measured in dMRI in each voxel is the Apparent Diffusion Coefficient (ADC) and it is well-established from imaging experiments that, in the brain, *in-vivo*, the ADC is dependent on the measured diffusion time. To aid in the understanding and interpretation of the ADC, using homogenization techniques, we derived a new asymptotic model for the dMRI signal from the Bloch-Torrey equation governing the water proton magnetization under the influence of diffusion-encoding magnetic gradient pulses. Our new model was obtained using a particular choice of scaling for the time, the biological cell membrane permeability, the diffusion-encoding magnetic field gradient strength, and a periodicity length of the cellular geometry. The ADC of the resulting model is dependent on the diffusion time. We numerically validated this model for a wide range of diffusion times for two dimensional geometrical configurations.

7.5.3. *Quantitative DLA-based Compressed Sensing for MEMRI Acquisitions*

P. Svehla, K.-V. Nguyen, J.-R. Li and L. Ciobanu

High resolution Manganese Enhanced Magnetic Resonance Imaging (MEMRI) has great potential for functional imaging of live neuronal tissue at single neuron scale. However, reaching high resolutions often requires long acquisition times which can lead to reduced image quality due to sample deterioration and hardware instability. Compressed Sensing (CS) techniques offer the opportunity to significantly reduce the imaging time. The purpose of this work is to test the feasibility of CS acquisitions based on Diffusion Limited Aggregation (DLA) sampling patterns for high resolution quantitative MEMRI imaging. Fully encoded and DLA-CS MEMRI images of Aplysia californica neural tissue were acquired on a 17.2T MRI system. The MR signal corresponding to single, identified neurons was quantified for both versions of the T1 weighted images. Results: For a 50% undersampling, DLA-CS leads to signal intensity differences, measured in individual neurons, of approximately 1.37% when compared to the fully encoded acquisition, with minimal impact on image spatial resolution. At the undersampling ratio of 50%, DLA-CS is capable of accurately quantifying signal intensities in MEMRI acquisitions. Depending on the image signal to noise ratio, higher undersampling ratios can be used to further reduce the acquisition time in MEMRI based functional studies of living tissues.

7.5.4. *The time-dependent diffusivity in the abdominal ganglion of Aplysia californica, comparing experiments and simulations*

K.-V. Nguyen, D. Le Bihan, L. Ciobanu and J.-R. Li

The nerve cells of the *Aplysia* are much larger than mammalian neurons. Using the *Aplysia* ganglia to study the relationship between the cellular structure and the diffusion MRI signal can shed light on this relationship for more complex organisms. We measured the dMRI signal at several diffusion times in the abdominal ganglion and performed simulations of water diffusion in geometries obtained after segmenting high resolution T2-weighted images and incorporating known information about the cellular structure from the literature. By fitting the experimental signal to the simulated signal for several types of cells in the abdominal ganglion at a wide range of diffusion times, we obtained estimates of the intrinsic diffusion coefficient in the nucleus and the cytoplasm. We also evaluated the reliability of using an existing formula for the time-dependent diffusion coefficient to estimate cell size.

7.5.5. *A two pool model to describe the IVIM cerebral perfusion*

G. Fournet, J.-R. Li, A.M. Cerjanic, B.P. Sutton, L. Ciobanu and D. Le Bihan

IntraVoxel Incoherent Motion (IVIM) is a magnetic resonance imaging (MRI) technique capable of measuring perfusion-related parameters. In this manuscript, we show that the mono-exponential model commonly used to process IVIM data might be challenged, especially at short diffusion times. Eleven rat datasets were acquired at 7T using a diffusion-weighted pulsed gradient spin echo sequence with b-values ranging from 7 to 2500 s/mm^2 at 3 diffusion times. The IVIM signals, obtained by removing the diffusion component from the raw MR signal, were fitted to the standard mono-exponential model, a bi-exponential model and the Kennan model. The Akaike information criterion used to find the best model to fit the data demonstrates that, at short diffusion times, the bi-exponential IVIM model is most appropriate. The results obtained by comparing the experimental data to a dictionary of numerical simulations of the IVIM signal in microvascular networks support the hypothesis that such a bi-exponential behavior can be explained by considering the contribution of two vascular pools: capillaries and somewhat larger vessels.

7.5.6. *The influence of acquisition parameters on the metrics of the bi-exponential IVIM model*

G. Fournet, J.-R. Li, D. Le Bihan and L. Ciobanu

The IntraVoxel Incoherent Motion (IVIM) MRI signal, typically described as a mono-exponential decay, can sometimes be better modeled as a bi-exponential function accounting for two vascular pools, capillaries and medium-size vessels. The goal of this work is to define precisely in which conditions the IVIM signal shape becomes bi-exponential and to understand the evolution of the IVIM outputs with different acquisition parameters. Rats were scanned at 7T and 11.7T using diffusion-weighted pulsed-gradient spin-echo (SE) and stimulated-echo (STE) sequences with different repetition times (TR) and diffusion encoding times. The obtained IVIM signals were fit to the mono- and bi-exponential models and the output parameters compared.

The bi-exponential and mono-exponential models converge at long diffusion encoding times and long TRs. The STE is less sensitive to inflow effects present at short TRs, leading to a smaller volume fraction for the fast pool when compared to the SE sequence. The two vascular components are more easily separated at short diffusion encoding times, short TRs and when using a SE sequence. The volume fractions of the two blood pools depend on the pulse sequence, TR and diffusion encoding times while the pseudo-diffusion coefficients are only affected by the diffusion encoding time.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Grant with ART-FI (June 2016- June 2017) on quantification of electromagnetic radiations inside the brain from partial measurements

- A CIFRE PhD thesis started in January 2015 with Dassault Aviations. The student is M. Aloïs Bissuel who is working on "linearized Navier-Stokes equations for optimization, fluttering and aeroacoustic".

- A CIFRE PhD thesis started in December 2015 with Safran Tech. The student is Mrs Perle Geoffroy who is working on "topology optimization by the homogenization method in the context of additive manufacturing".

8.2. Bilateral Grants with Industry

- The SOFIA project (SOlutions pour la Fabrication Industrielle Additive métallique) started in the summer of 2016. Its purpose is to make research in the field of metallic additive manufacturing. The industrial partners include Michelin, FMAS, ESI, Safran and others. The academic partners are different laboratories of CNRS, including CMAP at Ecole Polytechnique. The project is funded for 6 years by BPI (Banque Publique d'Investissement).

- FUI project Tandem. This three years project started in December 2012 and has been extended to September 2017 involves Bull-Amesys (coordinator), BOWEN (ERTE+SART), Ecole Polytechnique (CMAP), Inria, LEAT et VSM. It aims at constructing a radar system on a flying device capable of real-time imaging mines embedded in dry soils (up to 40 cm deep). We are in charge of numerical validation of the inverse simulator.

- FUI project Saxsize. This three years project started in October 2015 and involves Xenocs (coordinator), Inria (DEFI), Pyxalis, LNE, Cordouan and CEA. It is a followup of Nanolytix where a focus is put on SAXS quantifications of dense nanoparticle solutions.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

- ANR Metamath: Modelization and numerical simulation of wave propagation in metamaterials, program MN, September 2011- November 2016. This is a joint ANR with POEMS, Inria Scalay Ile de France project team (Coordinator, S. Fliss), DMIA, Département de Mathématiques de l'ISAE and IMATH, Laboratoire de Mathématiques de l'Université de Toulon. https://www.rocq.inria.fr/poems/metamath

- ANR CIACM "Computational Imaging of the Aging Cerebral Microvasculature", funded by ANR Program "US-French Collaboration". French Partners (Coordinating partner CEA Neurospin): CEA Neurospin (Coordinator Luisa Ciobanu), Inria Saclay (Coordinator Jing-Rebecca Li). US Partner: Univ of Illinois, bioengineering department (Coordinator Brad Sutton). Duration: Sept 2013-Sept 2016.

9.2. International Initiatives

9.2.1. Inria International Partners

9.2.1.1. Declared Inria International Partners

QUASI

Title: Qualitative Approaches to Scattering and Imaging

International Partner (Institution - Laboratory - Researcher):

University of Rutgers (United States) - Fioralba Cakoni

Duration: 2013 - 2017

Start year: 2013

We concentrate on the use of qualitative methods in acoustic and electromagnetic inverse scattering theory with applications to nondestructive evaluation of materials and medical imaging. In particular, we would like to address theoretical and numerical reconstruction techniques to solve the inverse scattering problems using either time harmonic or time dependent measurements of the scattered field. The main goal of research in this field is to not only detect but also identify geometric and physical properties of unknown objects in real time.

9.3. International Research Visitors

9.3.1. Visits of International Scientists

- Fioralba Cakoni (4 months)
- David Colton (1 week)
- Semra Ahmetola (11 months)
- Armin Lechleiter (1 week)
- Bojan Guzina (1 week)
- Helle Majander (12 months)

9.3.1.1. Internships

- Irene De Teresa-Trueba (University of Delaware) 3 months
- Jacob Rezac (University of Delaware) 3 months
- Marwa Kchaou (ENIT) 3 months
- BumsuKim (Ecole Polytechnique), from Jul 2016 until Nov 2016
- KevishNapal (Inria), from May 2016 until Oct 2016
- Hoang An Tran (Ecole Polytechnique), from Apr 2016 until Jun 2016

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- G. Allaire is scientific chair and one of the main organizers of the CEA/GAMNI seminar on computational fluid mechanics, IHP Paris (January 2016).

10.1.1.2. Member of the Organizing Committees

- L. Chesnel co-organized with Xavier Claeys and Sonia Fliss the workshop "Waves in periodic media and metamaterials" in Cargese http://uma.ensta-paristech.fr/conf/metamath/Metamath/Workshop.html
- L. Chesnel co-organize the seminar of the Centre de Mathématiques Appliquées of École Polytechnique.
- G. Allaire co-organized the PGMO conference (8-9 November 2016)
- J.R. Li is member of Organizing Committee of SIAM Conference on Computational Science and Engineering, 2017
- J.R. Li is organizer of Ecole d'ete d'excellence for Chinese Master's students funded by French Embassy in China, 2017.
- H. Haddar Co-organized the "International Conference on Computational Mathematics and Inverse Problems", Michgan, 15-19 August 2016

10.1.2. Scientific Events Selection

10.1.2.1. Member of the Conference Program Committees

- J.R. Li is member of the SIAM Committee on Programs and Conferences 2017-2019
- H. Haddar is memeber of the scientific committee of the conference series TAMTAM and Waves

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- G. Allaire is member of the editorial board of
    - book series "Mathématiques et Applications" of SMAI and Springer,
    - ESAIM/COCV, Structural and Multidisciplinary Optimization,
    - Discrete and Continuous Dynamical Systems Series B,
    - Computational and Applied Mathematics,
    - Mathematical Models and Methods in Applied Sciences (M3AS),
    - Annali dell'Universita di Ferrara,
    - OGST (Oil and Gas Science and Technology),
    - Journal de l'Ecole Polytechnique - Mathématiques,
    - Journal of Optimization Theory and Applications.
- H. Haddar is
    - member the editorial advisory board of Inverse Problems
    - Associate Editor of the SIAM Journal on Scientific Computing

10.1.3.2. Reviewer - Reviewing Activities

The members of the team reviewed numerous papers for numerous international journals. Too many to make a list.

10.1.4. Invited Talks

- G. Allaire
  - Congrès LEM2I à Hammamet, Tunisie (avril 2016).
  - Workshop "Variational Models of Fracture" à Calgary, Canada (mai 2016).
  - Congrès ECCOMAS à Heraklion, Crête (juin 2016).
  - European forum on additive manufacturing, Chatenay Malabry (juin 2016).
- H. Haddar
  - "ATAVI International Conference on Acoustics and Vibration ICAV'2016" from 21 to 23 March 2016 in Hammamet - Tunisia
  - Journées des Mathématiciens Tunisiens à l'Etranger 20-21 juillet 2016 Cité des Sciences de Tunis
  - Oberwolfach Workshop Theory and Numerics of Inverse Scattering Problems, 18 Sep - 24 Sep 2016.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Licence : Grégoire Allaire, Approximation Numérique et Optimisation, for students in the second year of Ecole Polytechnique curriculum: 8 lessons of 1h30.
- Licence : Houssem Haddar, Approximation Numérique et Optimisation, for students in the second year of Ecole Polytechnique curriculum: 8 TDs of 4h.
- Licence : Houssem Haddar, Variational analysis of partial differential equations, for students in the second year of Ecole Polytechnique curriculum: 8 TDs of 4h.
- Licence: Lucas Chesnel, "Elementary tools of analysis for partial differential equations", 29 equivalent TD hours, L3, Ensta ParisTech, Palaiseau, France
- Master : Grégoire Allaire, Optimal design of structures, for students in the third year of Ecole Polytechnique curriculum. 9 lessons of 1h30.
- Master : Grégoire Allaire, Transport et Diffusion, for students in the third year of Ecole Polytechnique curriculum. With F. Golse, 1/2 of 9 lessons of 1h30.
- Master : Houssem Haddar, Inverse problems, for Master (M2) students of Ecole Polytechnique and Paris 6 University, 1/2 of 9 lessons of 2h.
- Master: Lucas Chesnel, "The finite element method", 6 equivalent TD hours, M1, Ensta ParisTech, Palaiseau, France

10.2.2. Supervision

- Ph.D. : G. Fournet, Inclusion of blood flow in micro-vessels in a new dMRI signal model, October 2016, J.-R. Li and L. Ciobanu
- Ph.D.: S. Schiavi, Homogenized models for Diffusion MRI, December 2016, H. Haddar and J.-R. Li
- PhD : M. Giacomini, Shape optimization and Applications to aeronautics, December 2016, O. Pantz and K. Trabelsi
- PhD : A. Maury, shape optimization for non-linear structures, December 2016, G. Allaire and F. Jouve
- PhD : C. Patricot, coupling algorithms in neutronic/thermal-hydraulic/mechanics for numerical simulation of nuclear reactors, March 2016, G. Allaire and E. Hourcade

- PhD in progress : J.-L. Vié, optimization algorithms for topology design of structures, December 2016, G. Allaire and E. Cancès
- Ph.D. in progress: M. Lakhal, Time domain inverse scattering for buried objects, 2014, H. Haddar
- Ph.D. in progress: T.P. Nguyen, Direct and Inverse scattering from locally perturbed layers, 2013, H. Haddar
- Ph.D. in progress: B. Charfi, Identification of the sigular support of a GIBC, 2014, H. Haddar and S. Chaabane
- Ph.D. in progress: K. Van Nguyen, Modeling, simulation and experimental verification of water diffusion in neuronal network of the Aplysia ganglia, 2014, J.-R. Li and L. Ciobanu
- PhD in progress : A. Talpaert, the direct numerical simulation of vapor bubbles at low Mach number with adaptative mesh refinement, 2013, G. Allaire and S. Dellacherie
- PhD in progress : A. Bissuel, linearized Navier Stokes equations for optimization, floating and aeroaccoustic, 2014, G. Allaire
- PhD in progress :P. Geoffroy on topology optimization by the homogenization method in the context of additive manufacturing (Safran Tech, to be defended in 2019), G. Allaire.
- PhD in progress : K. Napal, Transmission eigenvalues and non destructive testing of concrete like materials , 2016, L. Chesnel H. Haddar and L. Audibert
- PhD in progress : M. Kchaou, Higher order homogenization tensors for DMRI modeling, 2016, H. Haddar, J.R Li and M. Moakher

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] M. GIACOMINI.*Quantitative a posteriori error estimators in Finite Element-based shape optimization*, Ecole polytechnique ; Université Paris-Saclay, December 2016, https://tel.archives-ouvertes.fr/tel-01418841.

[2] S. SCHIAVI.*Homogenized and analytical models for the diffusion MRI signal*, Université Paris Saclay ; École Polytechnique X, December 2016, https://tel.archives-ouvertes.fr/tel-01422174.

Articles in International Peer-Reviewed Journal

[3] G. ALLAIRE, O. BERNARD, J.-F. DUFRECHE, A. MIKELIC.*Ion transport through deformable porous media:derivation of the macroscopic equations using upscaling*, in "Computational and Applied Mathematics", February 2016 [*DOI :* 10.1007/s40314-016-0321-0], https://hal.archives-ouvertes.fr/hal-01278241.

[4] G. ALLAIRE, M. BRIANE, M. VANNINATHAN.*A comparison between two-scale asymptotic expansions and Bloch wave expansions for the homogenization of periodic structures*, in "SeMA Journal", 2016, vol. 73, n° 3, p. 237-259 [*DOI :* 10.1007/s40324-016-0067-z], https://hal.archives-ouvertes.fr/hal-01215580.

[5] L. AUDIBERT, H. HADDAR.*The Generalized Linear Sampling Method for limited aperture measurements*, in "SIAM Journal on Imaging Sciences", 2016, https://hal.archives-ouvertes.fr/hal-01422027.

[6] F. BENVENUTO, H. HADDAR, L. BLANDINE.*A robust inversion method according to a new notion of regularization for Poisson data with an application to nanoparticle volume determination*, in "SIAM Journal on Applied Mathematics", February 2016, vol. 76, n° 1, p. 276–292 [*DOI :* 10.1137/15M1024354], https://hal.inria.fr/hal-01217540.

[7] A.-S. BONNET-BENDHIA, C. CARVALHO, L. CHESNEL, P. CIARLET.*On the use of Perfectly Matched Layers at corners for scattering problems with sign-changing coefficients*, in "Journal of Computational Physics", October 2016, vol. 322, p. 224-247 [*DOI :* 10.1016/J.JCP.2016.06.037], https://hal.archives-ouvertes.fr/hal-01225309.

[8] F. CAKONI, I. DE TERESA, H. HADDAR, P. MONK.*Nondestructive testing of the delaminated interface between two materials*, in "SIAM Journal on Applied Mathematics", 2016, https://hal.inria.fr/hal-01373251.

[9] F. CAUBET, H. HADDAR, J.-R. LI, D. V. NGUYEN.*New Transmission Condition Accounting For Diffusion Anisotropy In Thin Layers Applied To Diffusion MRI*, in "Modelisation Mathématique et Analyse Numérique", 2016 [*DOI :* 10.1051/M2AN/2016060], https://hal.inria.fr/hal-01110298.

[10] S. CHAABANE, B. CHARFI, H. HADDAR.*Reconstruction of discontinuous parameters in a second order impedance boundary operator*, in "Inverse Problems", July 2016, vol. 32, n$^{\text{o}}$ 10 [*DOI :* 10.1088/0266-5611/32/10/105004], https://hal.inria.fr/hal-01349696.

[11] L. CHESNEL, X. CLAEYS.*A numerical approach for the Poisson equation in a planar domain with a small inclusion*, in "BIT Numerical Mathematics", March 2016, https://hal.inria.fr/hal-01109552.

[12] G. DOLLÉ, O. DURAN, N. FEYEUX, E. FRÉNOD, M. GIACOMINI, C. PRUD'HOMME.*Mathematical modeling and numerical simulation of a bioreactor landfill using Feel++*, in "ESAIM: Proceedings and Surveys", 2016, https://hal.archives-ouvertes.fr/hal-01258643.

[13] G. FOURNET, J.-R. LI, A. M. CERJANIC, B. P. SUTTON, L. CIOBANU, D. LE BIHAN.*A two-pool model to describe the IVIM cerebral perfusion*, in "Journal of Cerebral Blood Flow and Metabolism", 2016 [*DOI :* 10.1177/0271678X16681310], https://hal.archives-ouvertes.fr/hal-01429440.

[14] M. GIACOMINI, O. PANTZ, K. TRABELSI.*Certified Descent Algorithm for shape optimization driven by fully-computable a posteriori error estimators*, in "ESAIM: Control, Optimisation and Calculus of Variations", 2016 [*DOI :* 10.1051/COCV/2016021], https://hal.archives-ouvertes.fr/hal-01201914.

[15] H. HADDAR, Z. JIANG, M.-K. RIAHI.*A robust inversion method for quantitative 3D shape reconstruction from coaxial eddy-current measurements*, in "Journal of Scientific Computing", July 2016, 31 [*DOI :* 10.1007/S10915-016-0241-6], https://hal.inria.fr/hal-01110299.

[16] H. HADDAR, J.-R. LI, S. SCHIAVI.*A new macroscopic model for the diffusion MRI accounting for time-dependent diffusivity*, in "SIAM Journal on Applied Mathematics", May 2016, vol. 76, n$^{\text{o}}$ 3, p. 930–949 [*DOI :* 10.1137/15M1019398], https://hal.inria.fr/hal-01217537.

[17] H. HADDAR, J.-R. LI, S. SCHIAVI.*Adapting the Kärger model to account for finite diffusion-encoding pulses in diffusion MRI*, in "IMA Journal of Applied Mathematics", May 2016 [*DOI :* 10.1093/IMAMAT/HXW032], https://hal.inria.fr/hal-01217560.

[18] H. HADDAR, T. P. NGUYEN.*A volume integral method for solving scattering problems from locally perturbed infinite periodic layers*, in "Applicable Analysis", 2016, 29 [*DOI :* 10.1080/00036811.2016.1221942], https://hal.inria.fr/hal-01374892.

[19] F. POURAHMADIAN, B. B. GUZINA, H. HADDAR.*Generalized linear sampling method for elastic-wave sensing of heterogeneous fractures*, in "Inverse Problems", 2016, 28, https://hal.inria.fr/hal-01421916.

Scientific Books (or Scientific Book chapters)

[20] F. CAKONI, D. COLTON, H. HADDAR.*Inverse Scattering Theory and Transmission Eigenvalues*, CBMS Series, SIAM publications, 2016, vol. 88, https://hal.inria.fr/hal-01374893.

Other Publications

[21] G. ALLAIRE, E. CANCÈS, J.-L. VIE.*Second-order shape derivatives along normal trajectories, governed by Hamilton-Jacobi equations*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01326805.

[22] G. ALLAIRE, G. MICHAILIDIS.*Modal basis approaches in shape and topology optimization of frequency response problems*, August 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01354162.

[23] G. ALLAIRE, A. PIATNITSKI.*On the asymptotic behaviour of the kernel of an adjoint convection-diffusion operator in a long cylinder*, January 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01258747.

[24] G. ALLAIRE, J. RAUCH.*Instability of dielectrics and conductors in electrostatic fields*, May 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01321566.

[25] L. AUDIBERT.*Sampling method for sign changing contrast*, December 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01422024.

[26] A.-S. BONNET-BENDHIA, L. CHESNEL, S. NAZAROV.*Perfect transmission invisibility for waveguides with sound hard walls*, September 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01371163.

[27] C. CARVALHO, L. CHESNEL, P. CIARLET.*Eigenvalue problems with sign-changing coefficients*, November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01394856.

[28] L. A. CHESNEL, X. CLAEYS, S. A. NAZAROV.*Small obstacle asymptotics for a 2D semi-linear convex problem*, January 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01427617.

[29] L. CHESNEL, S. NAZAROV.*Team organization may help swarms of flies to become invisible in closed waveguides*, August 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01216135.

[30] F. FEPPON, G. MICHAILIDIS, M. SIDEBOTTOM, G. ALLAIRE, B. KRICK, N. VERMAAK.*Introducing a level-set based shape and topology optimization method for the wear of composite materials with geometric constraints*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01336301.

[31] G. FOURNET, J.-R. LI, D. LE BIHAN, L. CIOBANU.*The influence of acquisition parameters on the metrics of the bi-exponential IVIM model*, 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01429508.

[32] M. GIACOMINI.*An equilibrated fluxes approach to the Certified Descent Algorithm for shape optimization using conforming Finite Element and Discontinuous Galerkin discretizations*, November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01395529.

[33] M. GIACOMINI, O. PANTZ, K. TRABELSI.*Volumetric expressions of the shape gradient of the compliance in structural shape optimization*, January 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01441943.

[34] H. HADDAR, J.-R. LI, S. SCHIAVI.*Understanding the Time-Dependent Effective Diffusion Coefficient Measured by Diffusion MRI: the Intra-Cellular Case*, 2016, working paper or preprint, https://hal.inria.fr/hal-01421928.

[35] A. MAURY, G. ALLAIRE, F. JOUVE.*Shape optimisation with the level set method for contact problems in linearised elasticity*, January 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01435325.

[36] F. POURAHMADIAN, B. B. GUZINA, H. HADDAR.*A synoptic approach to the seismic sensing of heterogeneous fractures: from geometric reconstruction to interfacial characterization*, December 2016, working paper or preprint, https://hal.inria.fr/hal-01422085.

[37] P. SVEHLA, K.-V. NGUYEN, J.-R. LI, L. CIOBANU.*Quantitative DLA-based Compressed Sensing for MEMRI Acquisitions*, 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01429506.

Project-Team DISCO

Dynamical Interconnected Systems in COmplex Environments

IN COLLABORATION WITH: Laboratoire des signaux et systèmes (L2S)

IN PARTNERSHIP WITH:
CNRS

CentraleSupélec

RESEARCH CENTER
Saclay - Île-de-France

THEME
Optimization and control of dynamic systems

Table of contents

Project-Team DISCO

Creation of the Team: 2010 January 01, updated into Project-Team: 2012 January 01

Keywords:

Computer Science and Digital Science:
3.4.4. - Optimization and learning
3.4.5. - Bayesian methods
6.1.1. - Continuous Modeling (PDE, ODE)
6.1.3. - Discrete Modeling (multi-agent, people centered)
6.4.1. - Deterministic control
6.4.3. - Observability and Controlability
6.4.4. - Stability and Stabilization

Other Research Topics and Application Domains:
2.2.3. - Cancer
2.3. - Epidemiology
3.6. - Ecology
4.3.3. - Wind energy
5.2.3. - Aviation

The team is located at L2S, CentraleSupelec.

1. Members

Research Scientists
Catherine Bonnet [Team leader, Inria, Senior Researcher, HDR]
Frédéric Mazenc [Inria, Researcher, HDR]
Silviu-Iulian Niculescu [CNRS, Senior Researcher, HDR]

Faculty Members
Sorin Olaru [CentraleSupelec, Professor, HDR]
Guillaume Sandou [CentraleSupelec, Professor, HDR]
Ali Zemouche [En délégation de l'univ. Lorraine, Associate Professor, from Sep 2016, HDR]

PhD Students
Caetano Cardeliquio [CentraleSupelec, from Jun 2016]
Walid Djema [Inria]
Dina Irofti [CNRS, from Mar 2016]

Post-Doctoral Fellows
Le Ha Vy Nguyen [Inria]
Hakki Unal [Inria, until Aug 2016]

Visiting Scientists
André Fioravanti [Unicamp, since Nov 2016]
Stefanella Boatto [Univ Rio de Janeiro, since Oct 2016]
Yutaka Yamamoto [Univ Kyoto, Sep-Nov 2016]

Administrative Assistants
Katia Evrat [Inria, since Dec 2016]
Olga Mwana Mobulakani [Inria, until Nov 2016]

2. Overall Objectives

2.1. Objectives

The goal of the project is to better understand and well formalize the effects of complex environments on the dynamics of the interconnections, as well as to develop new methods and techniques for the analysis and control of such systems.

It is well-known that the interconnection of dynamic systems has as consequence an increased complexity of the behavior of the total system.

In a simplified way, as the concept of dynamics is well-understood, the interconnections can be seen as associations (by connections of materials or information flows) of distinct systems to ensure a pooling of the resources with the aim of obtaining a better operation with the constraint of continuity of the service in the event of a fault. In this context, the environment can be seen as a collection of elements, structures or systems, natural or artificial constituting the neighborhood of a given system. The development of interactive games through communication networks, control from distance (e.g. remote surgical operations) or in hostile environment (e.g. robots, drones), as well as the current trend of large scale integration of distribution (and/or transport and/or decision) and open information systems with systems of production, lead to new modeling schemes in problems where the dynamics of the environment have to be taken into account.

In order to tackle the control problems arising in the above examples, the team investigates new theoretical methods, develops new algorithms and implementations dedicated to these techniques.

3. Research Program

3.1. Modeling of complex environment

We want to model phenomena such as a temporary loss of connection (e.g. synchronisation of the movements through haptic interfaces), a nonhomogeneous environment (e.g. case of cryogenic systems) or the presence of the human factor in the control loop (e.g. grid systems) but also problems involved with technological constraints (e.g. range of the sensors). The mathematical models concerned include integro-differential, partial differential equations, algebraic inequalities with the presence of several time scales, whose variables and/or parameters must satisfy certain constraints (for instance, positivity).

3.2. Analysis of interconnected systems

- Robust stability of linear systems

  Within an interconnection context, lots of phenomena are modelled directly or after an approximation by delay systems. These systems might have fixed delays, time-varying delays, distributed delays ...

  For various infinite-dimensional systems, particularly delay and fractional systems, input-output and time-domain methods are jointly developed in the team to characterize stability. This research is developed at four levels: analytic approaches (H_∞-stability, BIBO-stablity, robust stability, robustness metrics) [1], [2], [5], [6], symbolic computation approaches (SOS methods are used for determining easy-to-check conditions which guarantee that the poles of a given linear system are not in the closed right half-plane, certified CAD techniques), numerical approaches (root-loci, continuation methods) and by means of softwares developed in the team [5], [6].

- Robustness/fragility of biological systems

Deterministic biological models describing, for instance, species interactions, are frequently composed of equations with important disturbances and poorly known parameters. To evaluate the impact of the uncertainties, we use the techniques of designing of global strict Lyapunov functions or functional developed in the team.

However, for other biological systems, the notion of robustness may be different and this question is still in its infancy (see, e.g. [66]). Unlike engineering problems where a major issue is to maintain stability in the presence of disturbances, a main issue here is to maintain the system response in the presence of disturbances. For instance, a biological network is required to keep its functioning in case of a failure of one of the nodes in the network. The team, which has a strong expertise in robustness for engineering problems, aims at contributing at the develpment of new robustness metrics in this biological context.

3.3. Stabilization of interconnected systems

- Linear systems: Analytic and algebraic approaches are considered for infinite-dimensional linear systems studied within the input-output framework.

In the recent years, the Youla-Kučera parametrization (which gives the set of all stabilizing controllers of a system in terms of its coprime factorizations) has been the cornerstone of the success of the H_∞-control since this parametrization allows one to rewrite the problem of finding the optimal stabilizing controllers for a certain norm such as H_∞ or H_2 as affine, and thus, convex problem.

A central issue studied in the team is the computation of such factorizations for a given infinite-dimensional linear system as well as establishing the links between stabilizability of a system for a certain norm and the existence of coprime factorizations for this system. These questions are fundamental for robust stabilization problems [1], [2].

We also consider simultaneous stabilization since it plays an important role in the study of reliable stabilization, i.e. in the design of controllers which stabilize a finite family of plants describing a system during normal operating conditions and various failed modes (e.g. loss of sensors or actuators, changes in operating points). Moreover, we investigate strongly stabilizable systems, namely systems which can be stabilized by stable controllers, since they have a good ability to track reference inputs and, in practice, engineers are reluctant to use unstable controllers especially when the system is stable.

- Nonlinear systems

The project aims at developing robust stabilization theory and methods for important classes of nonlinear systems that ensure good controllerperformance under uncertainty and time delays. The main techniques include techniques called backstepping and forwarding, contructions of strict Lyapunov functions through so-called "strictification" approaches [3] and construction of Lyapunov-Krasovskii functionals [4], [5], [6].

- Predictive control

For highly complex systems described in the time-domain and which are submitted to constraints, predictive control seems to be well-adapted. This model based control method (MPC: Model Predictive Control) is founded on the determination of an optimal control sequence over a receding horizon. Due to its formulation in the time-domain, it is an effective tool for handling constraints and uncertainties which can be explicitly taken into account in the synthesis procedure [7]. The team considers how mutiparametric optimization can help to reduce the computational load of this method, allowing its effective use on real world constrained problems.

The team also investigates stochastic optimization methods such as genetic algorithm, particle swarm optimization or ant colony [8] as they can be used to optimize any criterion and constraint whatever their mathematical structure is. The developed methodologies can be used by non specialists.

3.4. Synthesis of reduced complexity controllers

- PID controllers

  Even though the synthesis of control laws of a given complexity is not a new problem, it is still open, even for finite-dimensional linear systems. Our purpose is to search for good families of "simple" (e.g. low order) controllers for infinite-dimensional dynamical systems. Within our approach, PID candidates are first considered in the team [2], [67].

- Predictive control

  The synthesis of predictive control laws is concerned with the solution of multiparametric optimization problems. Reduced order controller constraints can be viewed as non convex constraints in the synthesis procedure. Such constraints can be taken into account with stochastic algorithms.

Finally, the development of algorithms based on both symbolic computation and numerical methods, and their implementations in dedicated Scilab/Matlab/Maple toolboxes are important issues in the project.

4. Application Domains

4.1. Analysis and Control of life sciences systems

The team is involved in life sciences applications. The two main lines are the analysis of bioreactors models and the modeling of cell dynamics in Acute Myeloblastic Leukemias (AML) in collaboration with St Antoine Hospital in Paris. A recent new subject is the modelling of Dengue epidemia.

4.2. Energy Management

The team is interested in Energy management and considers optimization and control problems in energy networks.

5. New Results

5.1. Characterizing the Codimension of Zero Singularities for Time-Delay Systems: A Link with Vandermonde and Birkhoff Incidence Matrices

Participants: Islam Boussaada, Silviu-Iulian Niculescu.

The analysis of time-delay systems mainly relies on detecting and understanding the spectral values bifurcations when crossing the imaginary axis. We have dealt with the zero singularity, essentially when the zero spectral value is multiple. The simplest case in such a configuration is characterized by an algebraic multiplicity two and a geometric multiplicity one, known as the Bogdanov-Takens singularity. Moreover, in some cases the codimension of the zero spectral value exceeds the number of the coupled scalar-differential equations. Nevertheless, to the best of the author's knowledge, the bounds of such a multiplicity have not been deeply investigated in the literature. It is worth mentioning that the knowledge of such an information is crucial for nonlinear analysis purposes since the dimension of the projected state on the center manifold is none other than the sum of the dimensions of the generalized eigenspaces associated with spectral values with zero real parts. Motivated by a control-oriented problems, we have provided an answer to this question for time-delay systems, taking into account the parameters' algebraic constraints that may occur in applications. We emphasize the link between such a problem and the incidence matrices associated with the Birkhoff interpolation problem. In this context, symbolic algorithms for LU-factorization for functional confluent Vandermonde as well as some classes of bivariate functional Birkhoff matrices are also proposed [11].

5.2. Tracking the Algebraic Multiplicity of Crossing Imaginary Roots for Generic Quasipolynomials: A Vandermonde-Based Approach

Participants: Islam Boussaada, Silviu-Iulian Niculescu.

A standard approach in analyzing dynamical systems consists in identifying and understanding the eigenvalues bifurcations when crossing the imaginary axis. Efficient methods for crossing imaginary roots identification exist. However, to the best of the author's knowledge, the multiplicity of such roots was not deeply investigated. We have emphasized [12] that the multiplicity of the zero spectral value can exceed the number of the coupled scalar delay-differential equations and a constructive approach Vandermonde-based allowing to an adaptive bound for such a multiplicity is provided. Namely, it is shown that the zero spectral value multiplicity depends on the system structure (number of delays and number of non zero coefficients of the associated quasipolynomial) rather than the degree of the associated quasipolynomial. We have extended the constructive approach in investigating the multiplicity of crossing imaginary roots $j\omega$ where $\omega \neq 0$ and establishes a link with a new class of functional confluent Vandermonde matrices. A symbolic algorithm for computing the LU-factorization for such matrices is provided. As a byproduct of the proposed approach, a bound sharper than the Polya-Szegö generic bound arising from the principle argument is established.

5.3. Coprimeness of fractional representations

Participants: Catherine Bonnet, Le Ha Vy Nguyen, Yutaka Yamamoto [Kyoto Univ].

Coprimeness of a fractional representation plays various crucial roles in many different contexts, for example, stabilization of a given plant, minimality of a state space representation, etc. It should be noted however that coprimeness depends crucially on the choice of a ring (or algebra) where such a representation is taken, which reflects the choice of a plant, and particular problems that one studies. Such relationships are particularly delicate and interesting when dealing with infinite-dimensional systems. We have disucssed various coprimeness issues for different rings, typically for H_∞ and pseudorational transfer functions. The former is related to H_∞-stabilizability, and the latter to controllability of behaviors. We have also given some intricate examples where a seemingly non-coprime factorization indeed turns out to be a coprime factorization over H_∞ [28], [29].

5.4. Output-feddback control design for time-delay systems

Participants: Catherine Bonnet, Caetano Cardeliquio, Matheus Souza [FEEC-UNICAMP], André Fioravanti [FEM-UNICAMP].

We presented new results on H_∞-control synthesis, via output-feedback, for time-delay linear systems. We extend the use of a finite order LTI system, called comparison system, to design a controller which depends not only on the output at the present time and maximum delay, but also on an arbitrary number of values between those. This approach allows us to increase the maximum stable delay without requiring any additional information.

5.5. Backstepping control design through the introduction of delays

Participants: Frederic Mazenc, Michael Malisoff [LSU], Jerome Weston [LSU].

We provided new backstepping results for time-varying systems with input delays. The results were obtained by the introduction of constant 'artificial' pointwise delays in the input. Thus they are significantly different from backstepping results for systems with delay in the input as presented in previous contributions.

I) The novelty of the contribution in [18] is in the bounds on the controls, and the facts that (i) one does not need to compute any Lie derivatives to apply the proposed controls, (ii) the controls have no distributed terms, and (iii) no differentiability conditions on the available controls for the subsystems are needed. The latter aspect is of paramount importance from an applied point of view.

II) In [43], we extended [18]. We provided new globally stabilizing backstepping controls for single input systems in a partially linear form. Instead of measuring the full state, the feedbacks use current and several time lagged values of a function of the state of the nonlinear subsystem (and have no distributed terms). We also allowed input delays. This improves on [18], since we allowed an arbitrary number of integrators whereas [18] is limited to one integrator.

5.6. Switched Nonlinear Systems

Participants: Frederic Mazenc, Yue-E Wang [Shaanxi Normal University], Xi-Ming Sun [Dalian University of Technology].

We considered in [26] a class of nonlinear time-varying switched control systems for which stabilizing feedbacks are available. We analyzed the effect of the presence of a delay in the input of switched nonlinear systems with an external disturbance. By contrast with most of the contributions available in the literature, we did not assume that all the subsystems of the switched system we consider are stable when the delay is present. Through a Lyapunov approach, we derived sufficient conditions in terms of size of the delay, ensuring the global exponential stability of the switched system. Moreover, under appropriate conditions, the input-to-state stability of the system with respect to an external disturbance was established.

5.7. Studies of systems with long delays

Participants: Frederic Mazenc, Michael Malisoff [LSU], Emilia Fridman [Tel-Aviv University].

We solved several problems of observer and control designs pertaining to the fundamental (and difficult) case where a delay in the input is too long for being neglected.

I) We considered in [17] the problem of stabilizing a linear continuous-time system with discrete-time measurements and a sampled input with a pointwise constant delay. In a first part, we designed a continuous-discrete observer which converges when the maximum time interval between two consecutive measurements is sufficiently small. In a second part, we constructed a dynamic output feedback by using a technique which is strongly reminiscent of the so called 'reduction model approach'. It stabilizes the system when the maximal time between two consecutive sampling instants is sufficiently small. No limitation on the size of the delay was imposed and an ISS property with respect to additive disturbances was established.

II) We solved stabilization problems for linear time-varying systems under input delays. We showed how changes of coordinates lead to systems with time invariant drifts, which are covered by the reduction model method and which lead to the problem of stabilizing a time-varying system without delay. For continuous-time periodic systems, we used Floquet's theory to find the changes of coordinates. We also proved an analogue for discrete time systems, through an original discrete-time extension of Floquet's theory [19].

III) In [21] and [42], we proposed a prediction based stabilization approach for a general class of nonlinear time-varying systems with pointwise delay in the input. It is based on a recent new prediction strategy, which makes it possible to circumvent the problem of constructing and estimating distributed terms in the expression for the stabilizing control laws. We observed that our result applies in cases where other recent results do not, including notably the case where a time-varying delay is present.

5.8. Extension of the Razumikhin's theorem

Participants: Frederic Mazenc, Michael Malisoff [LSU].

The Razumikhin's Theorem is a major extension of the Lypunov function theory, making possible to establish the global asymptotic stability of nonlinear systems with delays. It is especially efficient when the delays are time-varying. We provide in [41] an extension of this theorem for continuous-time time-varying systems with time-varying delays. Our result uses a novel 'strictification' technique for converting a nonstrict Lyapunov function into a strict one. Our examples show how our method can sometimes allow broader classes of allowable delays than the results in the literature.

5.9. Observer design for nonlinear systems

Participant: Ali Zemouche.

A new high-gain observer design method with lower gain compared to the standard high-gain observer was proposed. This new observer, called "HG/LMI" observer is obtained by combining the standard high-gain methodology with the LPVLMI-based technique. Through analytical developments, it is shown how the new observer provides a lower gain. A numerical example was used to illustrate the performance of the new "HG/LMI" observer. The aim of this research is the application of this new observer design to estimate some vehicle variables in autonomous vehicle applications.

5.10. Set invariance for discrete-time delay systems

Participants: Sorin Olaru, Mohammed Laraba [L2S], Silviu Niculescu, Franco Blanchini [Univ. Udine, Italy], Stefano Miani [Univ. Udine, Italy].

The existence of positively invariant sets for linear delay-difference equations was pursued in [15]. We made a survey effort and presented in a unified framework all known necessary and/or sufficient conditions for the existence of invariant sets with respect to dynamical systems described by linear discrete time-delay difference equations (dDDEs). Secondly, we address the construction of invariant sets in the original state space (also called D-invariant sets) by exploiting the forward mappings. The notion of D-invariance is appealing since it provides a region of attraction, which is difficult to obtain for delay systems without taking into account the delayed states in some appropriate extended state space model. The paper contains a sufficient condition for the existence of ellipsoidal D-contractive sets for dDDEs, and a necessary and sufficient condition for the existence of D-invariant sets in relation to linear time-varying dDDE stability. Another contribution is the clarification of the relationship between convexity (convex hull operation) and D-invariance of linear dDDEs. In short, it is shown that the convex hull of the union of two or more D-invariant sets is not necessarily D-invariant, while the convex hull of a non-convex D-invariant set is D-invariant. Positive invariance is an essential concept in control theory, with applications to constrained dynam-ical systems analysis, uncertainty handling as well as related control design problems. It serves as a basic tool in many topics, such as model predictive control, fault tolerant control and reference governor design. Furthermore, there exists a close link between classical stability theory and positive invariant sets. It is worth mentioning that, in Lyapunov theory, invariance is implicitly described.

5.11. Interpolation-based design for constrained dynamical systems

Participants: Sorin Olaru, Nam Nguyen [IFP, France], Per Olof Gutman [Technion, Israel].

A technique is presented in [49] leading to an explicit state feedback solution to the regulation problem for uncertain and/or time-varying linear discrete-time systems with state and control constraints. A piecewise affine control law is provided which not only guarantees recursive feasibility and robust asymptotic stability, but is also optimal for a region of the state space containing the origin.

5.12. Inverse optimality results for constrained control

Participants: Sorin Olaru, Ngoc Anh Nguyen [L2S], Pedro Rodriguez [L2S], Morten Hovd [NTNU Trondheim, Norway], Ioan Necoara [Univ. Politehnica Bucharest, Romania].

Parametric convex programming has received a lot of attention, since it has many applications in chemical engineering, control engineering, signal processing , etc. Further, inverse optimality plays an important role in many contexts, e.g., image processing, motion planning. In this context we introduced [10] a constructive solution of the inverse optimality problem for the class of continuous piecewise affine functions. The main idea is based on the convex lifting concept. Accordingly, an algorithm to construct convex liftings of a given convexly liftable partition have been put forward. Continuous piecewise affine function defined over a polytopic partition of the state space are known to be obtained as the solution of a parametric linear/quadratic programming problem. Regarding linear model predictive control, is shown that any continuous piecewise affine control law can be obtained via a linear optimal control problem with the control horizon at most equal to 2 prediction steps.

5.13. Robustness and fragility of Piecewise affine control laws

Participants: Sorin Olaru, Ngoc Anh Nguyen [L2S], Pedro Rodriguez [L2S], Morten Hovd [NTNU Trondheim, Norway], Georges Bitsoris [Univ. Patras, Greece].

We focus in [9]on the robustness and fragility problem for piecewise affine (PWA) control laws for discrete-time linear system dynamics in the presence of parametric uncertainty of the state space model. A generic geometrical approach will be used to obtain robustness/fragility margins with respect to the positive invariance properties. For PWA control laws defined over a bounded region in the state space, it is shown that these margins can be described in terms of polyhedral sets in parameter space. The methodology is further extended to the fragility problem with respect to the partition defining the controller. Finally, several computational aspects are presented regarding the transformation from the theoretical formulations to explicit representations (vertex/halfspace representation of polytopes) of these sets.

5.14. Distributed robust model predictive control

Participants: Sorin Olaru, Alexandra Grancharova [Technical University of Sofia, Bulgaria].

We presented in a suboptimal approach to distributed closed-loop min-max MPC for uncertain systems consisting of polytopic subsystems with coupled dynamics subject to both state and input constraints. The approach applies the dynamic dual decomposition method and reformulates the original centralized min-max MPC problem into a distributed optimization problem. The suggested approach was illustrated on a simulation example of an uncertain system consisting of two interconnected polytopic subsystems.

5.15. Algebraic Analysis Approach to Linear Functional Systems

Participants: Guillaume Sandou, Mohamed Lotfi Derouiche [Ecole nationale d'Ingénieurs de Tunis], Soufiene Bouallegue [Ecole nationale d'Ingénieurs de Tunis], Joseph Haggège Derouiche [Ecole nationale d'Ingénieurs de Tunis].

In this study, a new Model Predictive Controller (MPC) parameters tuning strategy is proposed using a LabVIEW-based perturbed Particle Swarm Optimization (pPSO). This original LabVIEW implementation of this metaheuristic algorithm is firstly validated on some test functions in order to show its efficiency and validity. The optimization results are compared with the standard PSO approach. The parameters tuning problem, i.e. the weighting factors on the output error and input increments of the MPC algorithm, is then formulated and systematically solved, using the proposed LabVIEW pPSA algorithm. The case of a Magnetic Levitation (MAGLEV) system is investigated to illustrate the robustness and superiority of the proposed pPSO-based tuning MPC approach. All obtained simulation results, as well as the statistical analysis tests for the formulated control problem with and without constraints, are discussed and compared with the Genetic Algorithm Optimization (GAO)-based technique in order to improve the effectiveness of the proposed pPSA-based MPC tuning methodology derouiche:hal-01347041.

5.16. Attitude dynamics, control and observation

Participants: Frederic Mazenc, Maruthi Akella [Univ of Texas], Sunpil Yang [Univ of Texas].

In [27], we addressed the rigid-body attitude tracking problem in the absence of angular velocity measurements. To achieve proportional-derivative feedback control, an angular velocity observer with global exponential convergence was designed based on the Immersion and Invariance (I&I) method. A dynamic scaling factor was introduced to circumvent the integrability condition typically arising in I&I design. Unlike the existing I&I observer for this problem, the estimated angular velocity is defined using a rotation matrix of the current quaternion state to avoid use of an additional filter for the angular velocity estimate. As a result, stability analysis became less complex and the observer structure was further simplified by efficient handling of the Coriolis effect in the observer error dynamics. In the case where proportional-derivative control is combined with the observer, asymptotic convergence of tracking errors was proved while establishing a separation property. Numerical simulations were provided to demonstrate the performance of the proposed observer and the output feedback controller.

5.17. Estimation for vehicle application

Participants: Ali Zemouche, Rajesh Rajamani [University of Minneapolis, USA], Gridsada Phanomchoeng [Chulalongkorn University, Thailand].

A new LMI (Linear Matrix Inequality) design technique is developed to address the problem of circle criterion based $\mathcal{H}_\infty$ observer design for nonlinear systems. The developed technique applies to both locally Lipschitz as well as monotonic nonlinear systems, and allows for nonlinear functions in both the process dynamics and output equations. The LMI design condition obtained is less conservative than all previous results proposed in literature for these classes of nonlinear systems. By judicious use of a modified Young's relation, additional degrees of freedom are included in the observer design. These additional decision variables enable improvements in the feasibility of the obtained LMI. Several recent results in literature are shown to be particular cases of the more general observer design methodology developed in this paper. Illustrative examples are used to show the effectiveness of the proposed methodology. The application of the method to slip angle estimation in automotive applications is discussed and experimental results are presented. This application was the main motivation of this work.

5.18. Observer-based stabilization for lateral vehicle control

Participants: Ali Zemouche, Rajesh Rajamani [University of Minneapolis, USA], Yan Wang [University of Minneapolis, USA].

Recently, motivated by autonomous vehicle control problem, a robust observer based estimated state feedback control design method for an uncertain dynamical system that can be represented as a LTI system connected with an IQC-type nonlinear uncertainty was developed. Different from existing design methodologies in which a convex semidefinite constraint is obtained at the cost of conservatism and unrealistic assumptions, the design of the robust observer state feedback controller is formulated in this paper as a feasibility problem of a bilinear matrix inequality (BMI) constraint. Unfortunately, the search for a feasible solution of a BMI constraint is a NP hard problem in general. The applicability of the linearization method, such as variable change method or congruence transformation, depends on the specific structure of the problem at hand and cannot be generalized. A new sequential LMI optimization method to search for a feasible solution was established. A vehicle lateral control problem was presented to demonstrate the applicability of the proposed algorithm to a real-world estimated state feedback control design.

5.19. Unified model for low-cost high-performance AC drives: the equivalent flux concept

Participants: Guillaume Sandou, Mohamad Koteich [Renault], Abdelmalek Maloum [Renault], Gilles Duc [CentraleSupélec].

This study presents a unified modeling approach of alternating current (AC) machines for low-cost high-performance drives. The Equivalent Flux concept is introduced. Using this concept, all AC machines can be seen as a non-salient synchronous machine with modified (equivalent) rotor flux. Therefore, complex salient-rotor machines models are simplified, and unified shaft-sensorless AC drives can be sought. For this purpose, a unified observer-based structure for rotor-flux position and speed estimation is proposed. The equivalent flux concept generalizes the existing concepts, such as the extended back-electromotive force, the fictitious flux and the active flux.

5.20. Supervision and rescheduling of a mixed CBTC traffic on a suburban railway line

Participants: Guillaume Sandou, Juliette Pochet [SNCF], Sylvain Baro [SNCF].

Railway companies need to achieve higher capacities on existing infrastructures such as high density suburban mainlines. Communication based train control (CBTC) systems have been widely deployed on dedicated

subway lines. However, deployment on shared rail infrastructure, where CBTC and non-CBTC trains run, leads to a mixed positioning and controlling system with different precision levels and restrictions. New performance and complexity issues are to arise. In this study, a method for rescheduling adapted to a CBTC system running in a mixed traffic, is introduced. The proposed method is based on a model predictive control (MPC) approach. In each step, an enhanced genetic algorithm with new mutation mechanisms solves the problem to optimize the cost function. It determines the dwell times and running times of CBTC trains, taking into account the non-CBTC trains planning and fixed-block localization. In addition, reordering can be allowed by modifying the problem constraints. The work is supported by a simulation tool developed by SNCF and adapted to mixed traffic study. The approach is illustrated with a case study based on a part of an East/West line in the Paris region network, proving the ability of the method to find good feasible solutions when delays occur in traffic [46].

5.21. Combined Feedback Linearization and MPC for Wind Turbine Power Tracking

Participants: Guillaume Sandou, Nicolo Gionfra [CentraleSupélec], Houria Siguerdidjane [Centrale-Supélec], Damien Faille [EDF], Philippe Loevenbruck [CentraleSupélec].

The problem of controlling a variable-speedvariable-pitch wind turbine in non conventional operating points is addressed.We aim to provide a control architecture for a general active power tracking problem for the entire operating envelope. The presented control enables to cope with system non linearities while handling state and input constraints, and avoiding singular points. Simulations are carried out based on a 600 kW turbine parameters. Montecarlo simulation shows that the proposed controller presents a certain degree of robustness with respect to the system major uncertainties [36].

5.22. Hierarchical Control of a Wind Farm for Wake Interaction Minimization

Participants: Guillaume Sandou, Nicolo Gionfra [CentraleSupélec], Houria Siguerdidjane [Centrale-Supélec], Damien Faille [EDF], Philippe Loevenbruck [CentraleSupélec].

The problem of controlling a wind farm for power optimization by minimizing the wake interaction among wind turbines is addressed. We aim to evaluate the real gain in farm power production when the dynamics of the controlled turbines are taken into account. The proposed local control enables the turbines to track the required power references in the whole operating envelope, and under the major uncertainties of the system. Simulations are carried out based on a wind farm of 600 kW turbines and they show the actual benefit of considering the wake effect in the optimization algorithm [54].

5.23. Control of a model of chemostat with delay

Participants: Frederic Mazenc, Michael Malisoff [LSU], Jerome Harmand [INRA].

We provided in [39] a new control design for models of chemostats, under constant substrate input concentrations, using piecewise constant delayed measurements of the substrate concentration. The growth functions can be uncertain and are not necessarily monotone. The dilution rate is the control. We used a new Lyapunov approach to derive conditions on the largest sampling interval and on the delay length to ensure asymptotic stabilization properties of a componentwise positive equilibrium point.

5.24. Mathematical Modelling of Acute Myeloid Leukemia

Participants: Catherine Bonnet, Jean Clairambault [MAMBA project-team], François Delhommeau [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Walid Djema, Emilia Fridman [Tel-Aviv University], Pierre Hirsch [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Frédéric Mazenc.

ALMA project focuses on analysis of healthy and unhealthy blood cell production. Dynamics of cell populations are modeled and mathematically analyzed in order to explain why some pathological disorders

may occur. The challenging problem that we are facing is to steadily extend both modelling and analysis aspects to constantly better represent this complex physiological mechanism, which is not yet fully understood. This year, we have progressed on this line [35] and particular emphasis has been placed on a new generation of differential systems, coupled to algebraic equations, modeling abnormal proliferation as observed in acute myeloid leukemia [65]. We have developed, in [34], Lyapunov-like techniques in order to derive global or local exponential stability conditions for that class of differential-difference hematopoietic models. A new model describing the coexistence between ordinary and mutated hematopoietic stem cells was introduced and analyzed in [33]. Above all, this was about giving theoretical conditions to guarantee the survival of healthy cells while eradicating unhealthy ones. Interpretation of mathematical results leads us to provide possibly innovative therapies by combining drugs infusions. By continuing on the path of models coupling healthy and malignant cells, we proposed a framework to investigate the phenomena of tumour dormancy, which goes beyond leukemias, to cover all types of cancer. Finally, in a recent study, we highlighted the role played by growth factors -hormone-like molecules- on the regulation of various biological features involved in hematopoietic mechanisms; that we interpret in the framework of switching systems with distributed delays.

5.25. Ananlysis of Dengue Fever SIR Model with time-varying parameters

Participants: Stefanella Boatto [Univ Feder Rio de Janeiro], Catherine Bonnet, Frédéric Mazenc.

Dengue fever is an infectious viral disease occuring in humans that is prevalent in parts of Central and South America, Africa, India and South-east Asia and which causes 390 millions of infections worldwilde. We have considered here a SIR model of Dengue fever with a periodically time-dependent infection rate. Such a model has been considered by other authors before but we focused here on different aspects such as the existence of a periodic stable orbit and the importance of the phase of the infection rates.

6. Bilateral Contracts and Grants with Industry

6.1. Bilateral Contracts with Industry

A collaboration with SAGEM Défense Sécurité on the stabilization of the lines of sight for pointing systems from optronic criterion using Bayesian optimization ended in December 2016 (CIFRE).

A collaboration with Renault on the observability study of AC machines ended in May 2016 (CIFRE).

A collaboration with SNCF on the supervision and rescheduling of a mixed CBTC traffic on a suburban railway line is currently undergoing (CIFRE).

A collaboration with EDF on the control of renewable energy parks is undergoing (financial support of a PhD student).

A collaboration with CEA and ADEME on the modelling anc control of district heating networks is undergoing (financial support of a PhD student).

7. Partnerships and Cooperations

7.1. Regional Initiatives

- DIGITEO Project (DIM LSC) ALMA3

  Project title: Mathematical Analysis of Acute Myeloid Leukemia (AML) and its treatments

  September 2014 - August 2017

  Coordinator: Catherine Bonnet

  Other partners: Inria Paris-Rocquencourt, France, L2S, France, UPMC, St Antoine Hospital Paris

Abstract: this project follows the regional projects ALMA (2010-2014) and ALMA2 (2011-2013). Starting from the work of J. L. Avila Alonso's PhD thesis in ALMA the aim of this project is to provide a refined coupled model of healthy and cancer cell dynamics in AML whose (stability) analysis will enable evaluation of polychemiotherapies delivered in the case of AML which have a high level of Flt-3 duplication (Flt-3-ITD).

7.2. National Initiatives

7.2.1. Industrial-Academic Institute

Guillaume Sandou is the head of the RISEGrid Institute. The Institute is dedicated to the study, modelling and simulation of smart electric distribution grids and their interactions with the whole electric power system. It is located in Supélec and gathers about 20 people (academic and industrial researchers, PhD students, post-doctoral researchers).

7.3. European Initiatives

7.3.1. FP7 & H2020 Projects

Program: ITN

Project acronym: TEMPO

Project title: Training in Embedded Predictive Control and Optimization

Duration: January 2014 - January 2018

Coordinator: Tor Arne Johanson; with Sorin Olaru (as French PI)

Other partners: U. Frieburg, Oxford, Imperial College; NTNU Trondheim; STUBA Bratislava; EPFL Lausanne; KU Leuven, Renault, ABB, Ampyx Power

Abstract: TEMPO is an international PhD program for highly motivated young scientists, where state-of-the-art research is combined with a comprehensive training program. The network is funded by the European Communityâs Seventh Framework program. The European Commission wants to make research careers more attractive to young people and therefore offers early-stage researchers (ESRs) a PhD program the opportunity to improve their research skills, join established research teams and enhance their career prospects via the Marie Curie Initial Training Networks (ITN) in the area of Embedded Predictive Control and Optimization.

Program: IEF

Project acronym: FUTURISM

Project title: Multiple sensor FaUlt ToleRant control for management of Interconnected nonlinear SysteMs

Duration: May 2014 - April 2016

Coordinator: Sorin Olaru

Abstract: The primary research objective of this project is the design and analysis of novel methods for diagnosing multiple sensor faults and compensating their effects on multi-sensory schemes used for controlling interconnected, nonlinear systems. The second main objective of this project is the application of these methods to complex systems.

7.3.2. Collaborations in European Programs, Except FP7 & H2020

Program: PHC STEFANIK 2016 (Slovakia)

Project acronym: AIMPC

Project title: Advanced techniques for practical implementation of model predictive control strategies

Duration: January 2016 - December 2017

Coordinator:Cristina Stoica (France), Martin Gulan (Slovakia)

Abstract: The proposed project is dedicated to the model predictive control with a particular empha-
sis on its practical implementations. The main objective is to explore new techniques allowing for an
efficient deployment of control algorithms on embedded, preferably low-cost microcontrollerbased
computing platforms. The inherent hardware memory/speed issues that become particularly chal-
lenging for fast real-time applications are to be addressed by appropriate acceleration and complex-
ity reduction techniques targeting either the implicit or the explicit control laws while preserving the
optimality of associated solutions. The run-time performance of the proposed control policies will
be experimentally verified and monitored in chosen existing applications.

Program: PHC BOSPHORE 2016 (Turkey)

Project title: Robust Control of Time Delayed Linear Parameter Varying Systems via Switched
Controllers.

Duration: January 2016 - December 2017

Coordinator: Frédéric Mazenc (France), Hitay Özbay (Turkey).

Abstract: The main goal of this project is to develop computational algorithms for robust controller
design for different classes of time delay systems appearing in various engineering applications
such as chemical processes, transportation systems and communications networks. The participants
will consider control problems of significant practical implications in this area: (i) developing new
computational techniques for simple (low order) reliable and scalable decentralized controllers
for control of (and control over) networks; and (ii) reducing conservatism in recently developed
dwell-time based stability results for the analysis of switched time delay systems. Moreover, design
of scalable low order controllers for reducing the effect of time delays is an important problem
investigated in this project. One of the objectives of this collaboration is to generalize the design
techniques already developed by the French and Turkish teams to larger classes of time delay
systems, in particular multi-input-multi-output (MIMO) systems with time varying delays.

Program: COST Action

Project acronym: FRACTAL

Project title: Fractional-order systems; analysis, synthesis and their importance for future design

Duration: November 2016 - October 2020

Coordinator: Jaroslav Koton Czech Republic

Abstract: Fractional-order systems have lately been attracting significant attention and gaining more
acceptance as generalization to classical integer-order systems. Mathematical basics of fractional-
order calculus were laid nearly 300 years ago and since that it has gained deeply rooted mathematical
concepts. Today, it is known that many real dynamic systems cannot be described by a system of
simple differential equation or of integer-order system. In practice we can encounter such systems in
electronics, signal processing, thermodynamics, biology, medicine, control theory, etc. The Action
will favour scientific advancement in above mentioned areas by coordinating activities of academic
research groups towards an efficient deployment of fractal theory to industry applications.

7.4. International Initiatives

Catherine Bonnet is the co-supervisor together with André Fioravanti of a PhD student of Unicamp (Brazil).

Guillaume Sandou is the co-supervisor of a PhD student in the Ecole nationale d'ingénieur de Tunis, on the
optimal tuning of MPC controllers using stochastic optimization methods.

7.4.1. Inria International Labs

7.4.1.1. Informal International Partners

- College of Mathematics and Information Science, Shaanxi Normal University, China

- School of Control Science and Engineering, Dalian University of Technology, Dalian, China

- Louisiana State University, Baton Rouge, USA

- School of Electrical Engineering at the Tel-Aviv University, Israel

- The University of Texas at Austin, Dept. of Aerospace Engineering & Engineering Mechanics, USA

- Blikent University, Turkey

- Universidad de Chile, Chile

- School of Mathematics, University of Leeds, U.K.

- University Federale Rio de Janeiro, Brazil

- UNICAMP, Brazil

- Kyoto University, Japan

7.4.2. Participation in Other International Programs

7.4.2.1. International Initiatives

STADE

Title: Stability and Dichotomies in Differential Equations (Ordinary & Delay).

International Partners (Institution - Laboratory - Researcher):

Universidad de Chile (Chile) - Mathematics Department - Gonzalo Robledo

Universidad de la Republica Uruguay (Uruguay) - Faculty of Engineering - Pablo Monzon

Duration: 2016 - 2017

Start year: 2016

See also: http://www.stade.cl/pages/list.html

The ship-flags of this project are the concepts of dichotomy and stability in an ODE & DDE framework. We intend to study some theoretical and applied problems involving these concepts and its relations. In particular, converse stability results (expressed in the existence of density functions), feedback stabilization, stability in delay differential equations and some applications to bioprocesses.

7.5. International Research Visitors

Gonzalo Robledo, Universidad de Chile, Chile, $14/11 - 28/11$.

Hitay Ozbay, Bilkent University, $26/10 - 02/11$.

Saed Ahmed, Bilkent University, $04/12 - 16/12$.

7.5.1. Visits of International Scientists

Stefanella Boatto, Federale University Rio de Janeiro, Brazil, 2 October-23 December

André Fioravanti, UNICAMP, Sao Paulo, Brazil, 24 November-31 December

Emilia Fridman, Tel-Aviv University, Israel, 23-30 September

Yutaka Yamamoto, Kyoto University, Japan, 6 September-19 November

7.5.2. Visits to International Teams

7.5.2.1. Research Stays Abroad

Matsumae International Foundation (MIF) fellowship - 3 months research visit of Sorin Olaru (June-September 2016) to Kyushu Institute of Technology (Hosted by Prof. Hiroshi Ito).

Mitacs Globalink Research Award – 3 months research visit of Dina Irofti (July – October 2016) to University of Lethbridge, Alberta, Canada (hosted by Marc R. Roussel).

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Organisation

8.1.1.1. General Chair, Scientific Chair

- Catherine Bonnet is together with Alexandre Chapoutot (ENSTA ParisTech) and Paolo Mason (L2S) the co-organizer of the Working Group Shy of Digicosme on the Plateau de Saclay.

- Catherine Bonnet was together with Alexandre Chapoutot and Laurent Fribourg (ENS Cachan) the co-organizer of the DigiCosme Spring School 2016 on Hybrid Systems, May 9-13, ENSTA ParisTech.

- Sorin Olaru was the Scientific organizer of the Workshop "Interpolation-based techniques for constrained control: from improved vertex control to robust model predictive control alternatives" at ECC 201

8.1.2. Scientific Events Selection

Frédéric Mazenc was Associate Editor for the conferences 2017 American Control Conference, Seattle, USA and the 55th IEEE Conference on Decision and Control, Las Vegas, USA, (2016).

Ali Zemouche was Associate Editor for the conferences 2017 American Control Conference, Seattle, USA, and the 55th IEEE Conference on Decision and Control, Las Vegas, USA, (2016).

8.1.2.1. Member of the Conference Program Committees

Catherine Bonnet was a member of the International Program Committee of the IFAC Conference on Time-Delay Systems - IFAC TDS2016.

Sorin Olaru was a member of the conference program committee: International Conference on System Theory, Control and Computing - ICSTCC 2016.

Guillaume Sandou was a member of the program committee of the 2016 IEEE Symposium on Computational Intelligence in Production and Logistics Systems, Athens, Greece.

Ali Zemouche was a member of the International Program Committee of the IFAC ACD16 Conference, Lille, November 2016.

Ali Zemouche is a member of the Technical Program Committee of the 2017 IEEE - ACC Conference, Seattle, May 2017.

8.1.2.2. Reviewer

The team reviewed many papers for international Conferences eg IEEE CDC 2016; IEEE ACC 2016, IFAC TDS 2016

8.1.3. Journal

8.1.3.1. Member of the Editorial Boards

Frédéric Mazenc is Member of the Mathematical Control and Related Fields editorial board.

Frédéric Mazenc is Member of the European Journal of Control editorial board.

Frédéric Mazenc is Associate Editor for the Asian Journal of control.

Frédéric Mazenc is Associate Editor for the Journal of Control and Decision.

Frédéric Mazenc is Associate Editor for IEEE Transactions on Automatic Control.

Sorin Olaru is a Member of the editorial board of IMA Journal of Mathematical Control and Information

Ali Zemouche is Member of the European Journal of Control editorial board.

Ali Zemouche is Associate Editor for SIAM Journal on Control and Optimization.

8.1.3.2. Reviewer - Reviewing Activities

The team reviewed many papers for international journals of mathematics, e.g. European Journal of Control, Automatica, IEEE Trans. Aut. Contr., SICON, IEEE Trans. on Control Systems Technology, Journal of Process Control, Asian Journal of Control, transactions on control systems and technology journal as well as a book for Springer.

8.1.4. Invited Talks

Frédéric Mazenc was a speaker of 'The 5th International Symposium on Positive Systems'. September 14th-16th 2016, Universita Campus Bio-Medico di Roma, Italy. Title of his talk: *Stability analysis of a differential-difference system through a linear Lyapunov functional design.*

Frédéric Mazenc was a plenary speaker of the workshop 'Stability and Control of Infinite-Dimensional Systems', 12-14 October 2016, Passau, Germany. Title of his talk: *New trajectory based approach for systems with delay: application to the reduction model technique.*

8.1.5. Leadership within the Scientific Community

Catherine Bonnet and Sorin Olaru are members of the IFAC Technical Committees Robust Control.

Catherine Bonnet is a member of the IFAC Technical Committees Distributed Parameter Systems and Non linear Control Systems. She is a member of the SIAG/CST (SIAM Activity group Control System Theory) steering committee.

For 'The 5th International Symposium on Positive Systems', September 14th-16th 2016, Universita Campus Bio-Medico di Roma, Frédéric Mazenc organized an invited session entitled *Positive systems with delay.*

8.1.6. Scientific Expertise

Catherine Bonnet is a member of the Evaluation Committee of Inria since September 2015.

Catherine Bonnet has been an expert for ANR.

Since 2014, Frédéric Mazenc is an expert for the FNRS (Belgium). His mission consists in evaluating research projects funded by this institution.

Since 2012, Frédéric Mazenc is a, expert for the ANVUR (National Agency for the Evaluation of Universities and Research Institutes, Italy). His mission consists in evaluating the contribution of Italian scientists.

Since 2011, Frédéric Mazenc is a, expert for the Romanian National Council for Development and Innovation (Romania). His mission consists in evaluating research projects funded by the this institution.

8.1.7. Research Administration

Catherine Bonnet is a Management Committee member of the COST action *Fractional-order systems; analysis, synthesis and their importance for future design*, member of the board of Directors of the consortium Cap'Maths, of the administration council of the association *Femmes et Mathématiques*, of the Inria Parity Committee (created in 2015) and of the *Cellule veille et prospective* of Inria.

Frédéric Mazenc and Sorin Olaru are members of the Conseil du Laboratoire of Laboratoire des Signaux et Systèmes (L2S).

Frédéric Mazenc is president of the commission scientifique du CRI Saclay-Ile-de-France.

Frédéric Mazenc is member of the Bureau du Comité des Projets du CRI Saclay-Ile-de-France.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Licence : Guillaume Sandou, Signals and Systems, 87h, L3, CentraleSupélec

Licence : Guillaume Sandou, Mathematics and programming, 18h, L3, CentraleSupélec

Licence : Sorin Olaru, Numerical methods and Optimization, 24h, niveau M1, SUPELEC, France

Licence : Sorin Olaru, Hybrid systems, 16h, M2, SUPELEC, France

Licence : Sorin Olaru, Automatic Control, 8h , M1, SUPELEC, France

Licence : Sorin Olaru, Signals and systems, 8h , L3, SUPELEC, France

Licence : Sorin Olaru, Embedded systems, 8h , M1, Centrale Paris, France

Master : Dina Irofti, Java programming, 40h, M2, Paris-Sud

Master : Dina Irofti, Industrial computing, 16h, M1, Paris-Sud

Master : Guillaume Sandou, Automatic Control, 8h, M1, CentraleSupélec

Master : Guillaume Sandou, Numerical methods and optimization, 28h, M1 and M2, Centrale-Supélec

Master : Guillaume Sandou, Modelling and system stability analysis, 21h, M2, CentraleSupélec

Master : Guillaume Sandou, Control of energy systems, 22h, M2, CentraleSupélec

Master : Guillaume Sandou, Robust control and mu-analysis, 9h, M2, CentraleSupélec

Master : Guillaume Sandou, Systems identification, 32h, M2, ENSTA

Master : Guillaume Sandou, System Analysis, 22h, M2, Ecole des Mines de Nantes

F. Mazenc: March 2016. Teaching (in English) for the *International Graduate School on Control* of the EECI (Master level). 21 hours). Subject : introduction to the ordinary differential equations, Lypunov design, control and observation of nonlinear dynamical systems.

8.2.2. Supervision

PhD in progress : Saeed Ahmed, Bilkent University, Stability analysis and control of switched systems with time-delay. Supervisor : Hitay Ozbay. Co-supervisor : Frédéric Mazenc.

PhD in progress : Nadine Aoun, Modélisation de réseaux de chaleur et gestion avancée multi-échelles de la production, de la distribution et de la demande. Supervisor: Guillaume Sandou.

PhD in progress : Caetano Cardeliquio, Stability and stabilization of (possibly fractional) systems with delays. French Supervisor : Catherine Bonnet, Brazilian Supervisor : André Fioravanti.

PhD in progress : Walid Djema, Analysis of an AML model enabling evaluation of polychemio-therapies delivered in the case of AML which have a high level of Flt-3 duplication (Flt-3-ITD). Supervisor : Catherine Bonnet. Co-supervisors : Jean Clairambault and Frédéric Mazenc.

PhD : Sophie Frasnedo, Optimisation globale des lois de commande des autodirecteurs sur critére optronique : application à un autodirecteur à double phase de stabilisation. Supervisors : Gilles Duc et Guillaume Sandou, soutenue le 6 décembre 2016

PhD in progress : Nicolo Gionfra, Optimisation du pilotage d'un parc d'énergies renouvelables avec stockage et du réseau de distribution sous-jacent. Supervisors: Houria Siguerdidjane et Guillaume Sandou.

PhD : Mohamad Koteich, Modélisation et Observabilité des Machines Electriques en vue de la commande sans capteur mécanique. Supervisors: Gilles Duc et Guillaume Sandou, soutenue le 18 mai 2016.

PhD in progress : Mohamed Lotfi Derouiche, Sur l'optimisation par métaheuristiques avancées de lois de commande prédictive non linéaire. Supervisor: Soufienne Bouallegue, Joseph Haggége et Guillaume Sandou.

PhD in progress : Juliette Pochet, Analyse de performance et de résilience d'une ligne de type RER équipée d'un automatisme CBTC. Supervisor: Guillaume Sandou.

8.2.3. Juries

Catherine Bonnet was a reviewer of the PhD thesis of Marine Jacquier entitled "*Mathematical modeling of the hormonal regulation of food intake and body weight - Application to caloric restriction and leptin resistance*", University of Lyon 1, February 5th 2016.

Catherine Bonnet was a member of several recruiting committees: Junior Researcher competition in Inria Grenoble - Rhône-Alpes, Senior Researcher competition at Inria, Professor competition at University of Perpignan, Preofessor competition at CentraleSupelec.

Frédéric Mazenc was a reviewer of the PhD thesis of Youssef Bourfia, entitled "*Modélisation et Analyse de Modèles en Dynamique Cellulaire avec Applications à des Problèmes Liés aux Cancers*", (University of Cadi Ayyad de Marrakech and University of Pierre et Marie Curie, December 28, 2016).

Frédéric Mazenc was an examiner of the HDR of Ali Zemouche, entitled "*State Observer Design and Stabilization of Nonlinear Systems via LMIs*". Centre de Recherche en Automatique de Nancy, UMR 7039 CNRS - Université de Lorraine, December 01, 2016.

Sorin Olaru has been appointed as evaluator for V. Grelet's PhD thesis at University of Lyon. The thesis was defended on February 18th 2016.

Guillaume Sandou was a reviewer of the PhD thesis Modélisation dynamique et gestion avancée de réseaux de chaleur, Loic Giraud, Université Grenoble Alpes

Guillaume Sandou was a member of the following PhD theses committees:

- Roman Le Goff Latimier, " *Gestion et dimensionnement d'une flotte de véhicules électriques associée Ã une centrale photovoltaique : co-optimisation stochastique et distribuée*", 26 septembre 2016.
- Sinziana Carloganu, "*Evaluation des produits d'effacements réalisés sur un ensemble de consommateurs d'électricité par identification d'un modéle global*", 8 décembre 2016.

Ali Zemouche was an examiner of two PhD theses under the supervision of Professor Hieu TRINH from Deakin University, Geelong, Australia:

- The PhD thesis of Minh Cuong Nguyen, entitled "*State Observer Design Methods for Lipschitz Time-Delay Systems*", Deakin University, Geelong, Australia, September 15, 2016.
- The PhD thesis of Ngoc Thanh Pham, entitled "*Robust Load Frequency Control of Interconnected Grids with Electric Vehicles*", Deakin University, Geelong, Australia, December 9, 2016.

9. Bibliography

Major publications by the team in recent years

[1] C. BONNET, J. PARTINGTON.*Analysis of fractional delay systems of retarded and neutral type*, in "Automatica", 2002, vol. 38, p. 1133–1138.

[2] C. BONNET, J. PARTINGTON.*Stabilization of some fractional delay systems of neutral type*, in "Automatica", 2007, vol. 43, p. 2047–2053.

[3] M. MALISOFF, F. MAZENC.*Constructions of Strict Lyapunov Functions*, Communications and Control Engineering Series, Springer-Verlag London Ltd., 2009.

[4] F. MAZENC, P. BLIMAN.*Backstepping design for time-delay nonlinear systems*, in "IEEE Transactions on Automatic Control", January 2006, vol. 51, n° 1, p. 149–154.

[5] W. MICHIELS, S.-I. NICULESCU.*Stability and Stabilization of Time-Delay Systems. An Eigenvalue-Based Approach*, Advances in Design and Control, SIAM: Philadelphia, 2007, vol. 12.

[6] S.-I. NICULESCU.*Delay Effects on Stability: a Robust Control Approach*, Lecture Notes in Control and Information Sciences, Springer, 2001, vol. 269.

[7] S. OLARU, D. DUMUR.*Avoiding constraints redundancy in predictive control optimization routines*, in "IEEE Trans. Automat. Control", 2005, vol. 50, n° 9, p. 1459–1465.

[8] G. SANDOU.*Particle swarm optimization: an efficient tool for the design of automatic control law*, in "European Control Conference", Budapest, Hungary, August 23rd-26th 2009.

Publications of the year

Articles in International Peer-Reviewed Journal

[9] N. ANH NGUYEN, S. OLARU, P. RODRIGUEZ-AYERBE, G. BITSORIS, M. HOVD.*Explicit robustness and fragility margins for linear discrete systems with piecewise affine control law*, in "Automatica", 2016, vol. 68, p. 334-343 [*DOI :* 10.1016/J.AUTOMATICA.2015.10.048], https://hal-centralesupelec.archives-ouvertes.fr/hal-01408554.

[10] N. ANH NGUYEN, S. S. OLARU, P. RODRIGUEZ-AYERBE, M. M. HOVD, I. I. NECOARA.*Constructive Solution of Inverse Parametric Linear/Quadratic Programming Problems*, in "Journal of Optimization Theory and Applications", 2016 [*DOI :* 10.1007/S10957-016-0968-0], https://hal-centralesupelec.archives-ouvertes.fr/hal-01408472.

[11] I. BOUSSAADA, S.-I. NICULESCU.*Characterizing the Codimension of Zero Singularities for Time-Delay Systems : A Link with Vandermonde and Birkhoff Incidence Matrices*, in "Acta Applicandae Mathematicae", October 2016, vol. 145, n° 1, p. 47-88 [*DOI :* 10.1007/S10440-016-0050-9], https://hal-centralesupelec.archives-ouvertes.fr/hal-01429961.

[12] I. BOUSSAADA, S.-I. NICULESCU.*Tracking the Algebraic Multiplicity of Crossing Imaginary Roots for Generic Quasipolynomials: A Vandermonde-Based Approach*, in "IEEE Transactions on Automatic Control", June 2016, vol. 61, n° 6, p. 1601-1606 [*DOI :* 10.1109/TAC.2015.2480175], https://hal-centralesupelec.archives-ouvertes.fr/hal-01425757.

[13] J. CHEN, P. FU, C. F. MÉNDEZ BARRIOS, S.-I. NICULESCU, H. ZHANG.*Stability Analysis of Polynomially Dependent Systems by Eigenvalue Perturbation*, in "IEEE Transactions on Automatic Control", 2017, vol. 62, accepté le 23 décembre 2016 [*DOI :* 10.1109/TAC.2016.2645758], https://hal-centralesupelec.archives-ouvertes.fr/hal-01431810.

[14] E. FRIDMAN, C. BONNET, F. MAZENC, W. DJEMA.*Stability of the cell dynamics in acute myeloid leukemia*, in "Systems and Control Letters", 2016 [*DOI :* 10.1016/J.SYSCONLE.2015.09.006], https://hal.inria.fr/hal-01257577.

[15] M.-T. LARABA, S. OLARU, S.-I. NICULESCU, F. BLANCHINI, G. GIORDANO, D. CASAGRANDE, S. MIANI.*Guide on set invariance for delay difference equations*, in "Annual Reviews in Control", 2016, vol. 41, p. 13 - 23 [*DOI :* 10.1016/J.ARCONTROL.2016.04.020], https://hal-centralesupelec.archives-ouvertes.fr/hal-01408474.

[16] X.-G. LI, S.-I. NICULESCU, A. CELA, L. ZHANG, X. LI.*A frequency-sweeping framework for stability analysis of time-delay systems*, in "IEEE Transactions on Automatic Control", 2017, vol. 64, accepté le 21 octobre 2016 [*DOI :* 10.11.09/TAC.2016.2633533], https://hal-centralesupelec.archives-ouvertes.fr/hal-01431725.

[17] F. MAZENC, E. FRIDMAN.*Predictor-based sampled-data exponential stabilization through continuous–discrete observers*, in "Automatica", 2016, vol. 63, 7 [*DOI :* 10.1016/J.AUTOMATICA.2015.10.016], https://hal.inria.fr/hal-01248072.

[18] F. MAZENC, M. MALISOFF.*New Control Design for Bounded Backstepping under Input Delays*, in "Automatica", April 2016, vol. 66, 8, https://hal.inria.fr/hal-01257338.

[19] F. MAZENC, M. MALISOFF.*Reduction model approach for linear time-varying systems with input delays based on extensions of Floquet theory*, in "Systems & Control Letters", 2016, vol. 94 [*DOI :* 10.1016/J.SYSCONLE.2016.05.013], https://hal.inria.fr/hal-01389860.

[20] F. MAZENC, M. MALISOFF.*Stability Analysis for Time-Varying Systems with Delay using Linear Lyapunov Functionals and a Positive Systems Approach*, in "IEEE Transactions on Automatic Control", 2016 [*DOI :* 10.1109/TAC.2015.2446111], https://hal.inria.fr/hal-01257336.

[21] F. MAZENC, M. MALISOFF.*Stabilization of Nonlinear Time-Varying Systems through a New Prediction Based Approach*, in "IEEE Transactions on Automatic Control", August 2016, 1 [*DOI :* 10.1109/TAC.2016.2600500], https://hal.inria.fr/hal-01389859.

[22] C. F. MÉNDEZ BARRIOS, S.-I. NICULESCU, F. LORENZ.*Sampling decomposition and asymptotic zeros behavior of sampled-data SISO systems. An eigenvalue-based approach*, in "IMA Journal on Mathematical Control and Information", December 2016, vol. 33, p. 1177-1197 [*DOI :* 10.1093/IMAMCI/DNV029], https://hal-centralesupelec.archives-ouvertes.fr/hal-01284740.

[23] L. H. V. NGUYEN, C. BONNET, A. FIORAVANTI.*Hinfty-stability analysis of fractional delay systems of neutral type*, in "SIAM Journal on Control and Optimization", 2016, vol. 54, n° 2, p. 740-759, https://hal.inria.fr/hal-01271889.

[24] H.-N. NGUYEN, P.-O. GUTMAN, S. OLARU, M. HOVD.*Explicit improved vertex control for uncertain, time-varying lineardiscrete-time systems with state and control constraints*, in "International Journal of Robust and Nonlinear Control", 2016, vol. 26, n° 12, p. 2652-2667 [*DOI :* 10.1002/RNC.3466], https://hal.archives-ouvertes.fr/hal-01397717.

[25] M. B. SALDIVAR, M. SABINE, S.-I. NICULESCU, H. MOUNIER, I. BOUSSAADA.*A control oriented guided tour in oilwell drilling vibration modeling*, in "Annual Reviews in Control", September 2016, vol. 42, p. 100-113 [*DOI :* 10.1016/J.ARCONTROL.2016.09.002], https://hal-centralesupelec.archives-ouvertes.fr/hal-01425845.

[26] Y.-E. WANG, X.-M. SUN, F. MAZENC.*Stability of switched nonlinear systems with delay and disturbance*, in "Automatica", 2016, vol. 69, p. 78 - 86 [*DOI :* 10.1016/J.AUTOMATICA.2016.02.015], https://hal.inria.fr/hal-01391683.

[27] S. R. YANG, M. R. AKELLA, F. R. MAZENC.*Immersion and Invariance Observers for Gyro-Free Attitude Control Systems*, in "Journal of Guidance, Control, and Dynamics", November 2016, vol. 39, p. 2570 - 2577 [*DOI :* 10.2514/1.G002095], https://hal.inria.fr/hal-01389858.

Invited Conferences

[28] C. BONNET, Y. YAMAMOTO.*Coprimeness of Fractional Representations*, in "55th IEEE Conference on Decision and Control", Las Vegas, United States, December 2016, https://hal.inria.fr/hal-01413310.

[29] C. BONNET, Y. YAMAMOTO.*Stabilizability of Some Neutral Delay Systems with Chains of Poles Clustering on the Imaginary Axis*, in "Mathematical Theory of Networks and Systems", Minneapolis, United States, July 2016, https://hal.inria.fr/hal-01413303.

International Conferences with Proceedings

[30] A. ALESSANDRI, A. ZEMOUCHE.*On the enhancement of high-gain observers for state estimation of nonlinear systems*, in "55th IEEE Conference on Decision and Control, CDC 2016", Las Vegas, NV, United States, December 2016, https://hal.archives-ouvertes.fr/hal-01418758.

[31] H. BIBI, F. BEDOUHENE, H. KHELOUFI, A. ZEMOUCHE, H. TRINH.*Observer-based control design via LMIs for a class of switched discrete-time linear systems with parameter uncertainties*, in "55th IEEE Conference on Decision and Control, CDC 2016", Las Vegas, NV, United States, December 2016, https://hal.archives-ouvertes.fr/hal-01418759.

[32] M. L. DEROUICHE, S. BOUALLÈGUE, J. HAGGÈGE, G. SANDOU.*LabVIEW Perturbed Particle Swarm Optimization Based Approach for Model Predictive Control Tuning*, in "4th IFAC International Conference on Intelligent Control and Automation Sciences (ICONS 2016)", Reims, France, June 2016, https://hal-centralesupelec.archives-ouvertes.fr/hal-01347041.

[33] W. DJEMA, F. MAZENC, C. BONNET, J. CLAIRAMBAULT, P. HIRSCH, F. DELHOMMEAU.*Stability of a Delay System Coupled to a Differential-Difference System Describing the Coexistence of Ordinary and Mutated Hematopoietic Stem Cells*, in "Conference on Decision and Control", Las Vegas, United States, December 2016, https://hal.inria.fr/hal-01389870.

[34] W. DJEMA, F. MAZENC, C. BONNET.*Analysis of a Nonlinear Delay Differential-Difference Biological Model*, in "IFAC Time Delay Systems", Istanbul, Turkey, July 2016, https://hal.inria.fr/hal-01389867.

[35] W. DJEMA, F. MAZENC, C. BONNET.*Stability of immature cell dynamics in healthy and unhealthy hematopoiesis*, in "2016 American Control Conference", Boston, United States, July 2016, p. 6121 - 6126 [*DOI :* 10.1109/ACC.2016.7526631], https://hal.inria.fr/hal-01389865.

[36] N. GIONFRA, H. SIGUERDIDJANE, G. SANDOU, D. FAILLE, P. LOEVENBRUCK.*Combined Feedback Linearization and MPC for Wind Turbine Power Tracking*, in "IEEE Multi-Conference on Systems and Control (MSC 2016)", Buenos Aires, Argentina, September 2016 [*DOI :* 10.1109/CCA.2016.7587821], https://hal-centralesupelec.archives-ouvertes.fr/hal-01381416.

[37] U. HAKKI, I. BOUSSAADA, S.-I. NICULESCU.*Further remarks on delay dynamics in Oregonator models*, in "IEEE International Conference on Control and Automation", Kathmandu, Nepal, June 2016 [*DOI :* 10.1109/ICCA.2016.7505261], https://hal-centralesupelec.archives-ouvertes.fr/hal-01445899.

[38] H. KHELOUFI, B. CHERIFA, A. ZEMOUCHE, F. BEDOUHENE, H. TRINH.*New decentralized control design for interconnected nonlinear discrete-time systems with nonlinear interconnections*, in "55th IEEE Conference on Decision and Control, CDC 2016", Las Vegas, NV, United States, December 2016, https://hal.archives-ouvertes.fr/hal-01418757.

[39] F. MAZENC, J. HARMAND, M. MALISOFF.*Stabilization in a chemostat with sampled and delayed measurements*, in "2016 American Control Conference", Boston , United States, July 2016, p. 1857 - 1862 [*DOI :* 10.1109/ACC.2016.7525189], https://hal.inria.fr/hal-01389864.

[40] F. MAZENC, M. MALISOFF.*Constructions of Strict Lyapunov Functions: Stability, Robustness, Delays, and State Constraints*, in "MTNS", Minneapolis, United States, July 2016, https://hal.inria.fr/hal-01389866.

[41] F. MAZENC, M. MALISOFF.*Extension of Razumikhin's theorem for time-varying systems with delay*, in "2016 American Control Conference", Boston , United States, July 2016, 6 [*DOI :* 10.1109/ACC.2016.7524896], https://hal.inria.fr/hal-01389863.

[42] F. MAZENC, M. MALISOFF.*New Prediction Approach for Stabilizing Time-Varying Systems under Time-Varying Input Delay*, in "CDC 2016 - 55th IEEE Conference on Decision and Control", Las Vegas, United States, December 2016, https://hal.inria.fr/hal-01389869.

[43] F. MAZENC, M. MALISOFF, J. WESTON.*New Bounded Backstepping Control Designs for Time-Varying Systems under Converging Input Converging State Conditions*, in "Conference on Decision and Control", Las Vegas, United States, December 2016, https://hal.inria.fr/hal-01389868.

[44] M. T. NGUYEN, L. RODRIGUES, C. STOICA MANIU, S. OLARU.*Discretized optimal control approach for dynamic Multi-Agent decentralized coverage*, in "IEEE Multi-Conference on Systems and Control", Buenos Aires, Argentina, September 2016, p. 335-340 [*DOI :* 10.1109/ISIC.2016.7579984], https://hal.archives-ouvertes.fr/hal-01390082.

[45] M. T. NGUYEN, C. STOICA MANIU, S. OLARU.*Decentralized constructive collision avoidance for Multi-Agent dynamical systems*, in "European Control Conference", Aalborg, Denmark, June 2016, p. 1526-1531, https://hal.archives-ouvertes.fr/hal-01390087.

[46] J. POCHET, S. BARO, G. SANDOU.*Supervision and rescheduling of a mixed CBTC traffic on a suburban railway line*, in "IEEE International Conference on Intelligent Rail Transportation (ICIRT 2016)", Birmingham, United Kingdom, August 2016 [*DOI :* 10.1109/ICIRT.2016.7588547], https://hal-centralesupelec.archives-ouvertes.fr/hal-01381366.

[47] A. QUADRAT.*A constructive algebraic analysis approach to Artstein's reduction of linear time-delay systems*, in "12th IFAC Workshop on Time Delay Systems", Ann Arbor, United States, Proceedings of 12th IFAC Workshop on Time Delay Systems, University of Michigan, May 2016, https://hal-centralesupelec.archives-ouvertes.fr/hal-01259862.

[48] N. STROE, G. COLIN, K. BEN-CHERIF, S. OLARU, Y. CHAMAILLARD.*Towards a generic control-oriented model for HEV predictive energy management*, in "8th IFAC International Symposium on Advances in Automotive Control (AAC2016)", Kolmarden, Sweden, IFAC-PapersOnLine, June 2016, p. 266-271, https://hal-univ-orleans.archives-ouvertes.fr/hal-01338065.

Conferences without Proceedings

[49] P. AHMADI-MOSHKENANI, T. A. JOHANSEN, S. OLARU.*On degeneracy in exploration of combinatorial tree in multi-parametric quadratic programming*, in "Decision and Control (CDC), 2016 IEEE 55th Conference on", Las Vegas, United States, December 2016, vol. 26, p. 2320 - 2326 [*DOI :* 10.1109/CDC.2016.7798609], https://hal-centralesupelec.archives-ouvertes.fr/hal-01408493.

[50] P. AHMADI-MOSHKENANI, S. OLARU, T. A. JOHANSEN.*Further Results on the Exploration of Combinatorial Tree in Multi-Parametric Quadratic Programming*, in "European Control Confrence 2016", Aalborg, Denmark, June 2016, https://hal-centralesupelec.archives-ouvertes.fr/hal-01429237.

[51] I. BALLESTEROS-TOLOSANA, S. OLARU, P. RODRIGUEZ-AYERBE, G. PITA-GIL, R. DEBORNE.*Comparison of optimization-based strategies for constrained control of Auto-Steering systems*, in "European Control Confrence 2016", Aalborg, Denmark, June 2016, https://hal-centralesupelec.archives-ouvertes.fr/hal-01429244.

[52] M. L. DEROUICHE, S. BOUALLÈGUE, J. HAGGÈGE, G. SANDOU.*Rapid Model Predictive Control Prototyping with LabVIEW/CDSim and CompactRIO Target*, in "4th International Conference on Control Engineering & Information Technology (CEIT-2016)", Hammamet, Tunisia, December 2016, https://hal-centralesupelec.archives-ouvertes.fr/hal-01415057.

[53] N. O. GIONFRA, G. SANDOU, H. SIGUERDIDJANE, D. FAILLE, P. LOEVENBRUCK.*Optimisation du pilotage d'un parc d'énergies renouvelables*, in "3ième Journée Scientifique de l'Institut RISEGrid", Palaiseau, France, September 2016, https://hal-centralesupelec.archives-ouvertes.fr/hal-01381714.

[54] N. O. GIONFRA, H. SIGUERDIDJANE, G. SANDOU, D. FAILLE.*Hierarchical Control of a Wind Farm for Wake Interaction Minimization*, in "IFAC Workshop on Control of Transmission and Distribution Smart Grids (CTDSG 2016)", Prague, Czech Republic, October 2016, https://hal-centralesupelec.archives-ouvertes.fr/hal-01381404.

[55] N. GIONFRA, H. SIGUERDIDJANE, G. SANDOU, D. FAILLE, P. LOEVENBRUCK.*Advanced Control for Wind Turbine Grid Connection Requirements*, in "World Congress and Exhibition on Wind & Renewable Energy", Berlin, Germany, July 2016, https://hal-centralesupelec.archives-ouvertes.fr/hal-01381643.

[56] S. MUNIR, M. HOVD, G. SANDOU, S. OLARU.*Controlled contractive sets for low-complexity constrained control*, in "IEEE Multi-Conference on Systems and Control", Buenos Aires, Argentina, September 2016, https://hal-centralesupelec.archives-ouvertes.fr/hal-01381706.

Scientific Books (or Scientific Book chapters)

[57] I. PRODAN, F. STOICAN, S. OLARU, S.-I. NICULESCU.*Mixed-Integer Representations in Control Design: Mathematical Foundations and Applications*, Springer, 2016 [*DOI :* 10.1007/978-3-319-26995-5], https://hal-supelec.archives-ouvertes.fr/hal-01257452.

[58] R. SIPAHI, S.-I. NICULESCU, F. ATAY.*Deterministic car-following traffic models. Delay effects and linear stability*, SpringerBriefs in Electrical and Computer Engineering (Control, Automation and Robotics), Springer : Heidelberg, 2017, accepté : novembre 2016, https://hal-centralesupelec.archives-ouvertes.fr/hal-01431320.

Research Reports

[59] N. A. NGUYEN, M. GULAN, S. OLARU, P. RODRIGUEZ-AYERBE.*Convex Liftings: Theory and Control Applications*, CentraleSupelec, June 2016, https://hal-centralesupelec.archives-ouvertes.fr/hal-01326804.

Other Publications

[60] N. ANH NGUYEN, P. RODRIGUEZ-AYERBE, S. OLARU.*Any Discontinuous PWA Function Is Optimal Solution to a Parametric Linear Programming Problem*, June 2016, working paper or preprint, https://hal-centralesupelec.archives-ouvertes.fr/hal-01259879.

[61] N. ANH NGUYEN, P. RODRIGUEZ-AYERBE, S. OLARU.*Inverse Parametric Linear/quadratic Programming Problem for Continuous PWA Functions Defined on Polyhedral Partitions of Polyhedra*, June 2016, working paper or preprint, https://hal-centralesupelec.archives-ouvertes.fr/hal-01259895.

[62] C. CARDELIQUIO, M. SOUZA, A. FIORAVANTI.*Stability Analysis and Output-Feedback Control Design for Time-Delay Systems* , December 2016, working paper or preprint, https://hal.inria.fr/hal-01409265.

[63] N. GIONFRA, H. SIGUERDIDJANE, G. SANDOU, D. FAILLE, P. LOEVENBRUCK.*Optimisation du contrôle d'un parc éolien*, October 2016, Vers la Transition Energétique, WTE 2016, Poster, https://hal-centralesupelec.archives-ouvertes.fr/hal-01381012.

[64] L. HETEL, C. FITER, H. OMRAN, A. SEURET, E. FRIDMAN, J.-P. RICHARD, S.-I. NICULESCU.*Recent developments on the stability of systems with aperiodic sampling: an overview ⋆*, September 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01363448.

References in notes

[65] M. ADIMY, A. CHEKROUN, T.-M. TOUAOULA.*Global asymptotic stability for an age-structured model of hematopoietic stem cell dynamics*, in "Applicable Analysis", February 2016, p. 1 - 12 [*DOI :* 10.1080/00036811.2016.1139698], https://hal.inria.fr/hal-01396691.

[66] H. KITANO.*Biological robustness*, in "Nature", 2004, vol. 5, p. 826–837.

[67] H. OZBAY, C. BONNET, A. FIORAVANTI.*PID Controller Design for Fractional-Order Systems with Time Delays*, in "Systems and Control Letters", January 2012, vol. 61, n° 1, p. 18-23, http://hal.inria.fr/hal-00766062.

Team EX-SITU

Extreme Interaction

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

RESEARCH CENTER
Saclay - Île-de-France

THEME
Interaction and visualization

Table of contents

Team EX-SITU

Creation of the Team: 2015 January 01

Keywords:

Computer Science and Digital Science:

5. - Interaction, multimedia and robotics
5.1. - Human-Computer Interaction
5.1.1. - Engineering of interactive systems
5.1.2. - Evaluation of interactive systems
5.1.5. - Body-based interfaces
5.1.6. - Tangible interfaces
5.1.7. - Multimodal interfaces
5.1.8. - 3D User Interfaces

Other Research Topics and Application Domains:

1.3. - Neuroscience and cognitive science
9.2.1. - Music, sound
9.5.1. - Psychology

1. Members

Research Scientists

Wendy Mackay [Team leader, Inria, Senior Researcher, HDR]
Theophanis Tsandilas [Inria, Researcher]

Faculty Members

Michel Beaudouin-Lafon [Univ. Paris-Sud, Professor, HDR]
Sarah Fdili Alaoui [Univ. Paris-Sud, Associate Professor]
Cedric Fleury [Univ. Paris-Sud, Associate Professor]

Technical Staff

Olivier Gladin [Engineer SED, Inria]
Rémi Hellequin [Engineer (IR), CNRS-Digiscope]
Amani Kooli [Engineer (IE), CNRS-Digiscope]
Jonathan Thorpe [Engineer (IR), CNRS-Digiscope]

PhD Students

Jessalyn Alvina [Inria]
Ignacio Avellino Martinez [Inria]
Marianela Ciolfi Felice [Univ. Paris-Sud]
Carla Griggio [Inria]
Shu-Yuan Hsueh [Univ. Paris-Sud, from Oct 2016]
Ghita Jalal [Univ. Paris-Sud]
German Leiva [Inria]
Wanyu Liu [Institut Telecom]
Nolwenn Maudet [Univ. Paris-Sud]
Yujiro Okuya [Univ. Paris-Sud]
Philip Tchernavskij [Univ. Paris-Sud, from Oct 2016]
Michael Wessely [Inria]

Oleksandr Zinenko [Inria, until Oct 2016]

Post-Doctoral Fellows
John Mccallum [Inria, from Dec 2016]
Joseph Malloch [Inria, until Sep 2016]

Visiting Scientists
James Hollan [UC San Diego, Professor, from Apr 2016 until Jun 2016]
Joanna Mcgrenere [University of British Columbia, Professor, until Jul 2016]

Administrative Assistant
Alexandra Merlin [Inria, Assistant]

Others
Brennan Jones [Inria, Intern, from Apr 2016 until Sep 2016]
Francesco Vitale [Inria, Intern, from Mar 2016 until Aug 2016]
Dimitri Belopopsky [Univ. Paris-Sud, Intern, from Jun 2016 until Sep 2016]
Sruthi Coimbatore Viswanathan [Inria, Intern, from Mar 2016 until Aug 2016]
Han Giang [Inria, Intern, from Mar 2016 until Aug 2016]
Shu-Yuan Hsueh [Inria, Intern, from Mar 2016 until Sep 2016]
Iva Karabatic [Univ. Paris-Sud, Intern, from Mar 2016 until May 2016]
Diana Lipcanu [Inria, Intern, from Mar 2016 until Aug 2016]
Pierre Mahe [Inria, Intern, from Mar 2016 until Aug 2016]
Niels Mourette [Inria, Intern, from Mar 2016 until Aug 2016]
Midas Nouwens [Inria, Intern, from Mar 2016]
Philip Tchernavskij [Inria, Intern, from May 2016 until Sep 2016]
Isha Van Baar [Inria, Intern, from Mar 2016 until Aug 2016]

2. Overall Objectives

2.1. Overall Objectives

Interactive devices are everywhere: we wear them on our wrists and belts; we consult them from purses and pockets; we read them on the sofa and on the metro; we rely on them to control cars and appliances; and soon we will interact with them on living room walls and billboards in the city. Over the past 30 years, we have witnessed tremendous advances in both hardware and networking technology, which have revolutionized all aspects of our lives, not only business and industry, but also health, education and entertainment. Yet the ways in which we interact with these technologies remains mired in the 1980s. The graphical user interface (GUI), revolutionary at the time, has been pushed far past its limits. Originally designed to help secretaries perform administrative tasks in a work setting, the GUI is now applied to every kind of device, for every kind of setting. While this may make sense for novice users, it forces expert users to use frustratingly inefficient and idiosyncratic tools that are neither powerful nor incrementally learnable.

ExSitu explores the limits of interaction — how extreme users interact with technology in extreme situations. Rather than beginning with novice users and adding complexity, we begin with expert users who already face extreme interaction requirements. We are particularly interested in creative professionals, artists and designers who rewrite the rules as they create new works, and scientists who seek to understand complex phenomena through creative exploration of large quantities of data. Studying these advanced users today will not only help us to anticipate the routine tasks of tomorrow, but to advance our understanding of interaction itself. We seek to create effective human-computer partnerships, in which expert users control their interaction with technology. Our goal is to advance our understanding of interaction as a phenomenon, with a corresponding paradigm shift in how we design, implement and use interactive systems. We have already made significant progress through our work on instrumental interaction and co-adaptive systems, and we hope to extend these into a foundation for the design of all interactive technology — to create a *physics of interaction*.

3. Research Program

3.1. Research Program

We characterize Extreme Situated Interaction as follows:

Extreme users. We study extreme users who make extreme demands on current technology. We know that human beings take advantage of the laws of physics to find creative new uses for physical objects. However, this level of adaptability is severely limited when manipulating digital objects. Even so, we find that creative professionals—artistists, designers and scientists—often adapt interactive technology in novel and unexpected ways and find creative solutions. By studying these users, we hope to not only address the specific problems they face, but also to identify the underlying principles that will help us to reinvent virtual tools. We seek to shift the paradigm of interactive software, to establish the laws of interaction that significantly empower users and allow them to control their digital environment.

Extreme situations. We develop extreme environments that push the limits of today's technology. We take as given that future developments will solve "practical" problems such as cost, reliability and performance and concentrate our efforts on interaction in and with such environments. This has been a successful strategy in the past: Personal computers only became prevalent after the invention of the desktop graphical user interface. Smartphones and tablets only became commercially successful after Apple cracked the problem of a usable touch-based interface for the iPhone and the iPad. Although wearable technologies, such as watches and glasses, are finally beginning to take off, we do not believe that they will create the major disruptions already caused by personal computers, smartphones and tablets. Instead, we believe that future disruptive technologies will include fully interactive paper and large interactive displays.

Our extensive experience with the Digiscope WILD and WILDER platforms places us in a unique position to understand the principles of distributed interaction that extreme environments call for. We expect to integrate, at a fundamental level, the collaborative capabilities that such environments afford. Indeed almost all of our activities in both the digital and the physical world take place within a complex web of human relationships. Current systems only support, at best, passive sharing of information, e.g., through the distribution of independent copies. Our goal is to support active collaboration, in which multiple users are actively engaged in the lifecycle of digital artifacts.

Extreme design. We explore novel approaches to the design of interactive systems, with particular emphasis on extreme users in extreme environments. Our goal is to empower creative professionals, allowing them to act as both designers and developers throughout the design process. Extreme design affects every stage, from requirements definition, to early prototyping and design exploration, to implementation, to adaptation and appropriation by end users. We hope to push the limits of participatory design to actively support creativity at all stages of the design lifecycle.

Extreme design does not stop with purely digital artifacts. The advent of digital fabrication tools and FabLabs has significantly lowered the cost of making physical objects interactive. Creative professionals now create hybrid interactive objects that can be tuned to the user's needs. Integrating the design of physical objects into the software design process raises new challenges, with new methods and skills to support this form of extreme prototyping.

Our overall approach is to identify a small number of specific projects, organized around four themes: *Creativity, Augmentation, Collaboration* and *Infrastructure*. Specific projects may address multiple themes, and different members of the group work together to advance these different topics.

4. Application Domains

4.1. Creative industries

We work closely with creative professionals in the arts and in design, including music composers, musicians, and sound engineers; painters and illustrators; dancers and choreographers; theater groups; graphic and industrial designers; and architects.

4.2. Scientific research

We work with creative professionals in the sciences and engineering, including neuroscientists and doctors; programmers and statisticians; chemists and astrophysicists; and researchers in fluid mechanics.

5. Highlights of the Year

5.1. Highlights of the Year

Michel Beaudouin-Lafon received an ERC Advanced Grant: *ONE – Unified Principles of Interaction*.

Ex-situ had a record of three research papers accepted at ACM/UIST 2016 and eleven research papers accepted at ACM/CHI 2017.

6. New Software and Platforms

6.1. New Software

6.1.1. WildOS

Participant: Michel Beaudouin-Lafon [correspondant].

WildOS is middleware designed to support applications that run in an interactive room, such as our WILD and WILDER rooms, with various interaction resources, including a tiled wall display, a motion tracking system, interactive tabletops, tablets, smartphones and custom-made or 3d printed interactive devices. The conceptual model of WildOS is a *platform*, such as the WILD or WILDER room, that can be described as a set of devices on which one or more applications can be run.

WildOS consists of a server running on a machine that has network access to all the machines involved in the platform, and a set of clients running on the various interaction resources, such as a display cluster or a tablet. Once *WildOS* is running, applications can be started and stopped and devices can be added to or removed from the platform.

WildOS relies on Web technologies, most notably Javascript and node.js, as well as node-webkit and HTML5. This makes it inherently portable (it is currently tested on Mac OS X and Linux). While applications can be developed only with these Web technologies, it is also possible to bridge to existing applications developed in other environments if they provide sufficient access for remote control. Sample applications include a web browser, an image viewer, a window manager, and the BrainTwister application developed in collaboration with neuroanatomists at NeuroSpin.

WildOS is used for several research projects at ExSitu and by other partners of the Digiscope project. It was also deployed on several of Google's interactive rooms in Mountain View, Dublin and Paris. It is available under on Open Source licence at https://bitbucket.org/mblinsitu/wildos.

- ACM: H.5.2 [User Interfaces]: Graphical user interfaces (GUI)
- Software benefit: helps development of multisurface applications.
- OS/Middleware: Crossplatform
- Required library or software: node.js, node-webkit
- Programming language: Javascript

6.1.2. *Unity Cluster*

Participants: Cédric Fleury [correspondant], Olivier Gladin, Jean-Baptiste Louvet.

Unity Cluster is middleware to distribute any Unity 3D (https://unity3d.com/) application on a cluster of computers that run in interactive rooms, such as our WILD and WILDER rooms, or immersive CAVES (Computer-Augmented Virtual Environments). Users can interact the the application with various interaction resources.

Unity Cluster provides an easy solution for running existing Unity 3D applications on any display that requires a rendering cluster with several computers. *Unity Cluster* is based on a master-slave architecture: The master computer runs the main application and the physical simulation as well as manages the input; the slave computers receive updates from the master and render small parts of the 3D scene. *Unity Cluster* manages data distribution and synchronization among the computers to obtain a consistent image on the entire wall-sized display surface.

Unity Cluster can also deform the displayed images according to the user's position in order to match the viewing frustum defined by the user's head and the four corners of the screens. This respects the motion parallax of the 3D scene, giving users a better sense of depth.

Unity Cluster is composed of a set of C Sharp scripts that manage the network connection, data distribution, and the deformation of the viewing frustum. In order to distribute an existing application on the rendering cluster, all scripts must be embedded into a Unity package that is included in an existing Unity project.

- ACM: C.2.4 [Distributed Systems]: Distributed applications, I.3.7 [3D Graphics and Realism]: Virtual reality
- Software benefit: adapts existing Unity 3D application to a rendering cluster of an interactive room.
- OS/Middleware: Crossplatform
- Required library or software: Unity 3D
- Programming language: C Sharp

6.2. Platforms

6.2.1. *WILDER*

Participants: Michel Beaudouin-Lafon [correspondant], Cédric Fleury, Olivier Gladin, Rémi Hellequin, Stéphane Huot, Amani Kooli, Monireh Sanaei, Gabriel Tezier, Jonathan Thorpe.

WILDER (Figure 1) is our second experimental ultra-high-resolution interactive environment, which follows the WILD platform developed in 2009 [2]. It features a wall-sized display with seventy-five 20" LCD screens, i.e. a 5m50 x 1m80 (18' x 6') wall displaying 14 400 x 4 800 = 69 million pixels, powered by a 10-computer cluster and two front-end computers. The platform also features a camera-based motion tracking system that lets users interact with the wall, as well as the surrounding space, with various mobile devices. The display uses a multitouch frame (the largest of its kind in the world) to make the entire wall touch sensitive.

WILDER was inaugurated in June, 2015. It is one of the ten platforms of the Digiscope Equipment of Excellence and, in combination with WILD and the other Digiscope rooms, provides a unique experimental environment for collaborative interaction.

In addition to using WILD and WILDER for our research, we have also developed software architectures and toolkits, such as WildOS and Unity Cluster, that enable developers to run applications on these multi-device, cluster-based systems.
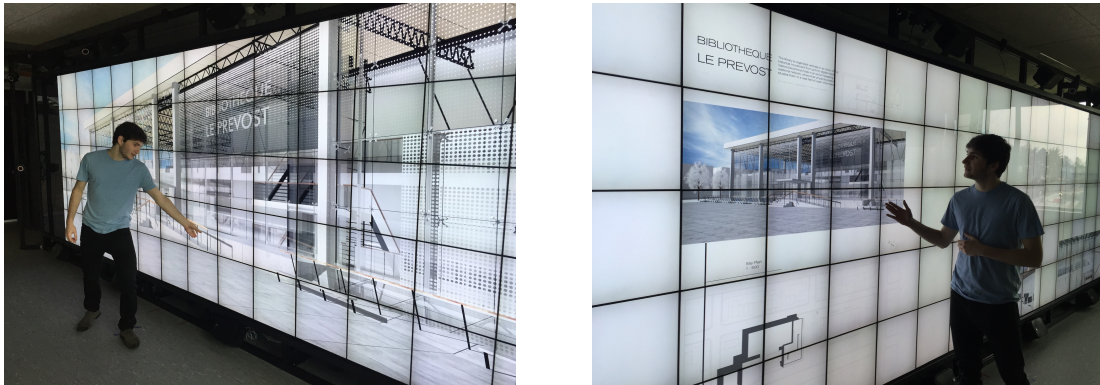
Figure 1. The WILDER platform.

7. New Results

7.1. Fundamentals of Interaction

Participants: Sarah Fdili Alaoui, Michel Beaudouin-Lafon, Cédric Fleury, Wendy Mackay, Theophanis Tsandilas.

In order to better understand fundamental aspects of interaction, ExSitu studies interaction in extreme situations. We conduct in-depth observational studies and controlled experiments which contribute to theories and frameworks that unify our findings and help us generate new, advanced interaction techniques.

StickyLines – Aligning and distributing graphical objects is a common, but cumbersome task. We studied graphic designers and regular users and identified three key problems with current tools: lack of persistence, unpredictability of the results, and inability to 'tweak' the layout. We created *StickyLines* [14], a tool that reifies guidelines into first-class objects: Users can create precise, predictable and persistent interactive alignment and distribution relationships, and can 'tweak' the alignment in a way that can be maintained for subsequent interactions (Figure 2). We ran a [2x2] within-participant experiment to compare *StickyLines* with standard commands and found that *StickyLines* performed up to 40% faster and required up to 50% fewer actions than traditional alignment and distribution commands for complex layouts. Finally, we gave *StickyLines* to six professional designers and found that not only did they quickly adopt it, they also identified novel uses, including creating complex compound guidelines and using them for both spatial and semantic grouping. This work demonstrate the power of reifying concepts, such as alignment and distribution, into first-class objects that can be directly manipulated and appropriated by end users.

UIST Video Browser – We created an interactive video browser that provides a rapid overview of the 30-second video previews of the ACM UIST conference papers, based on the conference schedule [16]. The web application was made available to the 600+ conference attendees, who could see an overview of upcoming talks, search by topic, and create personalized, shareable video playlists that capture the most interesting or relevant papers. Reifying playlists into first-class objects and applying instrumental interaction concepts helped create a fluid and efficient interface.

In(SITE) – We explored touch-based 3D interaction in the situation where users are immersed in a 3D virtual environment and move in front of a large multi-touch wall-sized display. We designed In(SITE) [20], a bimanual touch-based technique combined with object teleportation features which enables users to perform 3D object manipulation on a large vertical display (Figure 3). This technique was compared with a standard 3D interaction technique. The results showed that participants can reach the same level of performance for
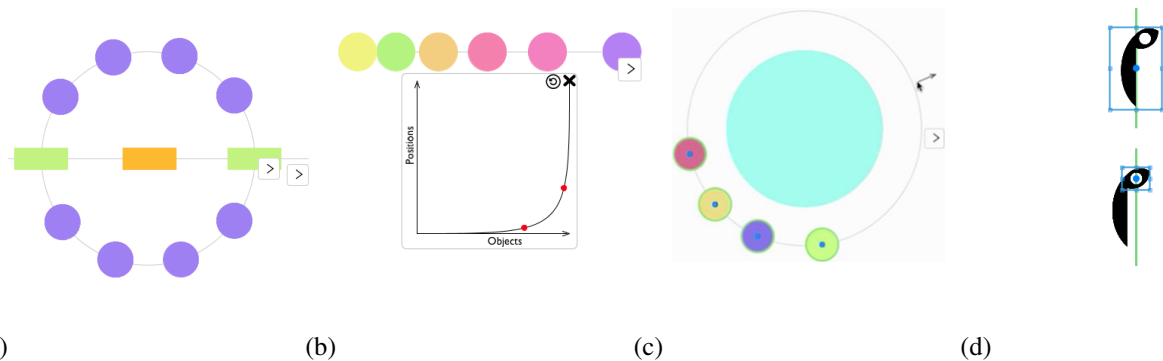
(a)   (b)   (c)   (d)

Figure 2. StickyLines reify alignment and distribution into first-class graphical objects that users can manipulate directly. (a) Circular and horizontal alignments. (b) Non-linear distribution. (c) Ghost guideline. (d) Tweaking an object's bounding box. .

completion time and a better precision for fine adjustments with the In(SITE) technique. They also revealed that combining object teleportation with both techniques improves translation tasks in terms of ease of use, fatigue, and user preference.
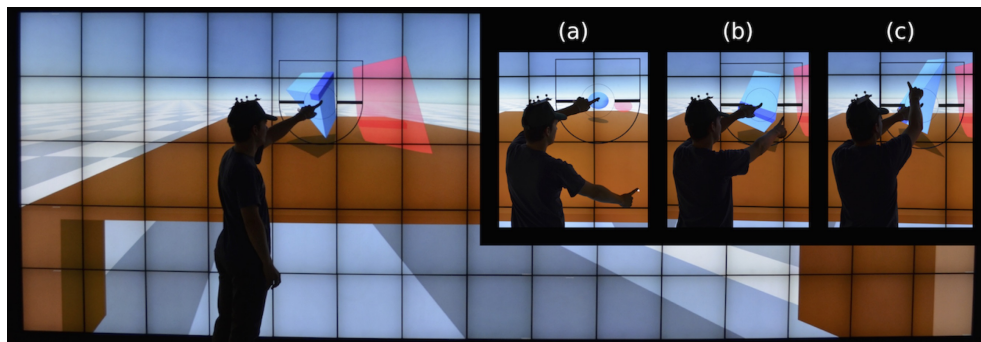
Figure 3. 3D manipulation on a multi-touch wall-sized display combining bimanual interaction and teleportation. The user is performing a xy translation (main pict.), z translation (a), roll rotation (b), and pitch & yaw rotation(c).

In collaboration with Inria Lille (MJOLNER group) and Univ. Strasbourg, we applied our design principles for instrumental interaction to create new interactive tools for the parallelization of programs, a highly specialized task that is currently done by expert developers. Current programming models, languages and tools do not help developers restructure existing programs for more effective execution. At the same time, automatic approaches are overly conservative and imprecise to achieve sufficient performance. We introduced *interactive program restructuring* [28], [11] to bridge the gap between semi-automatic program manipulation and software visualization. First, we extended a state-of-the-art polyhedral model for program representation so that it supports high-level program manipulation. Based on this model, we designed and evaluated a direct manipulation visual interface for program restructuring. This interface provides information about the program that was not immediately accessible in the code and allows to manipulate programs without rewriting code. By providing a visual and textual representation of an automatically computed program optimization that is easily modifiable and reusable by the developer, we create a sort of human-machine partnership where the developer can better take advantage of the power of the machine. An empirical study of developers using

this tool showed the value of program manipulation tools based on the instrumental interaction paradigm. This work illustrates how the combination of our conceptual approaches, namely instrumental interaction and human-computer partnership, can benefit extreme users such as developers of parallel programs.

Finally, we reviewed statistical methods for the analysis of user-elicited gestural vocabularies [24]. We showed that measures currently used to assess agreement between participants of a gesture elicitation study are problematic. We discussed the problem of chance agreement and showed how it can bias results. We reviewed chance-corrected agreement coefficients that are routinely used in inter-reliability studies and showed how to apply them to gesture elicitation studies. We also discussed how to compute interval estimates for these coefficients and how to use them for statistical inference.

7.2. Partnerships

Participants: Wendy Mackay, Jessalyn Alvina, Ghita Jalal, Joseph Malloch, Nolwenn Maudet.

ExSitu is interested in designing effective human-computer partnerships, in which expert users control their interaction with technology. Rather than treating the human users as the 'input' to a computer algorithm, we explore human-centered machine learning, where the goal is to use machine learning and other techniques to increase human capabilities. Much of human-computer interaction research focuses on measuring and improving productivity: our specific goal is to create what we call 'co-adaptive sytstems' that are discoverable, appropriable and expressive for the user. *Interactive program restructuring* [28] offers a concrete example, where expert programmers interact with dynamic visualisations of parallel programs to better understand and organize their code. Similarly, tools such as *Color Partner* generate color suggestions based the users input, helping the user guide their discovery of new color possibilities, and *Linkify* helps users create rules to define how visual properties should change under different user contexts (see Jalal's dissertation).

We hosted the 30-person *ERC CREATIV* workshop in Paris, to explore our concepts of *Co-adaptive Systems* (including human-centered machine learning); and *Instrumental Interaction* (including substrates) with prominent researchers from Stanford University, New York University, University of Aarhus, Goldsmiths College, University of Toulouse, IRCAM, University of British Columbia, UC San Diego, and UC Berkeley. Our long-term, admittedly ambitious, goal is to create a unified theory of interaction grounded in how people interact with the world. Our principles of co-adaptive systems and instrumental interaction offer a generative approach for supporting creative activities, from early exploration to implementation. The workshop launched several research projects that are currently in progress or will be published in 2017.

Human-Centred Machine Learning:

We begin by challenging some of the standard assumptions surrounding Machine Learning, clearly one of the most important and successful techniques in contemporary computer science. It involves the statistical inference of models (such as classifiers) from data. However, all too often, the focus is on impersonal algorithms that work autonomously on passively collected data, rather than on dynamic algorithms that progressively reveal their progress to support human users. We collaborated on a workshop at the CHI 2017 conference, entitled "Human-centred Machine Learning" [15] with colleagues from Ircam, Goldsmiths College, and Microsoft Research. We seek a different understanding of the 'human-in-the-loop', where the focus is less on the human user as input to an algorithm, but rather as an algorithm in service of a human user. Examining machine learning from a human-centred perspective includes explicitly recognising human work in the creation of these algorithms, as well as the situated use these algorithms by human work practices. A human-centred understanding of machine learning in human context can lead not only to more usable machine learning tools, but to new ways of framing learning computationally.

Supporting Expressivity:

We helped organize and participated in a workshop at CHI 2017 *Human Computer Interaction meets Computer Music* [27], where we described the results of the MIDWAY Equipe Associé project (with McGill University, Ex-Situ and the *MINT* EP at Inria, Lille.) We presented results of our extensive research with contemporary music composers, in particular our strategy for developing 'co-adaptive instruments'. This involves a paradigm

shift, where the goal of the technology is not necessarily the accuracy of a particular result, but rather, the human user's ability to express themselves through the technology.

We also explored the idea of rethinking the use of machine learning to support human-computer partnerships for everyday interaction. We built on gesture-typing, which offers users an efficient, easy-to-learn, and error-tolerant technique for producing typed text on a soft keyboard. Our focus was not on improving recognition accuracy, which we take as a given, but rather on how to make gesture-typed output more expressive. Experiment 1 demonstrated that users vary word gestures according to instructions (accurately, quickly or creatively) as well as specific characteristics of each word, including length, angle, and letter repetition. We show that users produce highly divergent gestures, with three easily detectable characteristics: curviness, size, and speed. We created the Expressive Keyboard [10] which maps these characteristics to color variations, thus allowing users to control both the content and the color of gesture-typed words (Figure 4). Experiment 2 demonstrates that users can successfully control their gestures to produce the desired colored output, and find it easier to react to visual feedback than explicitly controlling the characteristics of each gesture. Expressive keyboards can map gestural input to any of a variety of output characteristics, such as personalized handwriting and dynamic emoticons, to let users transform gesture variation into expressivity, without sacrificing accuracy.
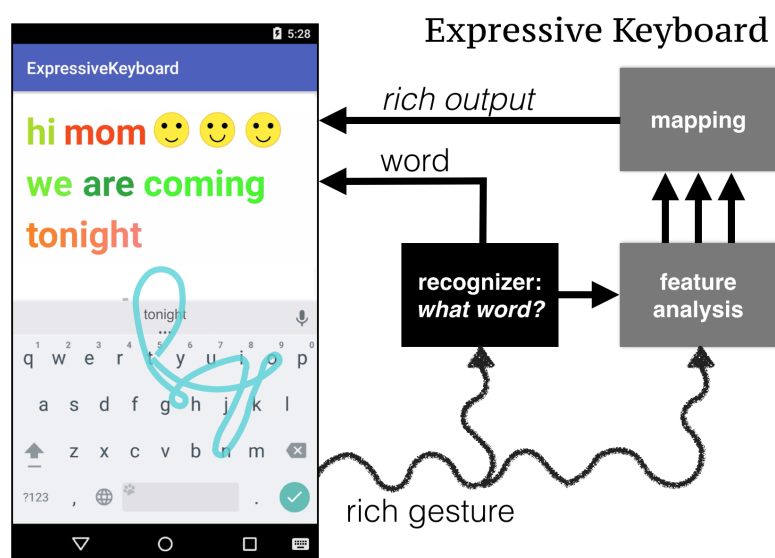
Figure 4. Expressive Keyboards produce accurate words, but also let users control multiple expressive output properties.

7.3. Creativity

Participants: Sarah Fdili Alaoui, Michel Beaudouin-Lafon, Ghita Jalal, Wendy Mackay, Joseph Malloch, Nolwenn Maudet, Michael Wessely, Theophanis Tsandilas.

ExSitu is interested in understanding the work practices of creative professionals, particularly artists, designers, and scientists, who push the limits of interactive technology.

We explore how concepts of substrate and co-adaptation can change how we design interactive technology for supporting creativity. Co-adaptation is the phenomenon in which users both adapt their behavior to the system's constraints, and appropriate the system for their own needs. We explore these concepts using

participatory design studies in creative contexts with expert and non-expert users. We study structuring layouts for graphic designers, sketching movement for choreographers, expressive movements for dancers and further explore expressive gesture of non-experts on mobile devices and possible interactions on hybrid stretchable interfaces. These studies require a multi-disciplinary design team that works closely with users throughout the design process. We create situations that cause users to reflect deeply about their activities in context and work with them to articulate the design problem. The experiments, prototypes and systems that we developed and deployed are illustrated below:

Graphic design: Our studies of the creative design practices of professional graphic designers show that designers appropriate visual properties of existing tools to create their own personal 'instruments'. Unfortunately, most professional design tools make this difficult: At best, they provide only indirect access, through property sheets or dialog boxes, to visual properties, such as color and style, rather treating them as as independent interactive objects. We developed a number of composition tools that demonstrate how to explicitly reify visual properties, using the concept of co-adaptive instruments. Ghita Jalal successfully defended her doctoral dissertation on this topic (see [9]).

We also examined artists' and designers' practices as they manipulate color and create layouts in their projects. We found that artists and designers select colors from personal representations. They manipulate color in the context of its surrounding graphical elements, and combine it with other visual properties such as texture. As they create their layouts, designers establish links among visual properties such as size, position, and layering of graphical elements. They define rules for how these properties change in space, across instances of the same composition, or in time, across related compositions. We also found that designers prefer tools that provide direct access to visual properties.

Choreography: We are interested in designing choreographic support tools because choreographers rarely have access to interactive tools that are designed specifically to support their creative process [13]. In order to design for such a technology, we interviewed six contemporary choreographers about their creative practice. We found that even though each process is unique, choreographers represent their ideas by applying a set of operations onto choreographic objects. Throughout different creative phases, choreographers compose by shifting among various degrees of specificity and vary their focal points from dancers to stage, to interaction, to the whole piece. Based on our findings, we presented a framework for articulating the higher-level patterns that emerge from these complex and idiosyncratic processes. We then articulated the resulting implications for the design of interactive tools to support the choreographic practice.

On generating choreographic ideas, we developed the Choreographer's Workbench, a full-body interactive system that aims to help choreographers explore and design dance movements during the ideation phase by creating a link between past recorded movement ideas and revealing their underlying relationships. The system explores how to increase the discoverability and appropriateness of movement ideas via feedforward visualization of movement characteristics.

We collaborated with the N+1 theater group on the "Grande Vitrine" art and science project, an interactive installation that takes place during the month of Christmas. It consisted of a virtual animated character with whom participants interact and a physical kinetic sculpture whose motions are triggered by participant interaction (Figure 5). The participant were expected to perform full-body movements and figure out the correct one that will help the animated character escape from the virtual screen into the physical motorized display. The installation tested the concept of "shaping" from experimental psychology where the participant is guided to make "successive approximations" in arriving at the correct gesture. It was installed at the theater of Évry, that has a display on the shopping mall in Évry for the entire month of December.

Finally, we collaborated with Simon Fraser University on an interactive installation called still, moving. The installation created a sonic experience that heightens self-awareness of our micro-movements in stillness. Sound created an intimate envelope that nurtures self-reflection and the experience of inward sensations. In still, moving, the audience was equipped with two Myo Armbands that capture their movements as well as their muscular activity. The physiological signals such as muscle tension and subtle accelerations were analyzed and mapped to a sound environment in order to increase perception of the inner self. The design of the relationship between movement and sound was evolving along the interaction, shifting the soundscape

Figure 5. Passersby interact with the animated Père Noël, first mirrors their behavior and then shapes it.

from reflective to challenging, guiding the audience in an exploration of novel and gradual relationship to weight and understanding of the complexity of the silent body.

Everyday creativity:

Finally, for non expert users, we developed an inexpensive method for fabricating *Stretchis*, highly stretchable interfaces that combine sensing and displaying capabilities [22]. This method enables designers and casual makers to embed transparent conductors and electroluminescence displays in stretchable PDMS substrates (Figure 6). We showed how to prototype stretchable user interfaces for a range of application scenarios by using standard design software and screen-printing techniques. Despite the use of inexpensive equipment, our results demonstrate that we can produce durable and highly stretchable sensors and displays that remain functional under strain levels of more than 100%.
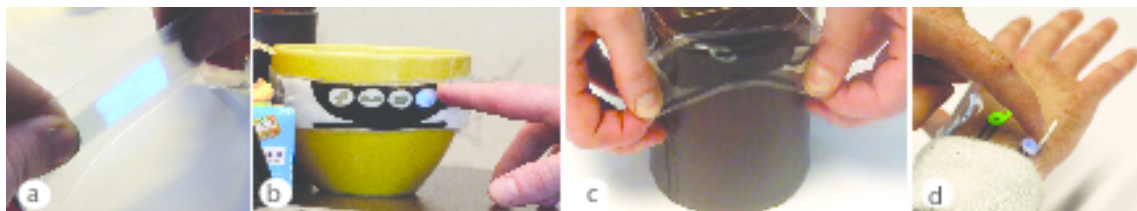
Figure 6. Stretchis are highly stretchable user interfaces that include touch and proximity sensors and electroluminescent displays (a). Stretchis are transparent (b); can be stretched to fit to the geometry of different physical objects (c); and can act as on-skin user interfaces (d).

7.4. Collaboration

Participants: Michel Beaudouin-Lafon, Cédric Fleury, Wendy Mackay, Can Liu, Ignacio Avellino Martinez.

ExSitu is interested in exploring new ways to support collaborative interaction, especially within and across large interactive spaces such as those of the Digiscope network (http://digiscope.fr/).

We studied how wall-sized displays support small groups of users working together on large amounts of data. We conducted observational studies showing that users adopt a range of collaboration styles, from loosely to closely coupled and that shared interaction techniques, in which multiple users perform a command collaboratively, support co-located collaborative work. In order to test the effect of such shared interaction techniques, we operationalize five collaborative situations with increasing levels of coupling in a data manipulation task [18]. The results show the benefits of shared interaction for close collaboration: it encourages collaborative manipulation, it is more efficient and preferred by users, and it reduces physical navigation and fatigue. We also identified the time costs caused by disruption and communication in loose collaboration and analyzed the trade-offs between parallelization and close collaboration. Altogether, these findings can inform the design of shared interaction techniques to support collaboration on wall-sized displays.

We are also interested in how to help teams of novice crafters prototype physical objects. To this end, we conducted a study [12] framed around two all-day design charrettes where novices performed a complete design process: ideation sketching, concept development and presentation, fabrication planning documentation and collaborative fabrication of hand-crafted prototypes. This structure allowed us to control key aspects of the design process while collecting rich data about creative tasks, including sketches on paper, physical models, and videos of collaboration discussions. Participants used a variety of drawing techniques to convey 3D concepts. They also extensively manipulated physical materials, such as paper, foam, and cardboard, both to support concept exploration and communication with design partners. Based on these observations, we proposed design guidelines for CAD tools targeted at novice crafters.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Grants with Industry

MultiHub (Microsoft donation, 2015-2016) – ExSitu was one of the ten academic institutions world wide awarded a hardware and monetary grant by Microsoft Research as part of its request for proposal to expand the potential applications of the Surface Hub across all aspects of society (http://research.microsoft.com/en-us/projects/surface-hub/). The goal of the MultiHub project is to enable interaction in the large, where groups of experts can interact with rich content and complex data while collaborating both locally and remotely in interactive, multi-surface environments. ExSitu was awarded two 55" Surface Hubs and $19,000 in cash.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. *MultiVis – Novel Interaction Models for Multi-surface Visualization*

       Type: Ph.D. grant

       Funding: DigiCosme Labex

       Duration: 2014-2017

       Coordinator: James Eagan (Institut Mines Telecom)

       Partners: Univ. Paris-Sud, Inria, CNRS, Institut Mines-Telecom

       Inria contact: Michel Beaudouin-Lafon

Abstract: The goal of this project is to design, evaluate, and implement novel interaction models that help users appropriate multiple computational surfaces in the sense-making process. Our initial approach is to operationalize and extend the instrumental interaction model to specifically accommodate the specific needs of the sense-making process for information visualization. This project funds Marc-Emmanuel Perrin, a joint PhD student between the VIA group at Institut Mines-Telecom and ExSitu.

9.1.2. MoveIT – Modeling the Speed/Accuracy Trade-Off of Human Aimed Movement with the Tools of Information Theory

Type: Ph.D. grant

Funding: DigiCosme Labex

Duration: 2015-2018

Coordinator: Olivier Rioul (Institut Mines Telecom)

Partners: Univ. Paris-Sud, Inria, CNRS, Institut Mines-Telecom

Inria contact: Michel Beaudouin-Lafon

Abstract: The goal of this project is to conduct fundamental studies of aimed movements based on information theory. The project studies the interaction phenomena involved in pointing, in order to discover novel, more effective pointing techniques. This project funds Wanyu Liu, a joint Ph.D. student between the COMELEC and VIA groups at Institut Mines Telecom and ExSitu.

9.1.3. SensoMotorCVE – Sensor-motor Interface for Collaborative Virtual Environments with Heterogeneous Devices: Application to Industrial Design

Type: Ph.D. grant

Funding: DigiCosme Labex

Duration: 2014-2017

Coordinator: Patrick Bourdot (LIMSI-CNRS)

Partners: Univ. Paris-Sud, Inria, CNRS

Inria contact: Cédric Fleury

Abstract: In the context of collaborative virtual environments, the goal of this project is to develop a sensorimotor interface model for CAD data manipulation that supports heterogeneous interactive systems such as wall-sized displays or immersive virtual reality rooms. This project funds Yujiro Okuya, a joint Ph.D. student between the VENISE group at LIMSI and ExSitu.

9.1.4. La Grande Vitrine des Choses

Type: Art-science grant

Funding: IDEX Paris-Saclay

Duration: 2015-2016

Coordinators: Michel Beaudouin-Lafon & Wendy Mackay

Partners: Univ. Paris-Sud, Inria, CNRS, Theater group $n + 1$

Abstract: Art-science project funded by "La Diagonale Paris-Saclay" to create, in collaboration with the theather group "n+1", an interactive store front in the form of an advent calendar, where users must discover which gestures to perform in order make an animated character open the next window. This installation raises the question of who is controlling whom: Participants think that their gestures directly control the character, but the system actually uses shaping techniques from experimental psychology that encourage users to make successive approximations to the correct gesture. The installation was demonstrated at the Fête de la Science in October, 2016, and was shown during the month of December, 2016 in the Evry shopping mall, next to the Agora Theater. It will also be shown in the Curiositas festival in Gif-sur-Yvette in May, 2017.

9.2. National Initiatives

9.2.1. Investissements d'Avenir

9.2.1.1. Digiscope - Collaborative Interaction with Complex Data and Computation

> Type: EQUIPEX (Equipement d'Excellence)
>
> Duration: 2011-2020
>
> Coordinator: Michel Beaudouin-Lafon
>
> Partners: FCS Paris-Saclay (coordinator), Université Paris-Sud, CNRS, CEA, Inria, Institut Mines-Telecom, Ecole Centrale Paris, Université Versailles - Saint-Quentin, ENS Cachan, Maison de la Simulation
>
> Overall budget: 22.5 Meuros, including 6.7 Meuros public funding from ANR
>
> Abstract: The goal of the project is to create ten high-end interactive rooms interconnected by high-speed networks and audio-video facilities to support remote collaboration across interactive visualization environments. The equipment will be open to outside users and targets four main application areas: scientific discovery, product lifetime management, decision support for crisis management, and education and training. Digiscope includes the existing WILD room, and funded the WILDER room. ExSitu contributes its expertise in the design and evaluation of advanced interaction techniques and the development of distributed software architectures for interactive systems. At the end of 2016, all ten rooms are operational, and the telepresence network is being deployed.

9.2.2. Institut Universitaire de France

9.2.2.1. The Instrumental Paradigm

> Type: IUF senior fellowship
>
> Duration: 2011-2016
>
> Principal investigator: Michel Beaudouin-Lafon
>
> Abstract: Tools or instruments are a natural way to interact with the real world, and can serve as a powerful metaphor to interact with on-line information. An instrument reifies interaction: it turns an interaction into a meaningful object for users, designers and developers. We envision a future where large, monolithic and closed applications are replaced by a rich ecology of instruments and information containers that can interoperate, giving users the power to shape their own environments. Our work on multisurface interaction [2] and Webstrates [5] illustrate this approach.

9.3. European Initiatives

9.3.1. European Research Council (ERC)

9.3.1.1. Creating Human-Computer Partnerships

> Program: ERC Advanced Grant
>
> Project acronym: CREATIV
>
> Project title: Creating Human-Computer Partnerships
>
> Duration: mois année début - mois année fin
>
> Coordinator: Wendy Mackay

Abstract: CREATIV explores how the concept of co-adaptation can revolutionize the design and use of interactive software. Co-adaptation is the parallel phenomenon in which users both adapt their behavior to the system's constraints, learning its power and idiosyncrasies, and appropriate the system for their own needs, often using it in ways unintended by the system designer. A key insight in designing for co-adaptation is that we can encapsulate interactions and treat them as first class objects, called interaction instruments This lets us focus on the specific characteristics of how human users express their intentions, both learning from and controlling the system. By making instruments co-adaptive, we can radically change how people use interactive systems, providing incrementally learnable paths that offer users greater expressive power and mastery of their technology. The initial goal of the CREATIV project is to fundamentally improve the learning and expressive capabilities of advanced users of creative software, offering significantly enhanced methods for expressing and exploring their ideas. The ultimate goal is to radically transform interactive systems for everyone by creating a powerful and flexible partnership between human users and interactive technology.

9.3.1.2. Unified Principles of Interaction

Program: ERC Advanced Grant
Project acronym: ONE
Project title: Unified Principles of Interaction
Duration: October 2016 - September 2020
Coordinator: Michel Beaudouin-Lafon
Abstract: The goal of ONE is to fundamentally re-think the basic principles and conceptual model of interactive systems to empower users by letting them appropriate their digital environment. The project addresses this challenge through three interleaved strands: empirical studies to better understand interaction in both the physical and digital worlds, theoretical work to create a conceptual model of interaction and interactive systems, and prototype development to test these principles and concepts in the lab and in the field. Drawing inspiration from physics, biology and psychology, the conceptual model combines *substrates* to manage digital information at various levels of abstraction and representation, *instruments* to manipulate substrates, and *environments* to organize substrates and instruments into digital workspaces.

9.4. International Initiatives

9.4.1. Inria Associate Teams Not Involved in an Inria International Labs

9.4.1.1. DECibel

Title: Discover, Express, Create – Interaction Technologies For Creative Collaboration
International Partner (Institution - Laboratory - Researcher):

University of California Berkeley (United States) - Electrical and Computer Engineering, Center for Magnetic Resonance Research - Bjoern Hartmann

Start year: 2016

The DECibel associated team includes Inria's ExSitu and the CITRIS Connected Communities Initiative (CCI) at UC Berkeley. ExSitu explores extreme interaction, working with creative professionals and scientists who push the limits of technology to develop novel interactive technologies that offer new strategies for creative exploration. ExSitu's research activities include: developing underlying theory (co-adaptive instruments and substrates), conducting empirical studies (participatory design with creative professionals), and implementing interactive systems (creativity support tools). The CITRIS Connected Communities Initiative investigates collaborative discovery and design through new technologies that enhance education, creative work, and public engagement. It develops interactive tools, techniques and materials for the rapid design and prototyping of novel interactive products, expertise sharing among designers, and citizen science investigations. DECibel will combine the strengths of these two groups to to investigate novel tools and technologies that support Discovery, Expressivity, and Creativity.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- Joanna McGrenere, Professor at University of British Columbia, Vancouver, Canada, visited ExSitu for her entire sabbatical, from September, 2015 to July, 2016.
- Jim Hollan, Professor at University of California at San Diego (UCSD), visited from April to June, 2016.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. Member of the Organizing Committees

- IHM 2016, *Conférence Francophone d'Interaction Homme-Machine*, Doctoral Consortium Co-Chair: Michel Beaudouin-Lafon, Jury: Wendy Mackay

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

- NIME 2016, *New Interfaces for Musical Expression*, Conference Workshop Chair: Joseph Malloch
- ACM/CHI 2016, *Music and HCI*, Conference Workshop Co-Organisers: Wendy Mackay, Joseph Malloch
- ACM/CHI 2016, *Human-Centred Machine Learning*, Conference Workshop Co-Organiser: Wendy Mackay
- ERC CREATIV Workshop *CREATIV*, Conference Workshop Organiser: Wendy Mackay
- MOCO 2016, *International Symposium on Movement and Computing*, Conference Steering Committee member: Sarah Fdili Alaoui

10.1.2.2. Member of the Conference Program Committees

- IEEE VR 2016, *Virtual Reality Conference*: Cédric Fleury
- GI 2016, *Graphics Interface Conference*: Cédric Fleury
- GRAPP 2016, *Conference on Computer Graphics Theory and Applications*: Cédric Fleury
- 3DCVE 2016, *IEEE VR Workshop on Collaborative Virtual Environments*: Cédric Fleury
- CHI 2017, *ACM CHI Conference on Human Factors in Computing Systems*: Michel Beaudouin-Lafon, Wendy Mackay
- CHI 2017, *ACM UIST User Interface Software and Technology*: Wendy Mackay
- MOCO 2016, *International Symposium on Movement and Computing*: Sarah Fdili Alaoui

10.1.2.3. Reviewer

- IJHCS *International Journal on Human-Computer Interaction*: Wendy Mackay
- ACM CSCW 2016 *ACM Conference on Computer-Supported Cooperative Work*: Wendy Mackay
- ACM CHI 2016 *ACM Conference on Human Factors in Computing Systems*: Marianela Ciolfi Felice, Germán Leiva, Oleksandr Zinenko, Ignacio Avellino, Nolwenn Maudet, Joseph Malloch, Theophanis Tsandilas, Michel Beaudouin-Lafon, Wendy Mackay, Sarah Fdili Alaoui
- ACM UIST 2016 *ACM Symposium on User Interface Software and Technology*: Joseph Malloch, Theophanis Tsandilas, Nolwenn Maudet, Ignacio Avellino, Jessalyn Alvina, Cédric Fleury, Michel Beaudouin-Lafon, Wendy Mackay

- UBICOMP 2016, *ACM International Joint Conference on Pervasive and Ubiquitous Computing*: Theophanis Tsandilas
- SIGGRAPH Asia 2016: Theophanis Tsandilas
- ACM MobileHCI 2016, *ACM Conference on Human-Computer Interaction with Mobile Devices and Services*: Jessalyn Alvina
- IEEE 3DUI 2016, *Symposium on 3D User Interfaces*: Cédric Fleury
- IEEE Visualization: Michel Beaudouin-Lafon
- IHM 2016, *Conférence Francophone d'Interaction Homme-Machine*: Cédric Fleury
- WAMCA 2016, *Workshop on Applications for Multi-Core Architectures*: Oleksandr Zinenko
- IMPACT 2016, *International Workshop on Polyhedral Compilation Techniques*: Oleksandr Zinenko
- TEI 2016, *International Conference on Tangible, Embedded and Embodied Interactions*: Sarah Fdili Alaoui
- DIS 2016, *Designing Interactive Systems*: Sarah Fdili Alaoui

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Editor for the Human-Computer Interaction area of the new ACM Books Series (published with Morgan & Claypool Publishers): Michel Beaudouin-Lafon (2013-)
- CACM, *Communications of the ACM Web Editorial Board*, ACM: Wendy Mackay (2008-)
- CACM *Communications of the ACM New Publications Board*, ACM: Wendy Mackay (2015-)
- TOCHI, *Transactions on Computer Human Interaction*, ACM: Michel Beaudouin-Lafon (2009-), Wendy Mackay (2016-)
- JIPS, *Journal d'Interaction Personne-Système*, AFIHM: Michel Beaudouin-Lafon (2009-)
- IJHCS, *International Journal of Human-Computer Study*, Elsevier: Michel Beaudouin-Lafon (Member of the Advisory Board, 2009-)
- JCSCW, *Journal of Computer Supported Cooperative Work*, Springer: Michel Beaudouin-Lafon (Member of the Advisory Board, 2010-)

10.1.3.2. Reviewer - Reviewing Activities

- Computers and Graphics: Theophanis Tsandilas
- TVCG, *Transactions on Visualization and Computer Graphics*, IEEE: Cédric Fleury
- UAIS, Universal Access in the Information Society: Michel Beaudouin-Lafon
- JAISE, Journal of Ambient Intelligence and Smart Environments: Michel Beaudouin-Lafon

10.1.4. Invited Talks

- Beyond Lab: Entre L'Homme et la Machine : Le Défi des Nouvelles Interfaces "*Interaction en Grand*", 8 February 2016: Michel Beaudouin-Lafon
- Beyond Lab: Entre L'Homme et la Machine : Le Défi des Nouvelles Interfaces "*Apprentissage Incrémental de Gestes*", 8 February 2016: Wendy Mackay
- Beyond Lab: Entre L'Homme et la Machine : Le Défi des Nouvelles Interfaces "*Movement Connoisseurship for User eXperience in Digital Art and HCI*", 8 February 2016: Sarah Fdili Alaoui
- Beyond Lab: Entre L'Homme et la Machine : Le Défi des Nouvelles Interfaces "*Papier Interactif pour la Créativité*", 8 February 2016: Theophanis Tsandilas
- FAST (Fusing Audio and Semantic Techologies for Intelligent Music Production and Consumption) Workshop Paris "*Creative Musical Expression with Human-Computer Partnerships*", 15 February 2016: Wendy Mackay

- ENS Cachan Semainaire d'Informatique "*Co-Adaptive Instruments. Can we reinvent the graphical user interface*", 2 March 2016: Wendy Mackay
- Café Techno – Inria Alumni & NUMA, "*Quelles interactions avec le monde numérique de demain ?*", March 2016: Michel Beaudouin-Lafon
- Les Outils du Design, ENSCI, "*Instrumental Interaction*", April 2016: Michel Beaudouin-Lafon
- École Thématique Maths-Info-HPC, Saint-Germain-au-Mont-d'Or, "*Polyhedral Code Transformations via Interactive Visualizations*", 10 May 2016: Oleksandr Zinenko
- Journée d'étude Scénariser la danse par le numérique, Université de Strasbourg "*Étude du mouvement expressif dans les interactions incorporée pour la danse*", 18 May 2016: Sarah Fdili Alaoui
- May Residency of the Moving Stories Project, Simon Fraser University "*Embodied Interaction for dance and tools for supporting choreography*", 28 May 2016: Sarah Fdili Alaoui
- ERC CREATIV Workshop "*Creating Human-Computer Partnerships*", 13 June 2016: Wendy Mackay
- ICCC International Conference on Computational Creativity 2016: ERC Comics Invited Address "*Creating CoAdaptive partnerships*", 29 June 2016: Wendy Mackay
- ENS Paris, "*Polyhedral program transformations via interactive visualization*", 1 June 2016: Oleksandr Zinenko
- ARC Raymond Hains, Ecole Européenne Supérieure Beaux Arts de Rennes, "*Construire et déconstruire les logiciels du design*", 11 octobre 2016: Nolwenn Maudet
- Journée MENESR-EHESS: Coding and computer science in our schools: new ways of learning? "*Entre les enfants et les ordinateurs: Human Computer Interaction*", 11 october 2016: Wendy Mackay
- CM Lab, Tokyo Tech, "*Designing Design Tools*", 20 october 2016: Nolwenn Maudet
- Google Japan, Tokyo, Japan, "*Expressive Keyboards: Expressive Keyboards: Enriching Gesture-Typing on Mobile Devices*", 20 October 2016: Jessalyn Alvina
- Rencontre Inria – Industrie Lille "*Interactions avec les objets ! et services numeriques*", 25 November 2016: Wendy Mackay
- TEDx Saclay, "*Dancing in digital*", 30 November 2016: Sarah Fdili Alaoui
- Rencontres Lurs: "*Droit à l'Erreur*", 25 August 2016: Wendy Mackay
- UniThé ou Café "*Créer des Partenariats Homme-Machine*", 17 November 2016: Wendy Mackay
- Infomuse, Casa Paganini "*Embodied Interaction for dance and tools for supporting choreograph*", 14 Decembre 2016: Sarah Fdili Alaoui

10.1.5. Scientific Expertise

- European Research Council, Panel member for Starting Grants: Michel Beaudouin-Lafon
- CNRS Mission pour l'Interdisciplinarité, Panel member for the call "Sciences Sociales et Cognitives des Comportements Collectifs": Michel Beaudouin-Lafon
- ACM/CHI Best Paper Award Committee "Best Paper Award Committee": Wendy Mackay
- ACM/SIGCHI "Lifetime Service Award Committee": Wendy Mackay

10.1.6. Research Administration

- Computer Science Department, Université Paris-Saclay: Wendy Mackay (member of steering committee)
- CNRS, Conseil Scientifique de l'Institut INS2I: Michel Beaudouin-Lafon (member)
- Telecom ParisTech: Michel Beaudouin-Lafon (member of research committee)
- IRCAM: Michel Beaudouin-Lafon (member of scientific committee)

- "Institut de la Société Numérique", IDEX Laboratory of Université Paris-Saclay: Michel Beaudouin-Lafon (member of steering committee)

- "Conseil de Laboratoire", LRI: Wendy Mackay, Cédric Fleury (members)

- "Conseil Scientifique", LRI: Michel Beaudouin-Lafon (member)

- "Commission Locaux", LRI: Theophanis Tsandilas (member)

- CCSU, "Commission Consultative de Spécialistes de l'Université", Computer Science department: Michel Beaudouin-Lafon, Wendy Mackay (members)

- COERLE "Comité Operationnel d'Evaluation des Risques Légaux et Ethiques", Inria: Wendy Mackay

- CERNI "Comité d'Ethique pour les Recherches Non Interventionnelles", Université Paris-Sud: Wendy Mackay

- "Commission de Développement Technologique", Inria: Theophanis Tsanidlas (member), until April 2016

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Interaction & HCID Masters: Wendy Mackay, *Career Seminar* 21 hrs, M2, Univ. Paris-Sud

Interaction & HCID Masters: Michel Beaudouin-Lafon, *Fundamentals of Human-Computer Interaction*, 21 hrs, M1/M2, Univ. Paris-Sud

Interaction & HCID Masters: Michel Beaudouin-Lafon, Wendy Mackay *Fundamentals of Situated Interaction*, 21 hrs, M1/M2, Univ. Paris-Sud

Interaction & HCID Masters: Wendy Mackay, *Design of Interactive Systems*, 21 hrs, M1/M2, Univ. Paris-Sud

Interaction & HCID Masters: Wendy Mackay, *Advanced Design of Interactive Systems*, 21 hrs, M1/M2, Univ. Paris-Sud

Interaction & HCID Masters: Wendy Mackay, *Business Development Lab* 21 hrs, M1, Univ. Paris-Sud

Interaction & HCID Masters: Michel Beaudouin-Lafon & Cédric Fleury, *Groupware and Collaborative Interaction*, 31.5 hrs, M1/M2, Univ. Paris-Sud

Interaction & HCID Masters: Sarah Fdili Alaoui, *Creative Design*, 21 hrs, M1/M2, Univ. Paris-Sud

Interaction & HCID Masters: Sarah Fdili Alaoui, *Design Project*, 21 hrs, M1/M2, Univ. Paris-Sud

Interaction & HCID Masters: Sarah Fdili Alaoui, *Gesture and Mobile Interaction*, 4 hrs, M1/M2, Univ. Paris-Sud

HCID Masters: Wendy Mackay & Michel Beaudouin-Lafon, *Fundamentals of Situated Interaction*, 21 hrs, M2, Univ. Paris-Sud

Interaction & HCID Masters: Sarah Fdili Alaoui, *Programmation des interfaces interactives avancées*, 15 hrs, L3, Univ. Paris-Sud

Polytech 5th year Apprentices: Sarah Fdili Alaoui, *Informatique Graphique et visualisation*, 12 hrs, M2, Univ. Paris-Sud

Polytech 5th year: Cédric Fleury, *Réalité Virtuelle et Interaction*, 48 hrs, M2, Univ. Paris-Sud

Polytech 3st year: Cédric Fleury, *Interaction Homme-Machine*, 18 hrs, M1, Univ. Paris-Sud

Polytech 3st year: Cédric Fleury, *Projet Java-Graphique-IHM*, 24 hrs, M1, Univ. Paris-Sud

Polytech 1st year: Cédric Fleury, *Introduction à l'Informatique*, 71 hrs, L1, Univ. Paris-Sud

M1 Informatique: Theophanis Tsandilas, *Évaluation des Interfaces*, 2 hrs, M1, Univ. Paris-Sud

10.2.2. Supervision

PhD students

PhD: Oleksandr Zinenko, *Interactive Program Restructuring*, Université Paris-Saclay, 25 November 2016. Advisors: Stéphane Huot (Inria Lille) & Cédric Bastoul (Université de Strasbourg)

PhD: Ghita Jalal, *Supporting visual composition with reified interactive properties*, Université Paris-Saclay, 16 December 2016. Advisor: Wendy Mackay

PhD in progress: Ignacio Avellino, *Remote Collaboration in Large Interactive Spaces*, September 2014. Advisors: Michel Beaudouin-Lafon & Cédric Fleury

PhD in progress: Jessalyn Alvina, *Mobile Co-Adaptive Instruments*, September 2014. Advisor: Wendy Mackay

PhD in progress: Nolwenn Maudet, *Substrates and Co-Adaptive Instruments*, September 2014. Advisors: Wendy Mackay & Michel Beaudouin-Lafon

PhD in progress: Marc-Emmanuel Perrin, *Novel Interaction Models for Multi-surface Visualization*, September 2014. Advisors: Michel Beaudouin-Lafon & James Eagan (Télécom ParisTech)

PhD in progress: Marianela Ciolfi Felice, *Substrates and Co-adaptive Instruments for Creativity*, September 2015. Advisors: Wendy Mackay & Sarah Fdili Alaoui

PhD in progress: Carla Griggio, *Interactive Human-Machine Learning*, September 2015. Advisor: Wendy Mackay

PhD in progress: Germán Leiva, *Interaction-driven Software Development*, September 2015. Advisor: Michel Beaudouin-Lafon

PhD in progress: Wanyu Liu, *Modeling the speed-accuracy trade-off of pointing tasks using the tools of information theory*, October 2015. Advisors: Olivier Rioul (Institut Mines Telecom) & Michel Beaudouin-Lafon

PhD in progress: Yujiro Okuya, *Sensorimotor interface for Collaborative Virtual Environments based on heterogeneous interactive devices: application to industrial design*, October 2015. Advisors: Patrick Bourdot (LIMSI-CNRS) & Cédric Fleury

PhD in progress: Michael Wessely, *Sketching and Physical Prototyping for Creative Fabrication Design*, November 2015. Advisors: Theophanis Tsandilas & Wendy Mackay

PhD in progress: Stacy (Shu-Yuan) Hseuh, *Embodied design for Human-Computer Co-creation*, November 2015. Advisors: Wendy Mackay & Sarah Fdili Alaoui

PhD in progress: Philip Tchernavskij, *Towards Unified Principles of Interaction*, October 2016. Advisor: Michel Beaudouin-Lafon

Masters students

Brennan Jones, University of Calgary, "Physical and virtual avatars for telepresence and remote collaboration": Ignacio Avellino, Cédric Fleury & Michel Beaudouin-Lafon

Francesco Vitale, Université Paris-Saclay, "Field study of operating systems upgrades": Joanna McGrenere, Wendy Mackay

Dimitri Belopopsky, Université Paris-Saclay, "La Grande Vitrine des Choses": Michel Beaudouin-Lafon, Wendy Mackay, Joseph Malloch

Sruthi Coimbatore Viswanathan, Université Paris-Saclay, "Touch Keyboard with Haptic Feedback": Michel Beaudouin-Lafon

Iva Karabatic, Université Paris-Saclay, "Chord input on a digital whiteboard": Wendy Mackay, Michel Beaudouin-Lafon

Diana Lipcanu, Université Paris-Saclay, "The Inteaction Musuem": Wendy Mackay, Michel Beaudouin-Lafon

Isha Van Baar, Université Paris-Saclay, "The Inteaction Musuem": Wendy Mackay, Michel Beaudouin-Lafon

Nam Giang, Université Paris-Saclay, "Interactive Video Preview Playlist": Wendy Mackay, Carla Griggio

Midas Nouwens, Université Paris-Saclay, "Communication Appliances to Support Communication for Distributed Families": Wendy Mackay, Carla Griggio

Pierre Mahe, IRCAM, "EMBRACe Expressive Movement Body fRAmework and Computing": Sarah Fdili Alaoui

Niels Mourette, IRCAM, "The augmented body from within": Sarah Fdili Alaoui

Shu-Yuan Hsueh, IRCAM & ENSCII, "Dynamic Guides for Embodied Interactiob": Wendy Mackay, Sarah Fdili Alaoui

Philip Tchernavskij, University of Aarhus, "Information Substrates for Asymmetrical Collaboration": Wendy Mackay, Michel Beaudouin-Lafon

10.2.3. Juries

Mathieu Le Goc, Inria Saclay, December 2016 (Advisor: Jean-Daniel Fekete): Michel Beaudouin-Lafon, President.

Alex Goguey, Inria Lille, October 2016 (Advisor: Géry Casiez): Michel Beaudouin-Lafon, reviewer.

Andéol Evain, IRISA, December 2016 (Advisor: Anatole Lécuyer): Michel Beaudouin-Lafon, reviewer.

Emiline Brulé, Télécom Paris, 6 June 2016 (Advisor: Annie Gentes): Wendy Mackay, reviewer.

10.3. Popularization

Organisation d'une Journée d'étude, 15 April 2016 @ ENSCI *La fabrique des outils, concevoir et pratiquer les logiciels de design*: Nolwenn Maudet

BeyondLab, 10 February 2016 *Meet-up sur les nouvelles interfaces*: Sarah Fdili Alaoui, Michel Beaudouin-Lafon, Wendy Mackay, and Theophanis Tsandilas

Inria Rencontres Industrielles 2016, "*Expressive Keyboards: Enriching Gesture-Typing on Mobile Devices*", 25 November 2016: Jessalyn Alvina, Wendy Mackay

Inria Rencontres Industrielles 2016, "*StickyLines: Reifying Alignment and Distribution of Graphical Objects*", 25 November 2016: Michel Beaudouin-Lafon, Wendy Mackay

ACM UIST Demonstration, "*Expressive Keyboards*", Jessalyn Alvina

ACM UIST Demonstration, "*StickyLines*", Marianela Ciolfi

ACM UIST Demonstration, "*Stretchis*", Michael Wessely

ACM UIST Demonstration, "*UIST Video Playlist*", Carla Griggio

ACM "*HCI Pioneers*", Michel Beaudouin-Lafon, Wendy Mackay

11. Bibliography

Major publications by the team in recent years

[1] I. AVELLINO, C. FLEURY, M. BEAUDOUIN-LAFON.*Accuracy of Deictic Gestures to Support Telepresence on Wall-sized Displays*, in "CHI '15 Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems", Seoul, South Korea, ACM (editor), ACM Conference on Human Factors in Computing Systems (CHI), April 2015, p. 2392-2396 [*DOI :* 10.1145/2702123.2702448], https://hal.archives-ouvertes.fr/hal-01242676.

[2] M. Beaudouin-Lafon, O. Chapuis, J. Eagan, T. Gjerlufsen, S. Huot, C. Klokmose, W. E. Mackay, M. Nancel, E. Pietriga, C. Pillias, R. Primet, J. Wagner.*Multi-surface Interaction in the WILD Room*, in "IEEE Computer", 2012, vol. 45, n⁰ 4, p. 48–56, http://dx.doi.org/10.1109/MC.2012.110.

[3] E. Ghomi, S. Huot, O. Bau, M. Beaudouin-Lafon, W. E. Mackay.*Arpège: Learning Multitouch Chord Gestures Vocabularies*, in "Interactive Tabletops and Surfaces (ITS '13)", St Andrews, United Kingdom, ACM (editor), ACM, October 2013, http://hal.inria.fr/hal-00870173.

[4] G. Jalal, N. Maudet, W. E. Mackay.*Color Portraits: From Color Picking to Interacting with Color*, in "Proceedings of the ACM International Conference on Human Factors in Computing Systems, 2015", Seoul, South Korea, May 2015 [*DOI : 10.1145/2702123.2702173*], https://hal.archives-ouvertes.fr/hal-01226494.

[5] C. Klokmose, J. Eagan, S. Baader, W. Mackay, M. Beaudouin-Lafon.*Webstrates: Shareable Dynamic Media*, in "ACM Symposium on User Interface Software and Technology (UIST)", Charlotte, United States, ACM (editor), UIST '15 Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology, November 2015, p. 280-290 [*DOI : 10.1145/2807442.2807446*], https://hal.archives-ouvertes.fr/hal-01242672.

[6] C. Liu, O. Chapuis, M. Beaudouin-Lafon, É. Lecolinet, W. E. Mackay.*Effects of Display Size and Navigation Type on a Classification Task*, in "CHI '14", Toronto, Canada, CHI '14, ACM, April 2014, p. 4147-4156 [*DOI : 10.1145/2556288.2557020*], https://hal.archives-ouvertes.fr/hal-00957269.

[7] M. Nancel, E. Pietriga, O. Chapuis, M. Beaudouin-Lafon.*Mid-air Pointing on Ultra-Walls*, in "ACM Transactions on Computer-Human Interaction", August 2015, vol. 22, n⁰ 5, p. 21:1–21:62 [*DOI : 10.1145/2766448*], https://hal.inria.fr/hal-01184544.

[8] T. Tsandilas, A. Bezerianos, T. Jacob.*SketchSliders: Sketching Widgets for Visual Exploration on Wall Displays*, in "Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems", Seoul, South Korea, ACM, April 2015, p. 3255-3264 [*DOI : 10.1145/2702123.2702129*], https://hal.archives-ouvertes.fr/hal-01144312.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[9] G. Jalal.*Réification des propriétés visuelles pour la composition visuelle*, Univ Paris-Sud, Orsay, December 2016.

International Conferences with Proceedings

[10] J. Alvina, J. Malloch, W. Mackay.*Expressive Keyboards: Enriching Gesture-Typing on Mobile Devices*, in "Proceedings of the 29th ACM Symposium on User Interface Software and Technology (UIST 2016)", Tokyo, Japan, ACM (editor), ACM, October 2016, p. 583 - 593 [*DOI : 10.1145/2984511.2984560*], https://hal.inria.fr/hal-01437054.

[11] L. Bagnères, O. Zinenko, S. Huot, C. Bastoul.*Opening Polyhedral Compiler's Black Box*, in "CGO 2016 - 14th Annual IEEE/ACM International Symposium on Code Generation and Optimization", Barcelona, Spain, March 2016, https://hal.inria.fr/hal-01253322.

[12] A. BOUSSEAU, T. TSANDILAS, L. OEHLBERG, W. E. MACKAY.*How Novices Sketch and Prototype Hand-Fabricated Objects*, in "Conference on Human Factors in Computing Systems (CHI)", San Jose, United States, ACM, May 2016 [*DOI :* 10.1145/2858036.2858159], https://hal.inria.fr/hal-01272187.

[13] M. CIOLFI FELICE, S. FDILI ALAOUI, W. MACKAY.*How Do Choreographers Craft Dance? Designing for a Choreographer-Technology Partnership*, in "Proceedings of the 3rd International Symposium on Movement and Computing", Thessaloniki, Greece, ACM (editor), July 2016 [*DOI :* 10.1145/2948910.2948941], https://hal.inria.fr/hal-01438682.

[14] M. C. CIOLFI FELICE, N. MAUDET, W. E. MACKAY, M. BEAUDOUIN-LAFON.*Beyond Snapping: Persistent, Tweakable Alignment and Distribution with StickyLines*, in "UIST '16 Proceedings of the 29th Annual Symposium on User Interface Software and Technology", Tokyo, Japan, Proceedings of the 29th Annual Symposium on User Interface Software and Technology, October 2016 [*DOI :* 10.1145/2984511.2984577], https://hal.archives-ouvertes.fr/hal-01410171.

[15] M. GILLIES, R. FIEBRINK, A. TANAKA, B. CARAMIAUX, J. GARCIA, F. BEVILACQUA, A. HELOIR, F. NUNNARI, W. MACKAY, S. AMERSHI, B. LEE, N. D 'ALESSANDRO, J. TILMANNE, T. KULESZA.*Human-Centered Machine Learning*, in "Conference on Human Factors in Computing Systems (CHI 2016)", San Jose, California, United States, Proceedings of the 2016 CHI Conference - Extended Abstracts on Human Factors in Computing Systems, ACM, May 2016, p. 3558-3565 [*DOI :* 10.1145/2851581.2856492], https://hal.inria.fr/hal-01437057.

[16] C. F. GRIGGIO, N. GIANG, G. F. LEIVA, W. MACKAY.*The UIST Video Browser: creating shareable playlists of video previews*, in "UIST 2016 - 29th ACM Symposium on User Interface Software and Technology", Tokyo, Japan, ACM, October 2016, p. 59 - 60, Demonstration [*DOI :* 10.1145/2984751.2985703], https://hal.inria.fr/hal-01437052.

[17] C. N. KLOKMOSE, J. N. EAGAN, S. BAADER, W. MACKAY, M. N. BEAUDOUIN-LAFON.*Webstrates: demonstrating the potential of Shareable Dynamic Media*, in "19th ACM conference on Computer-Supported Cooperative Work and Social Computing", San Francisco, CA, United States, CSCW '16 Companion Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion, ACM, February 2016, p. 61 - 64 [*DOI :* 10.1145/2818052.2874325], https://hal.archives-ouvertes.fr/hal-01435919.

[18] C. LIU, O. CHAPUIS, M. BEAUDOUIN-LAFON, E. LECOLINET.*Shared Interaction on a Wall-Sized Display in a Data Manipulation Task*, in "Proceedings of the 34th international conference on Human factors in computing systems", San Jose, United States, ACM (editor), CHI '16, SigCHI and ACM, May 2016, p. 1-12 [*DOI :* 10.1145/2858036.2858039], https://hal.inria.fr/hal-01275535.

[19] C. LIU, O. CHAPUIS, M. BEAUDOUIN-LAFON, E. LECOLINET.*CoReach: Cooperative Gestures for Data Manipulation on Wall-sized Displays*, in "Proceedings of the 35th international conference on Human factors in computing systems", Denver, United States, ACM (editor), CHI '17, May 2017 [*DOI :* 10.1145/3025453.3025594], https://hal.archives-ouvertes.fr/hal-01437091.

[20] J.-B. LOUVET, C. FLEURY.*Combining Bimanual Interaction and Teleportation for 3D Manipulation on Multi-Touch Wall-sized Displays*, in "Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology (VRST)", Munich, Germany, November 2016, p. 283 - 292 [*DOI :* 10.1145/2993369.2993390], https://hal.archives-ouvertes.fr/hal-01428914.

[21] M.-E. PERRIN, J. R. EAGAN, M. BEAUDOUIN-LAFON.*Human-oriented Infrastructures for Multi-surface Environments*, in "ACM Conference on Human Factors in Computing Systems (CHI '16)", San Jose, California, United States, ACM, May 2016, https://hal.archives-ouvertes.fr/hal-01435923.

[22] M. WESSELY, T. TSANDILAS, W. E. MACKAY.*Stretchis: Fabricating Highly Stretchable User Interfaces*, in "ACM Symposium on User Interface Software and Technology (UIST)", Tokyo, Japan, October 2016, p. 697-704 [*DOI :* 10.1145/2984511.2984521], https://hal.archives-ouvertes.fr/hal-01412289.

National Conferences with Proceeding

[23] J.-B. LOUVET, G. DUBUISSON DUPLESSIS, N. CHAIGNAUD, J.-P. KOTOWICZ, L. VERCOUTER.*Recherche collaborative de documents : comparaison assistance humaine/automatique*, in "IC2016: Ingénierie des Connaissances", Montpellier, France, June 2016, https://hal.archives-ouvertes.fr/hal-01442735.

Research Reports

[24] T. TSANDILAS, P. DRAGICEVIC.*Accounting for Chance Agreement in Gesture Elicitation Studies*, LRI - CNRS, University Paris-Sud, February 2016, n° 1584, 5, https://hal.archives-ouvertes.fr/hal-01267288.

Scientific Popularization

[25] M. BEAUDOUIN-LAFON.*50 ans d'interaction homme-machine : retours vers le futur*, in "Interstices", July 2016, https://hal.inria.fr/inria-00174327.

[26] M. BEAUDOUIN-LAFON.*Mieux penser les interfaces informatiques*, June 2016, 61, Journal du CNRS n°285, https://hal.archives-ouvertes.fr/hal-01435924.

Other Publications

[27] M. M. WANDERLEY, J. MALLOCH, J. GARCIA, W. E. MACKAY, M. BEAUDOUIN-LAFON, S. HUOT.*Human Computer Interaction meets Computer Music: The MIDWAY Project*, May 2016, 4, Workshop Music and HCI, in conjunction with ACM Conference on Human Factors in Computing Systems (CHI '16), https://hal.inria.fr/hal-01370588.

[28] O. ZINENKO, S. HUOT, C. BASTOUL.*Interactive Program Restructuring*, November 2016, working paper or preprint, https://hal.inria.fr/hal-01406294.

Project-Team GALEN

Organ Modeling through Extraction, Representation and Understanding of Medical Image Content

Table of contents

Project-Team GALEN

Creation of the Team: 2008 February 15, updated into Project-Team: 2014 January 01

Keywords:

Computer Science and Digital Science:

3.4. - Machine learning and statistics
3.4.1. - Supervised learning
3.4.2. - Unsupervised learning
3.4.3. - Reinforcement learning
3.4.4. - Optimization and learning
3.4.5. - Bayesian methods
3.4.6. - Neural networks
3.4.7. - Kernel methods
3.4.8. - Deep learning
5.3.1. - Compression
5.3.2. - Sparse modeling and image representation
5.3.3. - Pattern recognition
5.3.4. - Registration
5.4.1. - Object recognition
5.4.2. - Activity recognition
5.4.4. - 3D and spatio-temporal reconstruction
5.9. - Signal processing
5.9.1. - Sampling, acquisition
5.9.2. - Estimation, modeling
5.9.3. - Reconstruction, enhancement
5.9.4. - Signal processing over graphs
5.9.5. - Sparsity-aware processing
5.9.6. - Optimization tools
6.2.3. - Probabilistic methods
6.2.4. - Statistical methods
6.2.6. - Optimization
6.2.7. - High performance computing
6.3.1. - Inverse problems
7.2. - Discrete mathematics, combinatorics
7.3. - Optimization
7.8. - Information theory
7.9. - Graph theory
8.2. - Machine learning
8.3. - Signal analysis
8.5. - Robotics

Other Research Topics and Application Domains:

2.2.3. - Cancer

1. Members

Faculty Members

Nikos Paragios [Team leader, CentraleSupélec, Professor, HDR]
Jean-Christophe Pesquet [CentraleSupélec, Professor, from Oct 2016, HDR]
Iasonas Kokkinos [CentraleSupélec,Associate Professor, until Aug 2016, HDR]
Dimitrios Samaras [CentraleSupélec, Visiting Professor, from Oct 2016]
Emilie Chouzenoux [Univ. Paris Est, Associate Professor, from Sep 2016]
Evangelia Zacharaki [CentraleSupélec, Visiting Professor]

Technical Staff

Rafael Marini Silva [CentraleSupélec]

PhD Students

Stavros Alchatzidis [CentraleSupélec, until Sep 2016]
Eugène Belilovsky [CentraleSupélec]
Maxim Berman [CentraleSupélec]
Siddhartha Chandra [Inria]
Vivien Fecamp [until Mar 2016]
Enzo Ferrante [CentraleSupélec, until Mar 2016]
Riza Guler [Inria]
Hariprasad Kannan [CentraleSupélec]
Stefan Kinauer [CentraleSupélec]
Evgenios Kornaropoulos [CentraleSupélec]
Huu Dien Khue Le [CentraleSupélec]
Azar Louzi [CentraleSupélec, from Dec 2016]
Mihir Sahasrabudhe [CentraleSupélec]
Jiaqian Yu [CentraleSupélec]
Marie-Caroline Corbineau [CentraleSupélec, from Oct 2016]
Loubna El Gueddari [Université Paris-Sud, from Oct 2016]
Guillaume Chassagnon [CentraleSupélec, from Jun 2016]

Post-Doctoral Fellows

Puneet Dokania [Inria, until May 2016]
Stavros Tsogkas [CentraleSupélec, until Aug 2016]

Administrative Assistant

Alexandra Merlin [Inria]

Others

Natalia Leclercq [CentraleSupélec, Administrative Assistant ou Engineers]
Jie Ying Wu [Intern, from Mar 2016 until Jul 2016]
Zizhao Li [Intern, from May 26th 2016 until Aug 24th 2016]

2. Overall Objectives

2.1. GALEN@Centrale-Paris

Computational vision is one of the most challenging research domains in engineering sciences. The aim is to reproduce human visual perception through intelligent processing of visual data. The application domains span from computer aided diagnosis to industrial automation & robotics. The most common mathematical formulation to address such a challenge is through mathematical modeling. In such a context, first the solution of the desired vision task is expressed in the form of a parameterized mathematical model. Given such a model, the next task consists of associating the model parameters with the available observations, which is often called the model-to-data association. The aim of this task is to determine the impact of a parameter choice to the observations and eventually maximize/minimize the adequacy of these parameters with the visual observations. In simple words, the better the solution is, the better it will be able to express and fit the data. This is often achieved through the definition of an objective function on the parametric space of the model. Last, but not least given the definition of the objective function, visual perception is addressed through its optimization with respect to the model parameters. To summarize, computation visual perception involves three aspects, a task-specific definition of a parametric model, a data-specific association of this model with the available observations and last the optimization of the model parameters given the objective and the observations.

Such a chain processing inherits important shortcomings. The curse of dimensionality is often used to express the importance of the model complexity. In simple words, the higher the complexity of the model is, the better its expressive power will be with counter effect the increase of the difficulty of the inference process. Non-linearity is another issue to be addressed which simply states that the association between the model and the data is a (highly) non-linear function and therefore direct inference is almost infeasible. The impact of this aspect is enforced from the curse of non-convexity that characterizes the objective function. Often it lives in high-dimensional spaces and is ill posed making exact inference problematic (in many cases not possible) and computationally expensive. Last, but not least modularity and scalability is another important concern to be addressed in the context of computational vision. The use of task-specific modeling and algorithmic solutions make their portability infeasible and therefore transfer of knowledge from one task to another is not straightforward while the methods do not always scale well with respect either to the dimensionality of the representation or the data.

GALEN aims at proposing innovative techniques towards automatic structuring, interpretation and longitudinal modeling of visual data. In order to address these fundamental problems of computational perception, GALEN investigates the use of discrete models of varying complexity. These methods exhibit an important number of strengths such as their ability to be modular with respect to the input measurements (clinical data), the nature of the model (certain constraints are imposed from computational perspective in terms of the level of interactions), and the model-to-data association while being computational efficient.

3. Research Program

3.1. Shape, Grouping and Recognition

A general framework for the fundamental problems of image segmentation, object recognition and scene analysis is the interpretation of an image in terms of a set of symbols and relations among them. Abstractly stated, image interpretation amounts to mapping an observed image, X to a set of symbols Y. Of particular interest are the symbols Y^* that *optimally explain the underlying image*, as measured by a scoring function s that aims at distinguishing correct (consistent with human labellings) from incorrect interpretations:

$$Y^* = \mathrm{argmax}_Y s(X, Y) \tag{1}$$

Applying this framework requires (a) identifying which symbols and relations to use (b) learning a scoring function s from training data and (c) optimizing over Y in Eq.1.

One of the main themes of our work is the development of methods that jointly address (a,b,c) in a shape-grouping framework in order to reliably extract, describe, model and detect shape information from natural and medical images. A principal motivation for using a shape-based framework is the understanding that shape- and more generally, grouping- based representations can go all the way from image features to objects. Regarding aspect (a), image representation, we cater for the extraction of image features that respect the shape properties of image structures. Such features are typically constructed to be purely geometric (e.g. boundaries, symmetry axes, image segments), or appearance-based, such as image descriptors. The use of machine learning has been shown to facilitate the robust and efficient extraction of such features, while the grouping of local evidence is known to be necessary to disambiguate the potentially noisy local measurements. In our research we have worked on improving feature extraction, proposing novel blends of invariant geometric- and appearance- based features, as well as grouping algorithms that allow for the efficient construction of optimal assemblies of local features.

Regarding aspect (b) we have worked on learning scoring functions for detection with deformable models that can exploit the developed low-level representations, while also being amenable to efficient optimization. Our works in this direction build on the graph-based framework to construct models that reflect the shape properties of the structure being modeled. We have used discriminative learning to exploit boundary- and symmetry-based representations for the construction of hierarchical models for shape detection, while for medical images we have developed methods for the end-to-end discriminative training of deformable contour models that combine low-level descriptors with contour-based organ boundary representations.

Regarding aspect (c) we have developed algorithms which implement top-down/bottom-up computation both in deterministic and stochastic optimization. The main idea is that 'bottom-up', image-based guidance is necessary for efficient detection, while 'top-down', object-based knowledge can disambiguate and help reliably interpret a given image; a combination of both modes of operation is necessary to combine accuracy with efficiency. In particular we have developed novel techniques for object detection that employ combinatorial optimization tools (A* and Branch-and-Bound) to tame the combinatorial complexity, achieving a best-case performance that is logarithmic in the number of pixels.

In the long run we aim at scaling up shape-based methods to 3D detection and pose estimation and large-scale object detection. One aspect which seems central to this is the development of appropriate mid-level representations. This is a problem that has received increased interest lately in the 2D case and is relatively mature, but in 3D it has been pursued primarily through ad-hoc schemes. We anticipate that questions pertaining to part sharing in 3D will be addressed most successfully by relying on explicit 3D representations. On the one hand depth sensors, such as Microsoft's Kinect, are now cheap enough to bring surface modeling and matching into the mainstream of computer vision - so these advances may be directly exploitable at test time for detection. On the other hand, even if we do not use depth information at test time, having 3D information can simplify the modeling task during training. In on-going work with collaborators we have started exploring combinations of such aspects, namely (i) the use of surface analysis tools to match surfaces from depth sensors (ii) using branch-and-bound for efficient inference in 3D space and (iii) groupwise-registration to build statistical 3D surface models. In the coming years we intend to pursue a tighter integration of these different directions for scalable 3D object recognition.

3.2. Machine Learning & Structured Prediction

The foundation of statistical inference is to learn a function that minimizes the expected loss of a prediction with respect to some unknown distribution

$$\mathcal{R}(f) = \int \ell(f, x, y) dP(x, y), \tag{2}$$

where $\ell(f, x, y)$ is a problem specific loss function that encodes a penalty for predicting $f(x)$ when the correct prediction is y. In our case, we consider x to be a medical image, and y to be some prediction, e.g. the segmentation of a tumor, or a kinematic model of the skeleton. The loss function, ℓ, is informed by the costs associated with making a specific misprediction. As a concrete example, if the true spatial extent of a tumor is encoded in y, $f(x)$ may make mistakes in classifying healthy tissue as a tumor, and mistakes in classifying diseased tissue as healthy. The loss function should encode the potential physiological damage resulting from erroneously targeting healthy tissue for irradiation, as well as the risk from missing a portion of the tumor.

A key problem is that the distribution P is unknown, and any algorithm that is to estimate f from labeled training examples must additionally make an implicit estimate of P. A central technology of empirical inference is to approximate $\mathcal{R}(f)$ with the empirical risk,

$$\mathcal{R}(f) \approx \widehat{\mathcal{R}}(f) = \frac{1}{n} \sum_{i=1}^{n} \ell(f, x_i, y_i), \tag{3}$$

which makes an implicit assumption that the training samples (x_i, y_i) are drawn i.i.d. from P. Direct minimization of $\widehat{\mathcal{R}}(f)$ leads to overfitting when the function class $f \in \mathcal{F}$ is too rich, and regularization is required:

$$\min_{f \in \mathcal{F}} \lambda \Omega(\|f\|) + \widehat{\mathcal{R}}(f), \tag{4}$$

where Ω is a monotonically increasing function that penalizes complex functions.

Equation Eq. 4 is very well studied in classical statistics for the case that the output, $y \in \mathcal{Y}$, is a binary or scalar prediction, but this is not the case in most medical imaging prediction tasks of interest. Instead, complex interdependencies in the output space leads to difficulties in modeling inference as a binary prediction problem. One may attempt to model e.g. tumor segmentation as a series of binary predictions at each voxel in a medical image, but this violates the i.i.d. sampling assumption implicit in Equation Eq. 3. Furthermore, we typically gain performance by appropriately modeling the inter-relationships between voxel predictions, e.g. by incorporating pairwise and higher order potentials that encode prior knowledge about the problem domain. It is in this context that we develop statistical methods appropriate to structured prediction in the medical imaging setting.

3.3. Self-Paced Learning with Missing Information

Many tasks in artificial intelligence are solved by building a model whose parameters encode the prior domain knowledge and the likelihood of the observed data. In order to use such models in practice, we need to estimate its parameters automatically using training data. The most prevalent paradigm of parameter estimation is supervised learning, which requires the collection of the inputs x_i and the desired outputs y_i. However, such an approach has two main disadvantages. First, obtaining the ground-truth annotation of high-level applications, such as a tight bounding box around all the objects present in an image, is often expensive. This prohibits the use of a large training dataset, which is essential for learning the existing complex models. Second, in many applications, particularly in the field of medical image analysis, obtaining the ground-truth annotation may not be feasible. For example, even the experts may disagree on the correct segmentation of a microscopical image due to the similarities between the appearance of the foreground and background.

In order to address the deficiencies of supervised learning, researchers have started to focus on the problem of parameter estimation with data that contains hidden variables. The hidden variables model the missing information in the annotations. Obtaining such data is practically more feasible: image-level labels ('contains car','does not contain person') instead of tight bounding boxes; partial segmentation of medical images. Formally, the parameters $\mathbf{w}$ of the model are learned by minimizing the following objective:

$$\min_{\mathbf{w} \in \mathcal{W}} R(\mathbf{w}) + \sum_{i=1}^{n} \Delta(y_i, y_i(\mathbf{w}), h_i(\mathbf{w})). \tag{5}$$

Here, $\mathcal{W}$ represents the space of all parameters, n is the number of training samples, $R(\cdot)$ is a regularization function, and $\Delta(\cdot)$ is a measure of the difference between the ground-truth output y_i and the predicted output and hidden variable pair $(y_i(\mathbf{w}), h_i(\mathbf{w}))$.

Previous attempts at minimizing the above objective function treat all the training samples equally. This is in stark contrast to how a child learns: first focus on easy samples ('learn to add two natural numbers') before moving on to more complex samples ('learn to add two complex numbers'). In our work, we capture this intuition using a novel, iterative algorithm called self-paced learning (SPL). At an iteration t, SPL minimizes the following objective function:

$$\min_{\mathbf{w}\in\mathcal{W},\mathbf{v}\in\{0,1\}^n} R(\mathbf{w}) + \sum_{i=1}^{n} v_i \Delta(y_i, y_i(\mathbf{w}), h_i(\mathbf{w})) - \mu_t \sum_{i=1}^{n} v_i. \tag{6}$$

Here, samples with $v_i = 0$ are discarded during the iteration t, since the corresponding loss is multiplied by 0. The term μ_t is a threshold that governs how many samples are discarded. It is annealed at each iteration, allowing the learner to estimate the parameters using more and more samples, until all samples are used. Our results already demonstrate that SPL estimates accurate parameters for various applications such as image classification, discriminative motif finding, handwritten digit recognition and semantic segmentation. We will investigate the use of SPL to estimate the parameters of the models of medical imaging applications, such as segmentation and registration, that are being developed in the GALEN team. The ability to handle missing information is extremely important in this domain due to the similarities between foreground and background appearances (which results in ambiguities in annotations). We will also develop methods that are capable of minimizing more general loss functions that depend on the (unknown) value of the hidden variables, that is,

$$\min_{\mathbf{w}\in\mathcal{W},\theta\in\Theta} R(\mathbf{w}) + \sum_{i=1}^{n} \sum_{h_i\in\mathcal{H}} \Pr(h_i|x_i, y_i; \theta) \Delta(y_i, h_i, y_i(\mathbf{w}), h_i(\mathbf{w})). \tag{7}$$

Here, θ is the parameter vector of the distribution of the hidden variables h_i given the input x_i and output y_i, and needs to be estimated together with the model parameters $\mathbf{w}$. The use of a more general loss function will allow us to better exploit the freely available data with missing information. For example, consider the case where y_i is a binary indicator for the presence of a type of cell in a microscopical image, and h_i is a tight bounding box around the cell. While the loss function $\Delta(y_i, y_i(\mathbf{w}), h_i(\mathbf{w}))$ can be used to learn to classify an image as containing a particular cell or not, the more general loss function $\Delta(y_i, h_i, y_i(\mathbf{w}), h_i(\mathbf{w}))$ can be used to learn to detect the cell as well (since h_i models its location)

3.4. Discrete Biomedical Image Perception

A wide variety of tasks in medical image analysis can be formulated as discrete labeling problems. In very simple terms, a discrete optimization problem can be stated as follows: we are given a discrete set of variables $\mathcal{V}$, all of which are vertices in a graph $\mathcal{G}$. The edges of this graph (denoted by $\mathcal{E}$) encode the variables' relationships. We are also given as input a discrete set of labels $\mathcal{L}$. We must then assign one label from $\mathcal{L}$ to each variable in $\mathcal{V}$. However, each time we choose to assign a label, say, x_{p_1} to a variable p_1, we are forced to pay a price according to the so-called *singleton* potential function $g_p(x_p)$, while each time we choose to assign a pair of labels, say, x_{p_1} and x_{p_2} to two interrelated variables p_1 and p_2 (two nodes that are connected by an edge in the graph $\mathcal{G}$), we are also forced to pay another price, which is now determined by the so called *pairwise* potential function $f_{p_1 p_2}(x_{p_1}, x_{p_2})$. Both the singleton and pairwise potential functions are problem specific and are thus assumed to be provided as input.

Our goal is then to choose a labeling which will allow us to pay the smallest total price. In other words, based on what we have mentioned above, we want to choose a labeling that minimizes the sum of all the MRF potentials, or equivalently the MRF energy. This amounts to solving the following optimization problem:

$$\arg\min_{\{x_p\}}\mathcal{P}(g,f) = \sum_{p\in\mathcal{V}}g_p(x_p) + \sum_{(p_1,p_2)\in\mathcal{E}}f_{p_1p_2}(x_{p_1},x_{p_2}). \tag{8}$$

The use of such a model can describe a number of challenging problems in medical image analysis. However these simplistic models can only account for simple interactions between variables, a rather constrained scenario for high-level medical imaging perception tasks. One can augment the expression power of this model through higher order interactions between variables, or a number of cliques $\{C_i, i \in [1,n] = \{\{p_{i^1},\cdots,p_{i|C_i|}\}\}$ of order $|C_i|$ that will augment the definition of $\mathcal{V}$ and will introduce hyper-vertices:

$$\arg\min_{\{x_p\}}\mathcal{P}(g,f) = \sum_{p\in\mathcal{V}}g_p(x_p) + \sum_{(p_1,p_2)\in\mathcal{E}}f_{p_1p_2}(x_{p_1},x_{p_2}) + \sum_{C_i\in\mathcal{E}}f_{p_1\cdots p_n}(x_{p_{i^1}},\cdots,p_{x_{i|C_i|}}). \tag{9}$$

where $f_{p_1\cdots p_n}$ is the price to pay for associating the labels $(x_{p_{i^1}},\cdots,p_{x_{i|C_i|}})$ to the nodes $(p_1\cdots p_{i|C_i|})$. Parameter inference, addressed by minimizing the problem above, is the most critical aspect in computational medicine and efficient optimization algorithms are to be evaluated both in terms of computational complexity as well as of inference performance. State of the art methods include deterministic and non-deterministic annealing, genetic algorithms, max-flow/min-cut techniques and relaxation. These methods offer certain strengths while exhibiting certain limitations, mostly related to the amount of interactions which can be tolerated among neighborhood nodes. In the area of medical imaging where domain knowledge is quite strong, one would expect that such interactions should be enforced at the largest scale possible.

4. Application Domains

4.1. Testing for Difference in Functional Brain Connectivity

Paticipants: Eugene Belilovsky, Matthew Blaschko Collaboration with Inria Parietal: Gael Varoquaux

Proposed a new algorithm for determining the differences in functional brain connectivity between two populations. The aim of our work was to leverage assumptions and show a method that can efficiently provide significance results in the form of (p-values). We demonstrated that our approach works well in practice and simulation and can provide faithful p-values on complicated fMRI data.

4.2. Lung Tumor Detection and Characterization

Paticipants: Evgenios Kornaropoulos, Evangelia Zacharaki, Nikos Paragios

The use of Diffusion Weighted MR Imaging (DWI) is investigated as an alternative tool to radiologists for tumor detection, tumor characterization, distinguishing tumor tissue from non-tumor tissue, and monitoring and predicting treatment response. In collaboration with Hôpitaux Universitaires Henri-Mondor in Paris, France and Chang Gung Memorial Hospital – Linkou in Taipei, Taiwan we investigate the use of model-based methods of 3D image registration, clustering and segmentation towards the development of a framework for automatic interpretation of images, and in particular extraction of meaningful biomarkers in aggressive lymphomas [23][24]. In [23] we combine deformable group-wise registration with a physiological model in order to better estimate diffusion in Diffusion-Weighted MRI, whereas in [24] we explicitly model the diffusion coefficients by a high-order MRF-based joint deformable registration and labeling scheme.

4.3. Protein function prediction

Paticipants: Evangelia Zacharaki, Nikos Paragios (in collaboration with D. Vlachakis, University of Patras, Greece)

The massive expansion of the worldwide Protein Data Bank (PDB) provides new opportunities for computational approaches which can learn from available data and extrapolate the knowledge into new coming instances. The aim of our work in [14] was to exploit experimentally acquired structural information of enzymes through machine learning techniques in order to produce models that predict enzymatic function.

4.4. Imaging biomarkers for chronic lung diseases

Paticipants: Guillaume Chassagnon, Evangelia Zacharaki, Nikos Paragios

Diagnosis and staging of chronic lung diseases is a major challenge for both patient care and approval of new treatments. Among imaging techniques, computed tomography (CT) is the gold standard for in vivo morphological assessment of lung parenchyma currently offering the highest spatial resolution in chronic lung diseases. Although CT is widely used its optimal use in clinical practice and as an endpoint in clinical trials remains controversial. Our goal is to develop quantitative imaging biomarkers that allow (i) severity assessment (based on the correlation to functional and clinical data) and (ii) monitoring the disease progression. In the current analysis we focus on scleroderma and cystic fibrosis as models for restrictive and obstructive lung disease, respectively. Two different approaches are investigated: disease assessment by histogram or texture analysis and assessment of the regional lung elasticity through deformable registration. This work is in collaboration with the Department of Radiology, Cochin Hospital, Paris.

4.5. Co-segmentation and Co-registration of Subcortical Brain Structures

Paticipants: Enzo Ferrante, Nikos Paragios, Iasonas Kokkinos

New algorithms to perform co-segmentation and co-registration of subcortical brain structures on MRI images were investigated in collaboration with Ecole Polytechnique de Montreal and the Sainte-Justine Hospital Research Center from Montreal [40]. Brain subcortical structures are involved in different neurodegenerative and neuropsychiatric disorders, including schizophrenia, Alzheimers disease, attention deficit, and subtypes of epilepsy. Segmenting these parts of the brain enables a physician to extract indicators, facilitating their quantitative analysis and characterization. We are investigating how estimated maps of semantic labels (obtained using machine learning techniques) can be used as a surrogate for unlabelled data. We are exploring how to combine them with multi-population deformable registration to improve both alignment and segmentation of these challenging brain structures.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- Wacha Bounliphone and Eugène Belilovsky received the Université Paris-Saclay STIC Doctoral School Best Scientific Contribution Award
- Eugène Belilovsky received the MITACS-Inria Globalink Award
- Prof. Iasonas Kokkinos was invited as keynote speaker in Astronomical Data Analysis Summer School, Chania, Greece, May 2016
- Prof. Iasonas Kokkinos was invited as keynote speaker in Local features workshop, held in conjunction with ECCV, October 2016
- Dr. Evangelia Zacharaki was appointed as guest associate editor for the Medical Physics journal

- 2nd place at the 2016 IEEE GRSS Data Fusion Contest for the paper: Simultaneous Registration, Segmentation and Change Detection from Multisensor, Multitemporal Satellite Image Pairs [30].

- Finalists (not-awarded) of the Best Papers award at the IEEE conference ICIP'16 for the paper: A Block Parallel Majorize-Minimize Memory Gradient Algorithm [16].

- Oral presentation in the Neural Information Processing Systems (NIPS), 2016 conference of the paper: Testing for Differences in Gaussian Graphical Models: Applications to Brain Connectivity [15] (oral presentations: only 1% of more than 2000 submitted papers).

- Oral presentation in the British Machine Vision Conference (BMVC), 2016, of the paper: Efficient Learning for Discriminative Segmentation with Supermodular Losses [33] (oral presentations: 7% of submitted papers).

- Oral presentation in the International Conference on Artificial Intelligence and Statistics (AISTATS), 2016, of the paper: A Convex Surrogate Operator for General Non-Modular Loss Functions [32] (oral presentations: 11% of submitted papers).

5.1.2. Other

- Acceptance of the project entitled «Predicteurs performants de l'efficacite des agents anticancereux par apprentissage profond (deep learning) de donnees radiomiques et genomiques» as part of the program *Imagerie Médicale Computationnelle*. PI: Dr. Charles Ferte, Gustave Roussy, 94805 Villejuif.

6. New Software and Platforms

6.1. DISD

Dense Image and Surface Descriptors
FUNCTIONAL DESCRIPTION Scale-Invariant Descriptor, Scale-Invariant Heat Kernel Signatures DISD implements the SID, SI-HKS and ISC descriptors. SID (Scale-Invariant Descriptor) is a densely computable, scale- and rotation- invariant descriptor. We use a log-polar grid around every point to turn rotation/scalings into translation, and then use the Fourier Transform Modulus (FTM) to achieve invariance. SI-HKS (Scale-Invariant Heat Kernel Signatures) extract scale-invariant shape signatures by exploiting the fact that surface scaling amounts to multiplication and scaling of a properly sampled HKS descriptor. We apply the FTM trick on HKS to achieve invariance to scale changes. ISC (Intrinsic Shape Context) constructs a net-like grid around every surface point by shooting outwards and tracking geodesics. This allows us to build a meta-descriptor on top of HKS/SI-HKS that takes neighborhood into account, while being invariant to surface isometries.

- Participants: Iasonas Kokkinos and Eduard Trulls
- Contact: Iasonas Kokkinos
- URL: http://vision.mas.ecp.fr/Personnel/iasonas/descriptors.html

6.2. DPMS

FUNCTIONAL DESCRIPTION DPMS implements branch-and-bound object detection, cutting down the complexity of detection from linear in the number of pixels to logarithmic.

- Participant: Iasonas Kokkinos
- Contact: Iasonas Kokkinos
- URL: http://cvn.ecp.fr/personnel/iasonas/dpms.html

6.3. ECP

Part-Based Object Detection

FUNCTIONAL DESCRIPTION A DPM model gives rise to the energy function being optimized during inference and has to be learned offline beforehand. All functionality is accessible via a ROS interface. The work follows closely Felzenszwalb's "Object Detection with Discriminatively Trained Part Based Models", using the FFLD-library in the implementation and facilitating communication in a setup with NAO robots.

- Participants: Haithem Boussaid, Stefan Kinauer, Iasonas Kokkinos
- Contact: Stefan Kinauer
- URL: https://bitbucket.org/eu-reconfig/ecp

6.4. DROP

KEYWORDS: Health - Merging - Registration of 2D and 3D multimodal images - Medical imaging
FUNCTIONAL DESCRIPTION Drop is a software programme that registers images originating from one or more modes by quickly and efficiently calculating a non-rigid / deformable field of deformation. Drop is a new, quick and effective registration tool based on new algorithms that do not require a cost function derivative.

- Partner: Centrale Paris
- Contact: Nikos Paragios
- URL: http://campar.in.tum.de/Main/Drop

6.5. FastPD

KEYWORD: Medical imaging
FUNCTIONAL DESCRIPTION FastPD is an optimization platform in C++ for the computer vision and medical imaging community.

- Contact: Nikolaos Paragyios
- URL: http://www.csd.uoc.gr/~komod/FastPD/

6.6. GraPeS

Grammar Parser for Shapes
FUNCTIONAL DESCRIPTION It is a software for parsing facade images using shape grammars. Grapes implement a parsing methods based on Reinforcement Learning principles. It optimizes simultaneously the topology of the parse tree as well as the associated parameters. GraPeS comes along with predefined shape grammars as XML files and defines three kinds of rewards. However, it also offers the possibility to create new grammars and to provide custom rewards in text files, widening the scope of potential applications. The name of the software comes from the aspect of the parse tree of the binary split grammars involved in the process.

- Participant: Iasonas Kokkinos
- Contact: Iasonas Kokkinos
- URL: http://vision.mas.ecp.fr/Personnel/teboul/grapesPage/index.php

6.7. HOAP-SVM

High-Order Average Precision SVM
SCIENTIFIC DESCRIPTION

We consider the problem of using high-order information (for example, persons in the same image tend to perform the same action) to improve the accuracy of ranking (specifically, average precision). We develop two learning frameworks. The high-order binary SVM (HOB-SVM) optimizes a convex upper bound of the surrogate 0-1 loss function. The high-order average precision SVM (HOAP-SVM) optimizes a difference-of-convex upper bound on the average precision loss function. Authors of the research paper: Puneet K. Dokania, A. Behl, C. V. Jawahar and M. Pawan Kumar

FUNCTIONAL DESCRIPTION The software provides a convenient API for learning to rank with high-order information. The samples are ranked according to a score that is proportional to the difference of max-marginals of the positive and the negative class. The parameters of the score function are computed by minimizing an upper bound on the average precision loss. The software also provides an instantiation of the API for ranking samples according to their relevance to an action, using the poselet features. The following learning algorithms are included in the API:

(1) Multiclass-SVM (2) AP-SVM (3) High Order Binary SVM (HOB-SVM) (4) High Order AP-SVM (HOAP-SVM) (5) M4 Learning (unpublished work)

The API is developed in C/C++ by Puneet K. Dokania.

- Participants: Puneet Dokania and Pawan Kumar
- Contact: Puneet Dokania
- URL: http://puneetkdokania.github.io/projects/ranking-highorder/ranking-highorder.html

6.8. LBSD

Learning-Based Symmetry Detection
FUNCTIONAL DESCRIPTION LBSD implements the learning-based approach to symmetry detection. It includes the code for running a detector, alongside with the ground-truth symmetry annotations that we have introduced for the Berkeley Segmentation Dataset (BSD) benchmark.

- Participant: Stavros Tsogkas
- Contact: Stavros Tsogkas
- URL: https://github.com/tsogkas/oid_1.0

6.9. TeXMeG

FUNCTIONAL DESCRIPTION Texture, modulation, generative models, segmentation, TeXMeG is a front-end for texture analysis and edge detection platform in Matlab that relies on Gabor filtering and image demodulation. Includes frequency- and time- based definition of Gabor- and other Quadrature-pair filterbanks, demodulation with the Regularized Energy Separation Algorithm and Texture/Edge/Smooth classification based on MDL criterion.

- Participant: Iasonas Kokkinos
- Contact: Iasonas Kokkinos
- URL: http://cvsp.cs.ntua.gr/software/texture/

6.10. mrf-registration

KEYWORDS: Health - Medical imaging
FUNCTIONAL DESCRIPTION Deformable image and volume registration, is a deformable registration platform in C++ for the medical imaging community. This is the first publicly available platform which contains most of the existing metrics to perform registration under the same concept. The platform is used for clinical research from approximately 3,000 users worldwide.

- Participant: Nikos Paragios
- Contact: Nikos Paragios
- URL: http://www.mrf-registration.net/

6.11. Newton-MRF

FUNCTIONAL DESCRIPTION MAP inference in MRFs can be performed through the LP relaxation approach. This project was a study of the feasibility and benefit of Newton-type methods for solving the optimization problem that is obtained by smoothing the dual of the LP relaxation. The project TRN-MRF is a trust-region Newton method that can address inference in higher order MRFs, for cases in which decomposition according to individual cliques leads to practical convergence. The project QN-MRF works for problems that need decomposition according to larger sub-problems (like chains of higher order cliques) and it works as long as a node is shared by exactly two sub-problems.

- Participants: Hariprasad Kannan, Nikos Paragios, Nikos Komodakis
- Contact: Hariprasad Kannan
- URL: https://github.com/neostrider/TRN-MRF, https://github.com/neostrider/QN-MRF

6.12. Relative MMD

FUNCTIONAL DESCRIPTION This software can be used to compare two distributions to a reference distribution. Applications include model selection and exploratory data analysis.

- Participants: Eugene Belilovsky, Wacha Bounliphone, Matthew Blaschko (in collaboration with researchers at UCL and Deepmind)
- Contact: Eugene Belilovsky
- URL: https://github.com/eugenium/MMD

6.13. Deep Graph Structure Discovery

FUNCTIONAL DESCRIPTION Novel approach to graph structure discovery using neural network. This software provides a much faster approach to discover underlying conditional independence structure in small sample data.

- Participants: Eugene Belilovsky
- Contact: Eugene Belilovsky
- URL: https://github.com/eugenium/LearnGraphDiscovery

7. New Results

7.1. Learning Grammars for Architecture-Specific Facade Parsing

Paticipants: Nikos Paragios (in collaboration with researchers from Université Paris-Est, LIGM, ENPC)

In [5], we present a novel framework to learn a compact grammar from a set of ground-truth images. To this end, parse trees of ground-truth annotated images are obtained running existing inference algorithms with a simple, very general grammar. From these parse trees, repeated subtrees are sought and merged together to share derivations and produce a grammar with fewer rules. Furthermore, unsupervised clustering is performed on these rules, so that, rules corresponding to the same complex pattern are grouped together leading to a rich compact grammar.

7.2. Non-Rigid Surface Registration

Paticipants: Dimitris Samaras, Nikos Paragios

This work [13] casts surface registration as the problem of finding a set of discrete correspondences through the minimization of an energy function, which is composed of geometric and appearance matching costs, as well as higher-order deformation priors. Two higher-order graph-based formulations are proposed under different deformation assumptions.

7.3. Monocular Surface Reconstruction using 3D Deformable Part Models

Paticipants: Maxim Berman, Stefan Kinauer, Iasonas Kokkinos

In this work [22] we train and detect part-based object models in 2D images, recovering 3D position and shape information (per part positions), allowing for a 3D reconstruction of the object. The resulting optimization problem is solved via a Branch&Bound approach, yielding detection results within a fraction of a second.

7.4. Learning with Non-modular loss functions

Paticipants: Jiaqian Yu, Matthew Blaschko

We have proposed an alternating direction method of multipliers (ADMM) based decomposition method loss augmented inference, that only depends on two individual solvers for the loss function term and for the inference term as two independent subproblems. In this way, we can gain computational efficiency and achieve more flexibility in choosing our non-modular loss functions of interest. We have proposed a novel supermodular loss function that empirically achieved better performance on the boundary of the objects, finding elongated structure [33]. We also introduced a novel convex surrogate operator for general non-modular loss functions, which provides for the first time a tractable solution for loss functions that are neither supermodular nor submodular, e.g. Dice loss. This surrogate is based on a canonical submodular-supermodular decomposition for which we have demonstrated its existence and uniqueness. It is further proven that this surrogate is convex, piecewise linear, an extension of the loss function, and for which subgradient computation is polynomial time [32][31].

7.5. Asymptotic Variance of MMD and Relative MMD

Paticipants: Eugene Belilovsky, Wacha Bounliphone, Matthew Blaschko (in collaboration with researchers at UCL and Deepmind)

Kernel mean embeddings allow for comparisons of complex distributions. They have been recently heavily used in hypothesis testing to compare distributions as well as in the nascent field of deep generative modeling. In this work we derived the asymptotic variance of the MMD and the cross covariance between joint MMD. We showed how this can be used effectively for model selection in complex Deep Generative Models where the likelihood metric is not accessible. Our results on the asymptotic variance of the MMD have already been used by other researchers to propose an efficient method for optimal testing and improved training of generative models.

7.6. Deconvolution and Deinterlacing of Video Sequences

Paticipants: Emilie Chouzenoux and Jean-Christophe Pesquet (in collaboration with F. Abboud, PhD student, J.-H. Chenot and L. Laborelli, research engineers, Institut National de l'Audiovisuel)

Optimization methods play a central role in the solution of a wide array of problems encountered in various application fields, such as signal and image processing. Especially when the problems are highly dimensional, proximal methods have shown their efficiency through their capability to deal with composite, possibly non smooth objective functions. The cornerstone of these approaches is the proximity operator, which has become a quite popular tool in optimization. In this work, we propose new dual forward-backward formulations for computing the proximity operator of a sum of convex functions involving linear operators. The proposed algorithms are accelerated thanks to the introduction of a block coordinate strategy combined with a preconditioning technique. Numerical simulations emphasize the good performance of our approach for the problem of jointly deconvoluting and deinterlacing video sequences.

7.7. A Variational Bayesian Approach for Restoring Data Corrupted with Non-Gaussian Noise

Paticipants: Emilie Chouzenoux and Jean-Christophe Pesquet (in collaboration with Y. Marnissi, PhD student at Univ. Paris-Est Marne la Vallée and Y. Zheng, IBM Research China)

In this work, a methodology is investigated for signal recovery in the presence of non-Gaussian noise. In contrast with regularized minimization approaches often adopted in the literature, in our algorithm the regularization parameter is reliably estimated from the observations. As the posterior density of the unknown parameters is analytically intractable, the estimation problem is derived in a variational Bayesian framework where the goal is to provide a good approximation to the posterior distribution in order to compute posterior mean estimates. Moreover, a majorization technique is employed to circumvent the difficulties raised by the intricate forms of the non-Gaussian likelihood and of the prior density. We demonstrate the potential of the proposed approach through comparisons with state-of-the-art techniques that are specifically tailored to signal recovery in the presence of mixed Poisson-Gaussian noise. Results show that the proposed approach is efficient and achieves performance comparable with other methods where the regularization parameter is manually tuned from an available ground truth.

7.8. The Majorize-Minimize Subspace Algorithm and Block Parallelization

Paticipants: Emilie Chouzenoux and Jean-Christophe Pesquet (in collaboration with S. Cadoni, Master student at Univ. Paris-Est Marne la Vallée and Dr C. Chaux, Univ. Aix-Marseille)

State-of-the-art methods for solving smooth optimization problems are nonlinear conjugate gradient, low memory BFGS, and Majorize-Minimize (MM) subspace algorithms. The MM subspace algorithm which has been introduced more recently has shown good practical performance when compared with other methods on various optimization problems arising in signal and image processing. However, to the best of our knowledge, no general result exists concerning the theoretical convergence rate of the MM subspace algorithm. The paper [3] aims at deriving such convergence rates both for batch and online versions of the algorithm and, in particular, discusses the influence of the choice of the subspace. We also propose a Block Parallel Majorize-Minimize Memory Gradient (BP3MG) algorithm for solving large scale optimization problems in [16]. This algorithm combines a block coordinate strategy with an efficient parallel update. The proposed method is applied to a 3D microscopy image restora- tion problem involving a depth-variant blur, where it is shown to lead to significant computational time savings with respect to a sequential approach.

7.9. Stochastic Forward-Backward and Primal-Dual Approximation Algorithms with Application to Online Image Restoration

Paticipants: Jean-Christophe Pesquet (In collaboration with Pr. P. L. Combettes, North Carolina State university)

Stochastic approximation techniques have been used in various contexts in data science. We propose a stochastic version of the forward-backward algorithm for minimizing the sum of two convex functions, one of which is not necessarily smooth. Our framework can handle stochastic approximations of the gradient of the smooth function and allows for stochastic errors in the evaluation of the proximity operator of the nonsmooth function. The almost sure convergence of the iterates generated by the algorithm to a minimizer is established under relatively mild assumptions. We also propose a stochastic version of a popular primal-dual proximal splitting algorithm, establish its convergence, and apply it to an online image restoration problem.

7.10. Random primal-dual proximal iterations for sparse multiclass SVM

Paticipants: Jean-Christophe Pesquet (in collaboration with Pr. G. Chierchia, Univ. Paris-Est Marne la Vallée, and Dr. N. Pustelnik, ENS Lyon)

Sparsity-inducing penalties are useful tools in variational methods for machine learning. In this paper, we propose two block-coordinate descent strategies for learning a sparse multiclass support vector machine. The first one works by selecting a subset of features to be updated at each iteration, while the second one performs the selection among the training samples. These algorithms can be efficiently implemented thanks to the flexibility offered by recent randomized primal-dual proximal methods. Experiments carried out for the supervised classification of handwritten digits demonstrate the interest of considering the primal-dual approach in the context of block-coordinate descent. The efficiency of the proposed algorithms is assessed through a comparison of execution times and classification errors.

7.11. PALMA, an improved algorithm for DOSY signal processing

Paticipants: Emilie Chouzenoux (in collaboration with Prof. M.-A. Delsuc, IGBMC, Strasbourg, and A. Cherni, PhD student, Univ. Strasbourg)

NMR is a tool of choice for the measure of diffusion coefficients of species in solution. The DOSY experiment, a 2D implementation of this measure, has proven to be particularly useful for the study of complex mixtures, molecular interactions, polymers, etc. However, DOSY data analysis requires to resort to inverse Laplace transform, in particular for polydisperse samples. This is a known difficult numerical task, for which we present here a novel approach. A new algorithm based on a splitting scheme and on the use of proximity operators is introduced. Used in conjunction with a Maximum Entropy and λ_1 hybrid regularisation [39], this algorithm converges rapidly and produces results robust against experimental noise. This method has been called PALMA. It is able to reproduce faithfully monodisperse as well as polydisperse systems, and numerous simulated and experimental examples are presented in [35]. It has been implemented on the server [http://palma.labo.igbmc.fr] where users can have their datasets processed automatically.

7.12. Graph-based change detection and classification in satellite image pairs

Paticipants: Maria Vakalopoulou, Nikos Paragios

We proposed a scalable, modular, metric-free, single-shot change detection/registration method for remote sensing image pairs [11]. The framework exploits a decomposed interconnected graphical model formulation where in the presence of changes the iconic similarity constraints are relaxed. We employ a discretized, grid-based deformation space. State-of-the-art linear programming and duality principles have been used to optimize the joint solution space where local consistency is imposed on the deformation and the detection space. The proposed framework is working both in a unsupervised and supervised manner depending on the application. The developed method has been validated through large scale experiments on several multi-temporal very high resolution optical satellite datasets. Also a novel generic framework has been designed, developed and validated for addressing simultaneously the tasks of image registration, segmentation and change detection from multisensor, multiresolution, multitemporal satellite image pairs [30]. Our approach models the inter-dependencies of variables through a higher order graph. A patch-based deep learning strategy has been employed and used for segmentation likelihoods. The evaluation of the developed framework was performed on the '2016 IEEE GRSS Data Fusion Contest' dataset and indicate very promising results for all three different tasks.

7.13. Graphical models in artificial vision

Paticipants: Nikos Komodakis, M. Pawan Kumar, Stavros Alchatzidis, Enzo Ferrante, Evangelia Zacharaki, Nikos Paragios

Computer vision tasks are often reformulated as mathematical inference problems where the objective is to determine the set of parameters corresponding to the lowest potential of a task-specific objective function. Graphical models have been the most popular formulation in the field over the past two decades. In [7] we focus on the inference component of the problem and in particular we discuss in a systematic manner the most commonly used optimization principles in the context of graphical models. In [8] we briefly review hyper-graph representations as prominent tools in the casting of perception as a graph optimization problem. We discuss their strength and limitations, provide appropriate strategies for their inference and present their application to address a variety of problems in biomedical image analysis.

Multi-atlas segmentation has emerged in recent years as a simple yet powerful approach in medical image segmentation. It commonly comprises two steps: (1) a series of pairwise registrations that establish correspondences between a query image and a number of atlases, and (2) the fusion of the available segmentation hypotheses towards labeling objects of interest. In [2], we introduce a novel approach that solves simultaneously for the underlying segmentation labels and the multi-atlas registration. We propose a pairwise Markov Random Field approach,where registration and segmentation nodes are coupled towards simultaneously recovering all atlas deformations and labeling the query image.

7.14. Pattern analysis of EEG signals with epileptic activity

Paticipants: Evangelia Zacharaki (in collaboration with Prof. M. Megalooikonomou, University of Patras and M. Koutroumanidis, King's College, London)

We have addressed the needs of epileptic patients and healthcare professionals, aiming at the design and development of a non-intrusive personal health system for the monitoring and analysis of epilepsy-relevant multi-parametric data and the documentation of the epilepsy related symptoms. Specifically, we investigated the classification of epileptic and non-epileptic events from EEG based on temporal and spectral analysis and different fusion schemes [9]. We also studied the EEG brain activity during whole night sleep, since sleep is recognized as a major precipitator of epileptic activity [12].

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

- Program: ANR Blanc International

    Project acronym: ADAMANTIUS

    Project title: Automatic Detection And characterization of residual Masses in pAtients with lymphomas through fusioN of whole-body diffusion-weighTed mrI on 3T and 18F-flUorodeoxyglucoSe pet/ct

    Duration: 9/2012-8/2015

    Coordinator: CHU Henri Mondor - FR

- Program: ANR JCJC

    Project acronym: HICORE

    Project title: HIerarchical COmpositional REpresentations for Computer Vision

    Duration: 10/2010-9/2014

    Coordinator: ECP - FR

- Program: ANR JCJC

    Project acronym: LearnCost

    Project title: Learning Model Constraints for Structured Prediction

    Duration: 2014-2018

    Coordinator: Inria Saclay - FR

- Program: ITMOs Cancer & Technologies pour la santé d'Aviesan / INCa

    Project acronym: CURATOR

    Project title: Slice-to-Image Deformable Registration towards Image-based Surgery Navigation & Guidance

    Duration: 12/2013-11/2015

    Coordinator: ECP - FR

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. I-SUPPORT

    Title: ICT-Supported Bath Robots

Programm: FP7

Duration: March 2015 - March 2018

Coordinator: Robotnik Automation S.L.L.

Partners:

Bethanien Krankenhaus - Geriatrisches Zentrum - Gemeinnutzige GMBH (Germany)

Fondazione Santa Lucia (Italy)

Institute of Communication and Computer Systems (Greece)

Karlsruher Institut für Technologie (Germany)

Theofanis Alexandridis Kai Sia Ee (OMEGATECH) (Greece)

Robotnik Automation Sll (Spain)

Scuola Superiore di Studi Universitari E di Perfezionamento Sant'Anna (Italy)

Frankfurt University of Applied Sciences (Germany)

Inria contact: Iasonas Kokkinos

The I-SUPPORT project envisions the development and integration of an innovative, modular, ICT-supported service robotics system that supports and enhances older adults' motion and force abilities and assists them in successfully, safely and independently completing the entire sequence of bathing tasks, such as properly washing their back, their upper parts, their lower limbs, their buttocks and groin, and to effectively use the towel for drying purposes. Advanced modules of cognition, sensing, context awareness and actuation will be developed and seamlessly integrated into the service robotics system to enable the robotic bathing system to adapt to the frail elderly population' capabilities and the frail elderly to interact in a master-slave mode, thus, performing bathing activities in an intuitive and safe way. Adaptation and integration of state-of-the-art, cost-effective, soft-robotic manipulators will provide the hardware constituents, which, together with advanced human-robot force/compliance control that will be developed within the proposed project, will form the basis for a safe physical human-robot interaction that complies with the most up-to-date safety standards. Human behavioural, sociological, safety, ethical and acceptability aspects, as well as financial factors related to the proposed service robotic infrastructure will be thoroughly investigated and evaluated so that the I-SUPPORT end result is a close-to-market prototype, applicable to realistic living settings.

8.2.1.2. MOBOT

Title: Intelligent Active MObility Aid RoBOT integrating Multimodal Communication

Programm: FP7

Duration: February 2013 - January 2016

Coordinator: Technische Universität München

Partners:

Bartlomiej Marcin Stanczyk (Poland)

Athena Research and Innovation Center in Information Communication & Knowledge Technologies (Greece)

Bethanien Krankenhaus - Geriatrisches Zentrum - Gemeinnutzige (Germany)

Diaplasis Rehabilitation Center (Greece)

Ecole Centrale des Arts et Manufactures (France)

Technische Universitaet Muenchen (Germany)

Ruprecht-Karls-Universitaet Heidelberg (Germany)

Inria contact: Iasonas Kokkinos

Mobility disabilities are prevalent in our ageing society and impede activities important for the independent living of elderly people and their quality of life. The MOBOT project aims at supporting mobility and thus enforcing fitness and vitality by developing intelligent active mobility assistance robots for indoor environments that provide user-centred, context-adaptive and natural support. Our driving concept envisions cognitive robotic assistants that act (a) proactively by realizing an autonomous and context-specific monitoring of human activities and by subsequently reasoning on meaningful user behavioural patterns, as well as (b) adaptively and interactively, by analysing multi-sensory and physiological signals related to gait and postural stability, and by performing adaptive compliance control for optimal physical support and active fall prevention. Towards these targets, a multimodal action recognition system will be developed to monitor, analyse and predict user actions with a high level of accuracy and detail. The main thrust of our approach will be the enhancement of computer vision techniques with modalities such as range sensor images, haptic information as well as command-level speech and gesture recognition. Data-driven multimodal human behaviour analysis will be conducted and behavioural patterns will be extracted. Findings will be imported into a multimodal human-robot communication system, involving both verbal and nonverbal communication and will be conceptually and systemically synthesised into mobility assistance models taking into consideration safety critical requirements. All these modules will be incorporated in a behaviour-based and context-aware robot control framework. Direct involvement of end-user groups will ensure that actual user needs are addressed. Finally, user trials will be conducted to evaluate and benchmark the overall system and to demonstrate the vital role of MOBOT technologies for Europe's service robotics.

8.2.1.3. RECONFIG

Type: FP7

Defi: Cognitive Systems and Robotics

Instrument: Specific Targeted Research Project

Objectif: Cognitive Systems and Robotics

Duration: February 2013 - January 2016

Coordinator: Dimos Dimarogonas

Partner: KTH (SE)

Inria contact: Iasonas Kokkinos

The RECONFIG project aims at exploiting recent developments in vision, robotics, and control to tackle coordination in heterogeneous multi-robot systems. Such systems hold promise for achieving robustness by leveraging upon the complementary capabilities of different agents and efficiency by allowing sub-tasks to be completed by the most suitable agent. A key challenge is that agent composition in current multi-robot systems needs to be constant and pre-defined. Moreover, the coordination of heterogeneous multi-agent systems has not been considered in manipulative scenarios. We propose a reconfigurable and adaptive decentralized coordination framework for heterogeneous multiple & multi-DOF robot systems. Agent coordination is held via two types of information exchange: (i) at an implicit level, e.g., when robots are in contact with each other and can sense the contact, and (ii) at an explicit level, using symbols grounded to each embodiment, e.g, when one robot notifies one other about the existence of an object of interest in its vicinity.

8.2.1.4. Strategie

Title: Statistically Efficient Structured Prediction for Computer Vision and Medical Imaging

Programm: FP7

Duration: January 2014 - December 2017

Coordinator: Inria

Inria contact: Matthew Blaschko

'Inference in medical imaging is an important step for disease diagnosis, tissue segmentation, alignment with an anatomical atlas, and a wide range of other applications. However, imperfections in imaging sensors, physical limitations of imaging technologies, and variation in the human population mean that statistical methods are essential for high performance. Statistical learning makes use of human provided ground truth to enable computers to automatically make predictions on future examples without human intervention. At the heart of statistical learning methods is risk minimization - the minimization of the expected loss on a previously unseen image. Textbook methods in statistical learning are not generally designed to minimize the expected loss for loss functions appropriate to medical imaging, which may be asymmetric and non-modular. Furthermore, these methods often do not have the capacity to model interdependencies in the prediction space, such as those arising from spatial priors, and constraints arising from the volumetric layout of human anatomy. We aim to develop new statistical learning methods that have these capabilities, to develop efficient learning algorithms, to apply them to a key task in medical imaging (tumor segmentation), and to prove their convergence to optimal predictors. To achieve this, we will leverage the structured prediction framework, which has shown impressive empirical results on a wide range of learning tasks. While theoretical results giving learning rates are available for some algorithms, necessary and sufficient conditions for consistency are not known for structured prediction. We will consequently address this issue, which is of key importance for algorithms that will be applied to life critical applications, e.g. segmentation of brain tumors that will subsequently be targeted by radiation therapy or removed by surgery. Project components will address both theoretical and practical issues.'

8.2.2. Collaborations with Major European Organizations

8.2.2.1. CT dose and IQ optimization

Title: Development of a system helping in optimizing and tailoring computed tomography protocols

Program: Collaboration et encadrement de these CIFRE

Partners: GE Medical Systems, Université Paris-Est Créteil Val de Marne

Duration: December 2014 - November 2017

The purpose of the PhD is to create a method of CT dose and IQ optimization taking into account clinical indication, patient's characteristics and previous exams. The method will consist (i) in linking a clinical indication to IQ defined by quantitative physical metrics and (ii) in quantifying the patient's morphology from the analysis of local scans.

8.2.2.2. Imaging biomarkers in ILD

Title: Développement d' outils de Quantification en Imagerie dans les maladies respiratoires chroniques fibrosantes et obstructives

Program: Fondation de Coopération Scientifique

Partners: GE Medical Systems SCS

Duration: April 2016 - March 2019

We aim to test two different approaches for the development of new imaging biomarkers in interstitial lung disease (scleroderma): assessment of the regional lung elasticity through deformable registration and tissue characterization by textural analysis. We will also study structural changes in obstructive lung disease (cystic fibrosis) in order to develop new imaging biomarkers based on elastic registration, histogram and textural analysis.

8.2.2.3. SuBSAmPLE

Title: Identification and prediction of salient brain states through probabilistic structure learning towards fusion of imaging and genomic date

Partners: Fondation de Coopération Scientifique Digiteo

Duration: January 2012 - December 2016 (extended by 1 year)

8.3. International Initiatives

8.3.1. Inria International Partners

8.3.1.1. Informal International Partners
- UCL – Collaboration with Professor Arthur Gretton. Topic: non-parametric hypothesis testing
- University of Toronto – Collaboration with Professors Raquel Urtasun and Richard Zemel. Topic: Representation of Scene Graphs for Image Retrieval and Visual Question Answering
- MILA, Universite de Montreal - Collaboration with Kyle Kastner from the lab of Professor Yoshua Bengio. Topic: Efficient Inference of Graph Structures using Deep Learning
- Sup'Com Tunis – Collaborative research with Amel Benazza-Benhayia. Collaboration Topic: Multispectral imaging
- University of Patras, Greece – Collaborative research with V. Megalooikonomou. Collaboration Topic: Biosignal analysis.
- University of Oxford – Collaborative research with Andrea Vedaldi. Collaboration Topic: Deep Learning for Texture Recognition.
- Google Research – Collaborative research with George Papandreou. Collaboration Topic: Deep Learning for Semantic Segmentation.
- Universitad Polytecnical de Catalunia – Collaborative research with Francesc Moreno. Collaboration Topic: Deep Learning for Image Descriptors.
- Ecole Polytechnique Federale de Lausanne (EPFL) – Collaborative research with Eduard Trulls. Collaboration Topic: Deep Learning for Image Descriptors.
- University of California at Los Angeles – Collaborative research with Alan Yuille. Collaboration Topic: Deep Learning for Semantic Segmentation.
- University of Massachusetts at Amherst – Collaborative research with Subhransu Maji. Collaboration Topic: Deep Learning for Texture Recognition.
- Ryerson University – Collaborative research with Kostas Derpanis. Collaboration Topic: Deep Learning for Learning Segmentation.
- University of Pennsylvania - Collaborative research with Aristeidis Sotiras. Collaboration Topic: Higher Order Graphs in biomedical image analysis.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the Organizing Committees
- Pesquet, Jean-Christophe (co-organizer): Special session on Advanced Optimization Methods for Online Signal Processing at the European Signal Processing Conference (EUSIPCO) 2016

9.1.2. Scientific Events Selection

9.1.2.1. Chair of Conference Program Committees
- Kokkinos, Iasonas: European Conference on Computer Vision (ECCV) 2016, Asian Conference on Computer Vision (ACCV), IEEE Conference on Computer Vision and Pattern Recognition (CVPR)

9.1.2.2. Member of the Conference Program Committees
- Pesquet, Jean-Christophe: European Signal Processing Conference (EUSIPCO), IEEE International Conference on Image Processing (ICIP)

9.1.2.3. Reviewer

The members of the team reviewed numerous papers for several international conferences, such as for the annual conferences on Computer Vision and Pattern Recognition (CVPR), Medical Image Computing and Computer Assisted Intervention (MICCAI), Neural Information Processing Systems (NIPS) and International Conference on Learning Representations (ICLR), IEEE International Conference and Acoustics Speech and Signal Processing (ICASSP), IEEE International Conference on Image Processing (ICIP), IEEE Statistical Signal Processing workshop (SSP), European Signal Processing Conference.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Paragios, Nikos: Medical Image Analysis Journal (MedIA), SIAM Journal on Imaging Sciences
- Kokkinos, Iasonas: Image and Vision Computing Journal (IVC), Computer Vision and Image Understanding Journal (CVIU), Special Issue on Deep Learning for Computer Vision (guest editor)
- Zacharaki, Evangelia: Medical Physics (guest editor), International Journal of Radiology, Dataset Papers in Science (Radiology)

9.1.3.2. Reviewer - Reviewing Activities

- Pesquet, Jean-Christophe: IEEE Trans. on Signal Processing, IEEE Trans. on Image Processing, IEEE Trans. on Information Theory (IEEE-TI), Signal Processing, SIAM Journal on Optimization, SIAM Journal on Imaging Sciences, Journal of Mathematical Imaging and Vision, Journal of Optimization Theory and Applications
- Kokkinos, Iasonas: International Journal of Computer Vision, IEEE Trans. on Pattern Analysis and Machine Intelligence, Computer Vision and Image Understanding (CVIU)
- Chouzenoux, Emilie: IEEE Trans. on Image Processing, IEEE Trans. Image Processing, Journal of Mathematical Imaging and Vision (JMIV), Journal of Optimization Theory and Applications, Journal of Global Optimization
- Zacharaki, Evangelia: IEEE Trans. on Medical Imaging (T-MI), Medical Image Analysis (MedIA), Trans. on Biomedical Engineering, Neuroimage, Artificial Intelligence in Medicine, Expert Systems with Applications
- Ferrante, Enzo: IEEE Trans. on Medical Imaging (T-MI), Medical Image Analysis (MedIA), Computarized Medical Imaging and Graphics (CMIG)

9.1.4. Invited Talks

- Pesquet, Jean-Christophe: Modena (Optimization Techniques for Inverse Problems workshop), Mathematics Faculty of Wien University,
  Polytechnic University of Warsaw.
- Kokkinos, Iasonas: Local features workshop held in conjunction with ECCV, October 2016 (keynote speech),
  Astronomical Data Analysis Summer School, Chania, Greece, May 2016 (keynote speech),
  September 2016, Qualcomm-UvA Deep Vision Seminar
  Perceptual Organization in Computer Vision Workshop (CVPR), Jun 2016,
  Facebook Artificial Intelligence Research (FAIR) Paris, May 2016,
  ETH,May 2016,
  Oxford University, March 2016,
  Univerity College London, March 2016,
  Center for Machine Perception, Prague, March 2016,
  Simon Fraser University, February 2016.
- Chouzenoux, Emilie: Séminaire Parisien des Mathématiques Appliquées à l'Imagerie, IHP, Paris, November 2016,
  Cavalieri workshop on Optimization and Optimal Transport for Imaging, Inria Paris, October 2016.
- Zacharaki, Evangelia: LIGM, Université Paris Est, November 2016

9.2. Teaching - Supervision - Juries

Masters

Kokkinos, Iasonas

- Master: Machine Learning for Computer Vision, 24, M2, Ecole Normale Superieure-Cachan, FR
- Master: Introduction to Deep Learning, 24, M2, CentraleSupélec, FR

Zacharaki, Evangelia

- Master: Foundations in Machine Learning, 36, M2 DataScience, Centrale-Supélec , FR

Corbineau, Marie-Caroline and Pesquet, Jean-Christophe

- Master: Advanced course on Optimization, 30h, M1, CentraleSupélec,FR

Chouzenoux Emilie and Pesquet, Jean-Christophe

- Master: Foundations of Distributed and Large Scale Computing, 24h, M2 Data-Science, CentraleSupélec,FR

9.2.1. Supervision

PhD in progress : Eugène Belilovsky, Structured Output Prediction on Large Scale Neuroscience Data, Universite Paris-Saclay & KU Leuven, 2014-2017, M. Blaschko

PhD in progress : Jiaqian Yu, Structured Prediction Methods for Computer Vision and Medical Imaging, Universite Paris-Saclay, 2014-2017, M. Blaschko

PhD in progress : Wacha Bounliphone, Statistical tools for Imaging-Genetics data integration, 2013-2016, Universite Paris-Saclay & KU Leuven, M. Blaschko & A. Tenenhaus

PhD in progress : Diane Bouchacourt, Large Scale Diverse Learning for Structured Output Prediction, 2014-2017, M. Pawan Kumar

PhD in progress: Siddhartha Chandra, Efficient Learning and Optimization for 3D Visual Data, 2013-2016, I. Kokkinos, Pawan Kumar

PhD in progress: Stefan Kinauer, Surface-based representations for high-level vision tasks, 2013-2016, I. Kokkinos.

PhD in progress: Alp Guler, Learning 3D representations for high-level vision, 2016-2018, I. Kokkinos.

PhD in progress : Stavros Alchatzidis, Message Passing Methods, Parallel Architectures & Visual Processing, 2011-2014 (extended), Nikos Paragios

PhD in progress : Enzo Ferrante, 2D-to-3D Multi-Modal Deformable Image Fusion, 2012-2015 (extended), N. Paragios

PhD in progress : Vivien Fecamp, Linear-Deformable Multi-Modal Deformable Image Fusion, 2012-2015 (extended), N. Paragios

PhD in progress : Evgenios Kornaropoulos, Diffusion Coefficient: a novel computer aided biomarker, 2013-2016, N. Paragios

PhD in progress : Maxim Berman, Learning Higher Order Graphical Models, 2014-2017, N. Paragios, I. Kokkinos

PhD in progress : Hariprasad Kannan, Efficient Inference on Higher Order Graphs, 2014-2017, N. Paragios

PhD in progress : Huu Dien Khue Le, Graph-based Visual Perception : Theories and Applications, 2014-2017, N. Paragios

PhD in progress : Marie-Caroline Corbineau, Fast online optimization algorithms for machine learning and medical imaging, 2016-2019, supervised by Emilie Chouzenoux and J.-C. Pesquet

PhD in progress : Loubna El Gueddari, Parallel proximal algorithms for compressed sensing MRI reconstruction - Applications to ultra-high magnetic field imaging, 2016-2019, supervised by Emilie Chouzenoux and J.-C. Pesquet

PhD in progress : Azar Louzi, Fast online optimization algorithms for machine learning and computer vision, 2016-2019, supervised by Emilie Chouzenoux and J.-C. Pesquet

9.2.2. Juries

The faculty members of the team (N. Paragios, J.-C. Pesquet, I. Kokkinos, E. Chouzenoux) participated in several PhD Thesis Committees, HDR Committees and served as Grant Reviewers.

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] F. ABBOUD, E. CHOUZENOUX, J.-C. PESQUET, J.-H. CHENOT, L. LABORELLI.*Dual Block Coordinate Forward-Backward Algorithm with Application to Deconvolution and Deinterlacing of Video Sequences*, in "Journal of Mathematical Imaging and Vision", 2016, https://hal.archives-ouvertes.fr/hal-01418393.

[2] S. ALCHATZIDIS, A. SOTIRAS, E. I. ZACHARAKI, N. PARAGIOS.*A Discrete MRF Framework for Integrated Multi-Atlas Registration and Segmentation*, in "International Journal of Computer Vision", 2016 [*DOI :* 10.1007/S11263-016-0925-2], https://hal.archives-ouvertes.fr/hal-01359094.

[3] E. CHOUZENOUX, J.-C. PESQUET.*Convergence Rate Analysis of the Majorize-Minimize Subspace Algorithm*, in "IEEE Signal Processing Letters", 2016, vol. 23, n° 9, p. 1284 - 1288 [*DOI :* 10.1109/LSP.2016.2593589], https://hal.archives-ouvertes.fr/hal-01373641.

[4] M. CIMPOI, S. MAJI, I. KOKKINOS, A. VEDALDI. *Deep filter banks for texture recognition, description, and segmentation* , in "International Journal of Computer Vision", January 2016, https://hal.inria.fr/hal-01263622.

[5] R. GADDE, R. MARLET, N. PARAGIOS.*Learning Grammars for Architecture-Specific Facade Parsing*, in "International Journal of Computer Vision", 2016 [*DOI :* 10.1007/s11263-016-0887-4], https://hal.inria.fr/hal-01429246.

[6] V. G. KANAS, E. I. ZACHARAKI, G. A. THOMAS, P. O. ZINN, V. MEGALOOIKONOMOU, R. R. COLEN.*Learning MRI-based classification models for MGMT methylation status prediction in glioblastoma*, in "Computer Methods and Programs in Biomedicine", January 2017 [*DOI :* 10.1016/J.CMPB.2016.12.018], https://hal.inria.fr/hal-01423323.

[7] N. KOMODAKIS, M. P. KUMAR, N. PARAGIOS.*Hyper-Graphs Inference through Convex Relaxations and Move Making Algorithms: Contributions and Applications in Artificial Vision*, in "Foundations and Trends in Computer Graphics and Vision", May 2016, vol. 10, n° 1, p. 1-102 [*DOI :* 10.1561/0600000066], https://hal.archives-ouvertes.fr/hal-01429215.

[8] N. PARAGIOS, E. FERRANTE, B. GLOCKER, N. KOMODAKIS, S. PARISOT, E. I. ZACHARAKI.*(Hyper)-Graphical Models in Biomedical Image Analysis*, in "Medical Image Analysis", 2016 [*DOI :* 10.1016/J.MEDIA.2016.06.028], https://hal.archives-ouvertes.fr/hal-01359107.

[9] E. PIPPA, V. G. KANAS, E. I. ZACHARAKI, V. TSIRKA, M. KOUTROUMANIDIS, V. MEGALOOIKONOMOU.*EEG-based Classification of Epileptic and Non-epileptic Events using Multi-array Decomposition*, in "International Journal of Monitoring and Surveillance Technologies Research", January 2017, https://hal.archives-ouvertes.fr/hal-01359125.

[10] E. PIPPA, E. I. ZACHARAKI, M. KOUTROUMANIDIS, V. MEGALOOIKONOMOU.*Data fusion for paroxysmal events' classification from EEG*, in "Journal of Neuroscience Methods", January 2017, vol. 275, p. 55-65 [*DOI :* 10.1016/J.JNEUMETH.2016.10.004], https://hal.inria.fr/hal-01426373.

[11] M. VAKALOPOULOU, K. KARANTZALOS, N. KOMODAKIS, N. PARAGIOS.*Graph-Based Registration, Change Detection, and Classification in Very High Resolution Multitemporal Remote Sensing Data*, in "IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing", 2016, vol. 9, p. 2940 - 2951 [*DOI :* 10.1109/JSTARS.2016.2557081], https://hal.inria.fr/hal-01413419.

[12] E. I. ZACHARAKI, I. MPORAS, K. GARGANIS, V. MEGALOOIKONOMOU.*Spike pattern recognition by supervised classification in low dimensional embedding space*, in "Brain Informatics", 2016 [*DOI :* 10.1007/s40708-016-0044-4], https://hal.archives-ouvertes.fr/hal-01359155.

[13] Y. ZENG, C. WANG, X. GU, D. SAMARAS, N. PARAGIOS.*Higher-Order Graph Principles towards Non-Rigid Surface Registration*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", December 2016, vol. 38, n° 12, p. 2416 - 2429 [*DOI :* 10.1109/TPAMI.2016.2528240], https://hal.archives-ouvertes.fr/hal-01429216.

International Conferences with Proceedings

[14] A. AMIDI, S. AMIDI, D. VLACHAKIS, N. PARAGIOS, E. I. ZACHARAKI.*A Machine Learning Methodology for Enzyme Functional Classification Combining Structural and Protein Sequence Descriptors*, in "Bioinformatics and Biomedical Engineering", Granada, Spain, April 2016, p. 728-738 [*DOI :* 10.1007/978-3-319-31744-1_63], https://hal.archives-ouvertes.fr/hal-01359157.

[15] E. BELILOVSKY, G. VAROQUAUX, M. B. BLASCHKO.*Testing for Differences in Gaussian Graphical Models: Applications to Brain Connectivity*, in "Neural Information Processing Systems (NIPS) 2016", Barcelona, Spain, December 2016, https://hal.inria.fr/hal-01248844.

[16] S. CADONI, E. CHOUZENOUX, J.-C. PESQUET, C. CHAUX.*A block parallel majorize-minimize memory gradient algorithm*, in "IEEE International Conference on Image Processing (ICIP 2016)", Phoenix, AZ, United States, September 2016, p. 3194 - 3198 [*DOI :* 10.1109/ICIP.2016.7532949], https://hal.archives-ouvertes.fr/hal-01418417.

[17] S. CHANDRA, I. KOKKINOS.*Fast, Exact and Multi-Scale Inference for Semantic Image Segmentation with Deep Gaussian CRFs*, in "European Conference on Computer Vision", Amsterdam, Netherlands, September 2016 [*DOI :* 10.5244/C.29.187], https://hal.inria.fr/hal-01410872.

[18] A. CHERNI, E. CHOUZENOUX, M.-A. DELSUC.*Proximity Operators for a Class of Hybrid Sparsity+Entropy Priors. Application to DOSY NMR Signal Reconstruction*, in "International Symposium on Signal, Image, Video and Communications (ISIVC 2016)", Tunis, Tunisia, November 2016, vol. In Proceedings of the International Symposium on Signal, Image, Video and Communications (ISIVC 2016), https://hal.archives-ouvertes.fr/hal-01421741.

[19] G. CHIERCHIA, N. PUSTELNIK, J.-C. PESQUET.*Random primal-dual proximal iterations for sparse multiclass SVM*, in "IEEE International Workshop on Machine Learning for Signal Processing", Vietri sul Mare, Italy, September 2016 [*DOI :* 10.1109/MLSP.2016.7738833], https://hal.archives-ouvertes.fr/hal-01419697.

[20] P. L. COMBETTES, J.-C. PESQUET.*Stochastic forward-backward and primal-dual approximation algorithms with application to online image restoration*, in "European Signal and Image Processing Conference (EUSIPCO 2016)", Budapest, Hungary, August 2016, p. 1813 - 1817 [*DOI :* 10.1109/EUSIPCO.2016.7760561], https://hal.archives-ouvertes.fr/hal-01422154.

[21] A. GULER, N. KARDARIS, S. CHANDRA, V. PITSIKALIS, C. WERNER, K. HAUER, C. TZSAFESTAS, P. MARAGOS, I. KOKKINOS.*Human Joint Angle Estimation and Gesture Recognition for Assistive Robotic Vision*, in "ACVR, ECCV", Amsterdam, Netherlands, G. HUA, H. JÉGOU (editors), LNCS - Lecture Notes in Computer Science, Springer, October 2016, vol. 9914, p. 415 - 431 [*DOI : 10.1007/978-3-319-48881-3_29*], https://hal.inria.fr/hal-01410854.

[22] S. KINAUER, M. BERMAN, I. KOKKINOS.*Monocular Surface Reconstruction using 3D Deformable Part Models*, in ""Geometry Meets Deep Learning" in ECCV 2016", Amsterdam, Netherlands, ECCV 2016 Workshops, Part III, LNCS 9915 proceedings, October 2016, vol. 9915, p. 296-308 [*DOI : 10.1007/978-3-319-49409-8_24*], https://hal.inria.fr/hal-01416479.

[23] E. N. KORNAROPOULOS, E. I. ZACHARAKI, P. ZERBIB, C. LIN, A. RAHMOUNI, N. PARAGIOS.*Deformable group-wise registration using a physiological model: application to DIffusion-Weighted MRI*, in "IEEE International Conference on Image Processing (ICIP)", Phoenix, Arizona, United States, Deformable group-wise registration using a physiological model: application to DIffusion-Weighted MRI, September 2016, https://hal.archives-ouvertes.fr/hal-01324238.

[24] E. N. KORNAROPOULOS, E. I. ZACHARAKI, P. ZERBIB, C. LIN, A. RAHMOUNI, N. PARAGIOS.*Optimal estimation of diffusion in DW-MRI by high-order MRF-based joint deformable registration and diffusion modeling*, in "WBIR 2016 - 7th International Workshop on Biomedical Image Registration", Las Vegas, Nevada, United States, Optimal estimation of diffusion in DW-MRI by high-order MRF-based joint deformable registration and diffusion modeling, July 2016, https://hal.archives-ouvertes.fr/hal-01324229.

[25] A. OSOKIN, J.-B. ALAYRAC, I. LUKASEWITZ, P. K. DOKANIA, S. LACOSTE-JULIEN.*Minding the Gaps for Block Frank-Wolfe Optimization of Structured SVMs*, in "International Conference on Machine Learning (ICML 2016)", New York, United States, 2016, Appears in Proceedings of the 33rd International Conference on Machine Learning (ICML 2016). 31 pages, https://hal.archives-ouvertes.fr/hal-01323727.

[26] R. PORCHETTO, F. STRAMANA, E. FERRANTE, N. PARAGIOS.*Rigid Slice-To-Volume Medical Image Registration through Markov Random Fields*, in "Bayesian and Graphical Models for Biomedical Imaging Workshop, BAMBI (MICCAI 2016, Athens, Greece)", Athens, Greece, October 2016, https://hal.inria.fr/hal-01355110.

[27] M. SHAKERI, E. FERRANTE, S. TSOGKAS, S. LIPPE, S. KADOURY, I. KOKKINOS, N. PARAGIOS.*Prior-based Coregistration and Cosegmentation*, in "MICCAI 2016", Athens, Greece, 2016, The first two authors contributed equally, https://hal.inria.fr/hal-01349189.

[28] M. SHAKERI, S. TSOGKAS, E. FERRANTE, S. LIPPE, S. KADOURY, N. PARAGIOS, I. KOKKINOS.*Sub-cortical brain structure segmentation using F-CNN's*, in "ISBI 2016: International Symposium on Biomedical Imaging", Prague, Czech Republic, 2016, https://hal.inria.fr/hal-01265500.

[29] A. TZALAVRA, K. DALAKLEIDI, E. I. ZACHARAKI, N. TSIAPARAS, F. CONSTANTINIDIS, N. PARAGIOS, K. S. NIKITA.*Comparison of Multi-resolution Analysis Patterns for Texture Classification of Breast Tumors Based On DCE-MRI*, in "7th Int. Workshop on Machine Learning in Medical Imaging (MICCAI workshop)", Athens, Greece, October 2016, vol. 10019, p. 296-304 [*DOI : 10.1007/978-3-319-47157-0_36*], https://hal.archives-ouvertes.fr/hal-01359118.

[30] M. VAKALOPOULOU, C. PLATIAS, M. PAPADOMANOLAKI, N. PARAGIOS, K. KARANTZA-LOS.*Simultaneous registration, segmentation and change detection from multisensor, multitemporal*

satellite image pairs, in "International Geoscience and Remote Sensing Symposium (IGARSS)", Beijing, China, July 2016 [*DOI :* 10.1109/IGARSS.2016.7729469], https://hal.inria.fr/hal-01413373.

[31] J. B. Yu, M. Blaschko.*A Convex Surrogate Operator for General Non-Modular Loss Functions*, in "The 25th Belgian-Dutch Conference on Machine Learning", Kortrijk, Belgium, September 2016, https://hal.inria. fr/hal-01417108.

[32] J. Yu, M. Blaschko.*A Convex Surrogate Operator for General Non-Modular Loss Functions*, in "The 19th International Conference on Artificial Intelligence and Statistics", Cadiz, Spain, Proceedings of the 19th International Conference on Artificial Intelligence and Statistics, May 2016, https://hal.inria.fr/hal-01299519.

[33] J. Yu, M. Blaschko.*Efficient Learning for Discriminative Segmentation with Supermodular Losses*, in "British Machine Vision Conference", York, United Kingdom, Proceedings of the British Machine Vision Conference, September 2016, https://hal.inria.fr/hal-01349000.

Conferences without Proceedings

[34] S. Chandra, G. Chrysos, I. Kokkinos.*Surface Based Object Detection in RGBD Images*, in "British Machine Vision Conference", Swansea, Wales, United Kingdom, September 2016 [*DOI :* 10.5244/C.29.187], https://hal.inria.fr/hal-01263930.

Research Reports

[35] A. Cherni, E. Chouzenoux, M.-A. Delsuc.*PALMA, an improved algorithm for DOSY signal processing*, LIGM - Laboratoire d'Informatique Gaspard-Monge, 2016, https://hal.archives-ouvertes.fr/hal-01421737.

[36] Y. Marnissi, Y. Zheng, E. Chouzenoux, J.-C. Pesquet.*A Variational Bayesian Approach for Restoring Data Corrupted with Non-Gaussian Noise*, Laboratoire Informatique Gaspard Monge, 2016, https://hal. archives-ouvertes.fr/hal-01418399.

Other Publications

[37] E. Belilovsky, K. Kastner, G. Varoquaux, M. Blaschko.*Learning to Discover Graphical Model Structures*, May 2016, working paper or preprint, https://hal.inria.fr/hal-01306491.

[38] W. Bounliphone, M. B. Blaschko.*A U-statistic Approach to Hypothesis Testing for Structure Discovery in Undirected Graphical Models*, April 2016, working paper or preprint, https://hal.inria.fr/hal-01298279.

References in notes

[39] A. Cherni, E. Chouzenoux, M.-A. Delsuc.*Proximity Operators for a Class of Hybrid Sparsity+Entropy Priors. Application to DOSY NMR Signal Reconstruction*, in "International Symposium on Signal, Image, Video and Communications (ISIVC 2016)", Tunisia, 2016.

[40] M. Shakeri, S. Tsogkas, E. Ferrante, S. Lippe, S. Kadoury, N. Paragios, I. Kokkinos.*Sub-cortical brain structure segmentation using F-CNN's*, in "ISBI 2016: International Symposium on Biomedical Imaging", Prague, Czech Republic, 2016, https://hal.inria.fr/hal-01265500.

Project-Team GAMMA3

Automatic mesh generation and advanced methods

IN PARTNERSHIP WITH:
Université de Technologie de Troyes

RESEARCH CENTER
Saclay - Île-de-France

THEME
Numerical schemes and simulations

Table of contents

Project-Team GAMMA3

Creation of the Project-Team: 2010 January 01

Keywords:

Computer Science and Digital Science:

2.5. - Software engineering
6.1. - Mathematical Modeling
6.2. - Scientific Computing, Numerical Analysis & Optimization
7.1. - Parallel and distributed algorithms
7.5. - Geometry, Topology

Other Research Topics and Application Domains:

5.2.3. - Aviation
5.2.4. - Aerospace

1. Members

Research Scientists

Paul Louis George [Team leader, Inria, Senior Researcher]
Frederic Alauzet [Inria, Researcher, HDR]
Patrick Laug [Inria, Researcher, HDR]
Adrien Loseille [Inria, Researcher]

Faculty Members

Houman Borouchaki [UT Troyes, Professor]
Dominique Barchiesi [UT Troyes, Professor]
Abel Cherouat [UT Troyes, Professor]
Thomas Grosges [UT Troyes, Associate Professor]
David Marcum [MSU, Professor]
Laurence Moreau [UT Troyes, Associate Professor]

Technical Staff

Alexis Loyer [Inria]
Slimane Amara [UT Troyes]

PhD Students

Bastien Andrieu [ONERA]
Rémi Feuillet [ENSTA, from Apr 2016]
Loïc Frazza [Ecole Polytechnique]
Eléonore Gauci [Inria]
Aichun Zhu [UT Troyes, until Feb 2016]

Post-Doctoral Fellow

Olivier Coulaud [Inria]

Administrative Assistants

Celine Fortabat [Inria, Jan 2016]
Emmanuelle Perrot [Inria]

Others

Loic Marechal [Engineer]
Clément Roge [Inria, Stagiaire, from Jun 2016 until Jul 2016]

2. Overall Objectives

2.1. Introduction

Une branche importante des sciences de l'ingénieur s'intéresse aux calculs des solutions d'équations aux dérivées partielles très variées (en mécanique du solide, en mécanique des fluides, en modélisation de problèmes thermiques, ...) par la méthode des éléments ou des volumes finis. Ces méthodes utilisent comme support spatial des calculs un maillage du domaine sur lequel les équations sont formulées. Par suite, les algorithmes (de construction) de maillages occupent un rôle important dans toute simulation par la méthode des éléments ou des volumes finis d'un problème modélisé en équations aux dérivées partielles. En particulier, la précision, voire la validité, des solutions calculées est liée aux propriétés du maillage utilisé [22].

L'équipe-projet GAMMA3 a été créé en 2010 à la suite du projet GAMMA. L'équipe est bilocalisée avec une partie à l'UTT (Troyes) et l'autre à Rocquencourt puis Saclay. Les thèmes du projet regroupent un ensemble d'activités concernant les points indiqués ci-dessus, en particulier, l'aspect génération automatique de maillages afin de construire les supports utilisés par les méthodes d'éléments ou de volumes finis. Sont également étudiés les aspects de modélisation géométrique, de post-traitement et de visualisation des résultats issus de tels calculs [23].

L'évolution de la demande en termes de génération automatique de maillages implique une évolution des méthodes classiques de création de maillages vers des méthodes permettant de construire des maillages contrôlés. Les maillages doivent donc être soit isotropes, le contrôle portant sur des tailles souhaitées, soit anisotropes, le contrôle portant à la fois sur des directions et des tailles selon ces dernières.

Le développement d'algorithmes de maillages gouvernés sert de support naturel à la conception de boucles de maillages adaptatifs qui, via un estimateur d'erreurs *a posteriori*, permettent de contrôler la qualité des solutions. Les estimateurs d'erreurs sont issus d'applications en mécanique des fluides (Inria) et du solide (UTT). Leurs validations reposent sur le développement de solveurs avancés, en particulier, dans ces disciplines. Ces deux points (estimateurs et solveurs) constituent au moins la moitié de nos recherches.

Ces préoccupations amènent à considérer le problème du maillage des domaines de calculs en eux-mêmes tout comme celui du maillage ou du remaillage des courbes et surfaces, frontières de ces domaines.

La taille, en termes de nombre de nœuds, des maillages nécessaires pour certaines simulations, amène à travailler sur la parallélisation des processus de calculs. Cette problématique conduit également à s'intéresser à l'aspect multi-cœurs au niveau des algorithmes de maillages proprement dits.

Simultanément, le volume des résultats obtenus dans de telles simulations, nécessite d'envisager le post-traitement de ces résultats en parallèle ou par des méthodes appropriées.

Par ailleurs, de nombreux problèmes partent de saisies *scanner* (ou autre système discret) des géométries à traiter et demandent d'en déduire des maillages de surfaces aptes à être, par la suite, traités par les méthodes classiques (de remaillage, d'optimisation, de calculs).

Enfin, la maturité de certaines méthodes (victimes de leur succès) conduit les utilisateurs à demander plus et à considérer des problèmes de maillage ou des conditions d'utilisations extrêmes induisant des algorithmes *a priori* inattendus.

Les objectifs du projet GAMMA3 consistent à étudier l'ensemble des points mentionnés ci-dessus afin de rendre automatique le calcul de la solution d'un problème donné avec une précision imposée au départ. Par ailleurs, certaines des techniques utilisées dans les problématiques de maillage sont utilisables dans d'autres disciplines (compression d'images pour ne citer qu'un seul exemple).

3. New Software and Platforms

3.1. ABL4FLO

FUNCTIONAL DESCRIPTION

KEYWORDS: Boundary layer - Hybrid meshes

SCIENTIFIC DESCRIPTION

Automatic boundary layer mesh generation for complex geometries

FUNCTIONAL DESCRIPTION

ABL4FLO is designed to generate 3D adapted boundary layer meshes by using a cavity-based operator.

- Participant: Adrien Loseille
- Contact: Adrien Loseille
- URL: https://www.rocq.inria.fr/gamma/Adrien.Loseille/index.php?page=softwares

3.2. AMA4FLO

FUNCTIONAL DESCRIPTION

KEYWORDS: Anisotropic mesh adaptation - Surface and volume remeshing - Non manifold geometries

SCIENTIFIC DESCRIPTION

Robust and automatic generation of anisotropic meshes in 3D

FUNCTIONAL DESCRIPTION

AMA4FLO is designed to generate adapted meshes with respect to a provided anisotropic sizing field. The surface and the volume mesh is adapted simultaneously to guarantee that a 3D valid mesh is provided on output.

- Participant: Adrien Loseille
- Contact: Adrien Loseille
- URL: https://www.rocq.inria.fr/gamma/Adrien.Loseille/index.php?page=softwares

3.3. BL2D

KEYWORDS: Automatic mesher - Delaunay - Anisotropic - Planar domain

SCIENTIFIC DESCRIPTION

The meshing method is of controlled Delaunay type, isotropic or anisotropic. The internal point generation follows an advancing-front logic, and their connection is realized as in a classical Delaunay approach. Quadrilaterals are obtained by a pairing process. The direct construction of degree 2 elements has been made possible via the control of the domain boundary mesh, in order to ensure the desired compatibility.

FUNCTIONAL DESCRIPTION

Planar mesh generator (isotropic or anisotropic, adaptive).

- Participants: Houman Borouchaki and Patrick Laug
- Contact: Patrick Laug
- URL: https://www.rocq.inria.fr/gamma/Patrick.Laug/logiciels/logiciels.html

3.4. BL2D-ABAQ

KEYWORDS: Automatic mesher - Delaunay - Anisotropic - Planar domain - Error estimation - Interpolation

SCIENTIFIC DESCRIPTION

The meshing method is the same as BL2D (see above) in an adaptive process. An *a posteriori* error estimation of a solution at the nodes of the current mesh results in a size map. A new mesh satisfying these size specifications (made continuous) is built, and the solution is interpolated on the new mesh.

FUNCTIONAL DESCRIPTION

Planar mesh generator (isotropic or anisotropic, adaptive) for deformable domains, interacting with the ABAQUS solver.

- Participants: Houman Borouchaki, Patrick Laug and Abel Cherouat
- Contact: Patrick Laug
- URL: https://www.rocq.inria.fr/gamma/Patrick.Laug/logiciels/logiciels.html

3.5. BLGEOL

KEYWORDS: Automatic mesher - Hex-dominant - Geologic structures

SCIENTIFIC DESCRIPTION

The aim is to generate hex-dominant meshes of geologic structures complying with different geometric constraints: surface topography (valleys, reliefs, rivers), geologic layers and underground workings. First, a reference 2D domain is obtained by projecting all the line constraints into a horizontal plane. Different size specifications are given for rivers, outcrop lines and workings. 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, taking into account the surfaces found (interfaces between two layers, top or bottom faces of underground workings). The generation of volume elements follows a global order established on the whole set of surfaces to ensure the conformity of the resulting mesh.

FUNCTIONAL DESCRIPTION

Hex-dominant mesher of geologic structures and storage facilities.

- Participants: Patrick Laug and Houman Borouchaki
- Contact: Patrick Laug
- URL: https://www.rocq.inria.fr/gamma/Patrick.Laug/logiciels/logiciels.html

3.6. BLMOL

KEYWORDS: Automatic mesher - Molecular surface

SCIENTIFIC DESCRIPTION

To model a molecular surface, each constituting atom is idealized by a simple sphere. First, a boundary representation (B-rep) of the surface is obtained, i.e. a set of patches and the topological relations between them. Second, an appropriate parametrization and a metric map are computed for each patch. Third, meshes of the parametric domains are generated with respect to an induced metric map, using a combined advancing-front generalized-Delaunay approach. Finally these meshes are mapped onto the entire surface.

FUNCTIONAL DESCRIPTION

Molecular surface mesher.

- Participants: Houman Borouchaki and Patrick Laug
- Contact: Patrick Laug
- URL: https://www.rocq.inria.fr/gamma/Patrick.Laug/logiciels/logiciels.html

3.7. BLSURF

KEYWORDS: Automatic mesher - Parametric surface - CAD surface

SCIENTIFIC DESCRIPTION

An indirect method for meshing parametric surfaces conforming to a user-specifiable size map is used. First, from this size specification, a Riemannian metric is defined so that the desired mesh should have unit length edges with respect to the related Riemannian space (the so-called "unit mesh"). Then, based on the intrinsic properties of the surface, the Riemannian structure is induced into the parametric space. Finally, a unit mesh is generated completely inside the parametric space such that it conforms to the metric of the induced Riemannian structure. This mesh is constructed using a combined advancing-front Delaunay approach applied within a Riemannian context.

FUNCTIONAL DESCRIPTION

CAD surface mesher.

- Participants: Houman Borouchaki and Patrick Laug
- Contact: Patrick Laug
- URL: https://www.rocq.inria.fr/gamma/Patrick.Laug/logiciels/logiciels.html

3.8. FEFLOA-REMESH

KEYWORDS: Anisotropic mesh adaptation - Surface remeshing - Cavity-based operator
SCIENTIFIC DESCRIPTION

Automatic generation of metric-aligned and metric-orthogonal anisotropic meshes in 3D

FUNCTIONAL DESCRIPTION

FEFLOA-REMESH is intended to generate adapted 2D, surface and volume meshes by using a unique cavity-based operator. The metric-aligned or metric-orthogonal approach is used to generate high quality surface and volume meshes independently of the anisotropy involved.

- Participants: Adrien Loseille and Frédéric Alauzet
- Contact: Adrien Loseille
- URL: https://www.rocq.inria.fr/gamma/Adrien.Loseille/index.php?page=softwares

3.9. GAMANIC 3D

KEYWORDS: Tetrahedral mesh - Delaunay - Anisotropic size and direction control - Automatic Mesher
SCIENTIFIC DESCRIPTION

Automatic tetrahedral mesher based on an anisotropic Delaunay type point insertion method. A metric field is provided specifying the desired size (edge length) and directional properties.

FUNCTIONAL DESCRIPTION

GAMANIC3D is a volume mesher governed by a (anisotropic) size and directional specification metric field.

- Participants: Houman Borouchaki, Paul Louis George, Frederic Hecht, Eric Saltel, Frédéric Alauzet and Adrien Loseille
- Contact: Paul Louis George
- URL: http://www.meshgems.com/volume-meshing.html

3.10. GAMHIC 3D

KEYWORDS: Tetrahedral mesh - Delaunay - Isotropic size control - Automatic Mesher
SCIENTIFIC DESCRIPTION

Automatic tetrahedral mesher based on the Delaunay point insertion method. A metric field is provided specifying the desired size (edge length).

FUNCTIONAL DESCRIPTION

GAMHIC3D is a volume mesher governed by a (isotropic) size specification metric field.

- Participants: Houman Borouchaki, Paul Louis George, Frederic Hecht, Eric Saltel, Frédéric Alauzet and Adrien Loseille
- Contact: Paul Louis George
- URL: http://www.meshgems.com/volume-meshing.html

3.11. GHS3D

KEYWORDS: Tetrahedral mesh - Delaunay - Automatic Mesher
SCIENTIFIC DESCRIPTION

Automatic tetrahedral mesher based on the Delaunay point insertion method.

FUNCTIONAL DESCRIPTION

GHS3D is an automatic volume mesher

- Participants: Paul Louis George, Houman Borouchaki, Eric Saltel, Frédéric Alauzet, Adrien Loseille and Frederic Hecht
- Contact: Paul Louis George
- URL: http://www.meshgems.com/volume-meshing.html

3.12. HEXOTIC

KEYWORDS: Hexahedral mesh - Octree - Automatic mesher
SCIENTIFIC DESCRIPTION

Automatic full hexahedral mesher primarily based on an octree.

FUNCTIONAL DESCRIPTION

HEXOTIC is an automatic hexahedral mesher

- Contact: Loïc Maréchal
- URL: https://www.rocq.inria.fr/gamma/gamma/Membres/CIPD/Loic.Marechal/Research/Hexotic.html

3.13. Metrix

KEYWORD: Scientific calculation
SCIENTIFIC DESCRIPTION

Compute a metric field from a given solution field using various error estimates.

FUNCTIONAL DESCRIPTION

Metrix computes metric field from a given solution field using various error estimates. Available error estimates are feature-based and goal-oriented based error estimates for steady or unsteady fields. Metrix also performs operations on metrics: gradation, intersection, natural metric of a mesh.

- Participants: Frédéric Alauzet and Adrien Loseille
- Contact: Frédéric Alauzet
- URL: https://www.rocq.inria.fr/gamma/Frederic.Alauzet/code_eng.html

3.14. Nimbus 3D

KEYWORDS: Surface reconstruction - Point cloud
SCIENTIFIC DESCRIPTION

Given a point cloud, a surface is constructed primarily based on a Delaunay approach.

FUNCTIONAL DESCRIPTION

Nimbus3D is a surface reconstruction method piece of software

- Participants: Paul Louis George and Houman Borouchaki
- Contact: Paul Louis George
- URL: http://www.meshgems.com/volume-meshing.html

3.15. VIZIR

KEYWORDS: Mesh and solution visualization
SCIENTIFIC DESCRIPTION

Interactive mesh and solution visualization for linear, and high order curved elements

FUNCTIONAL DESCRIPTION

VIZIR is intended to visualize and modify interactively simplicial, hybrid and high order curved meshes.

- Participants: Julien Castelneau, Adrien Loseille and Alexis Loyer
- Contact: Adrien Loseille
- URL: http://www-roc.inria.fr/gamma/gamma/vizir/

3.16. Wolf

KEYWORD: Scientific calculation
SCIENTIFIC DESCRIPTION

General solver platform containing all the Wolf modules: Wolf-Bloom, Wolf-Elast, Wolf-Interpol, Wolf-MovMsh, Wolf-Nsc, Wolf-Shrimp, Wolf-Spyder and Wolf-Xfem, for details see thereafter.

FUNCTIONAL DESCRIPTION

Wolf is a general solver platform containing all the Wolf modules: Wolf-Bloom, Wolf-Elast, Wolf-Interpol, Wolf-MovMsh, Wolf-Nsc, Wolf-Shrimp, Wolf-Spyder and Wolf-Xfem.

- Participants: Frédéric Alauzet and Adrien Loseille
- Contact: Frédéric Alauzet
- URL: https://www.rocq.inria.fr/gamma/Frederic.Alauzet/code_eng.html

3.17. Wolf-Bloom

KEYWORD: Scientific calculation
SCIENTIFIC DESCRIPTION

Structured boundary layer mesh generator using a pushing approach.

FUNCTIONAL DESCRIPTION

Wolf-Bloom is a structured boundary layer mesh generator using a pushing approach. It start from an existing volume mesh and insert a structured boundary layer by pushing the volume mesh. The volume mesh deformation is solved with an elasticity analogy. Mesh-connectivity optimizations are performed to control volume mesh element quality.

- Participants: Frédéric Alauzet, Adrien Loseille and Dave Marcum
- Contact: Frédéric Alauzet
- URL: https://www.rocq.inria.fr/gamma/Frederic.Alauzet/code_eng.html

3.18. Wolf-Elast

KEYWORD: Scientific calculation

SCIENTIFIC DESCRIPTION

Linear elasticity solver using a P1 Finite-Element method.

FUNCTIONAL DESCRIPTION

Wolf-Elast is a linear elasticity solver using a P1 Finite-Element method. The Young and Poisson coefficient can be parametrized. The linear system is solved using the Conjugate Gradient method with the LUSGS preconditioner.

- Participants: Frédéric Alauzet and Adrien Loseille
- Contact: Frédéric Alauzet
- URL: https://www.rocq.inria.fr/gamma/Frederic.Alauzet/code_eng.html

3.19. Wolf-Interpol

KEYWORD: Scientific calculation
SCIENTIFIC DESCRIPTION

Software transferring scalar, vector and tensor fields from one mesh to another one.

FUNCTIONAL DESCRIPTION

Wolf-Interpol is a tool to transfer scalar, vector and tensor fields from a mesh to another one. Polynomial interpolation (from order 2 to 4) or conservative interpolation operators can be used. Wolf-Interpol also extract solutions along lines or surfaces.

- Participants: Frédéric Alauzet and Adrien Loseille
- Contact: Frédéric Alauzet
- URL: https://www.rocq.inria.fr/gamma/Frederic.Alauzet/code_eng.html

3.20. Wolf-MovMsh

KEYWORD: Scientific calculation
SCIENTIFIC DESCRIPTION

Moving mesh algorithm coupled with mesh-connectivity optimization.

FUNCTIONAL DESCRIPTION

Wolf-MovMsh is a moving mesh algorithm coupled with mesh-connectivity optimization. Mesh deformation is computed by means of a linear elasticity solver or a RBF interpolation. Smoothing and swapping mesh optimization are performed to maintain good mesh quality. It handles rigid or deformable bodies, and also rigid or deformable regions of the domain.

- Participants: Frédéric Alauzet and Adrien Loseille
- Contact: Frédéric Alauzet
- URL: https://www.rocq.inria.fr/gamma/Frederic.Alauzet/code_eng.html

3.21. Wolf-Nsc

KEYWORD: Scientific calculation
SCIENTIFIC DESCRIPTION

Numerical flow solver solving the compressible Navier-Stokes equations.

FUNCTIONAL DESCRIPTION

Wolf-Nsc is numerical flow solver solving steady or unsteady turbulent compressible Euler and Navier-Stokes equations. The available turbulent models are the Spalart-Almaras and the Menter SST k-omega. A mixed finite volume - finite element numerical method is used for the discretization. Second order spatial accuracy is reached thanks to MUSCL type methods. Explicit or implicit time integration are available. It also resolved dual (adjoint) problem and compute error estimate for mesh adaptation.

- Participants: Frédéric Alauzet and Adrien Loseille
- Contact: Frédéric Alauzet
- URL: https://www.rocq.inria.fr/gamma/Frederic.Alauzet/code_eng.html

3.22. Wolf-Shrimp

KEYWORD: Scientific calculation
SCIENTIFIC DESCRIPTION

Mesh partitioner for parallel mesh generation and parallel computation.

FUNCTIONAL DESCRIPTION

Wolf-Shrimp is a generic mesh partitioner for parallel mesh generation and parallel computation. It can partition planar, surface (manifold and non manifold), and volume domain. Several partitioning methods are available: Hilbert-based, BFS, BFS with restart. It can work with or without weight function and can correct the partitions to have only one connected component.

- Participants: Frédéric Alauzet and Adrien Loseille
- Contact: Frédéric Alauzet
- URL: https://www.rocq.inria.fr/gamma/Frederic.Alauzet/code_eng.html

3.23. Wolf-Spyder

KEYWORD: Scientific calculation
SCIENTIFIC DESCRIPTION

Metric-based mesh quality optimizer using vertex smoothing and edge/face swapping.

FUNCTIONAL DESCRIPTION

Wolf-Spyder is a metric-based mesh quality optimizer using vertex smoothing and edge/face swapping.

- Participants: Frédéric Alauzet and Adrien Loseille
- Contact: Frédéric Alauzet
- URL: https://www.rocq.inria.fr/gamma/Frederic.Alauzet/code_eng.html

3.24. Wolf-Xfem

KEYWORD: Scientific calculation
SCIENTIFIC DESCRIPTION

Tool providing the mesh of the intersection between a surface mesh and a volume mesh in the goal of simulating mechanical fractures.

FUNCTIONAL DESCRIPTION

Wolf-Xfem is a tool providing the mesh of the intersection between a surface mesh and a volume mesh.

- Participants: Frédéric Alauzet
- Contact: Frédéric Alauzet
- URL: https://www.rocq.inria.fr/gamma/Frederic.Alauzet/code_eng.html

4. New Results

4.1. Remaillage adaptatif pour la mise en forme de tôles minces et de composites

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

Au cours des simulations numériques de mise en forme en 3D, les grandes déformations mises en jeu font que le maillage subit de fortes distorsions. Il est alors nécessaire de remailler continuellement la pièce afin de pouvoir capturer les détails géométriques des surface en contact, adapter la taille du maillage à la solution physique et surtout pouvoir effectuer la simulation jusqu'à la fin du procédé de mise en forme. Lorsque la pièce est comprise entre des outils rigides (cas de l'emboutissage), aux problèmes de remaillage s'ajoutent aussi des difficultés sur la gestion du contact entre les pièces. Une méthode couplant une stratégie de remaillage adaptatif et une technique de projection a été développée. Afin de pouvoir réaliser des simulations numériques de composites tissés, une procédure spécifique a été ajoutée au remailleur afin de pouvoir raffiner les éléments finis bi-composants (association d'éléments finis de barre et de membrane orientés matérialisant le comportement de fibres chaîne et trame).

Ce travail a donnée lieu à 1 article.

4.2. Le formage incrémental : étude expérimentale, numérique et remaillage adaptatif

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

Le formage incrémental est un procédé de mise en forme récent permettant de mettre en forme des tôles minces grâce au déplacement d'un outil hémisphérique dont la trajectoire est pilotée par une machine à commande numérique. Ce procédé peu couteux est une alternative intéressante à l'emboutissage traditionnel pour les entreprises réalisant des pièces de petite taille à usage unique ou en petite série comme les entreprises biomédicales (prothèses, implants personnalisés..). Cependant, il reste encore des développements importants sur le plan numérique et expérimental pour que ce procédé soit industrialisable : problèmes d'état de surface, de non-respect de la géométrie, risques de rupture. Nous avons étudié numériquement et expérimentalement ce procédé de formage incrémental : développement d'une méthode de remaillage adaptée à ce procédé, optimisation des paramètres du procédé, étude du formage incrémental à chaud, étude du formage incrémental robotisé.r

Ce travail a donnée lieu à 2 articles et 5 participations à des conférences internationales.

4.3. Reconstruction de surface 3D à partir d'images numériques 2D

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

Ces travaux portent sur la reconstruction 3D d'objets à partir de plusieurs photos prises via des caméras calibrées avec des points de vue différents couvrant la totalité de la surface de l'objet. La méthodologie générale consiste à apparier les pixels correspondants de deux photos et obtenir des positions 3D via une technique de triangulation.. L'idée originale réside dans une nouvelle méthodologie automatique d'appariement de pixels. Elle comprend trois étapes : un motif présentant un maillage triangulaire aléatoire est projeté sur l'objet 3D, le maillage est identifié sur chaque photo et la technique de triangulation est appliquée aux sommets de ce maillage. La méthodologie de reconstruction 3D a été appliquée à la modélisation géométrique du buste féminin afin d'envisager des simulations de comportements statique et dynamique de ce buste. Ces travaux ont conduit aussi à la conception et la réalisation d'une cabine d'acquisition permettant de prendre 24 prises de vue de manière simultanée depuis un ordinateur extérieur à la cabine.

Ce travail a donnée lieu à 1 article et 2 participations à des conférences internationales.

4.4. Modélisation numérique, remaillage adaptatif et optimisation pour la morphologie de nanofils

Participants: Laurence Moreau [correspondant], Thomas Grosges.

L'objectif était de développer une méthode permettant de détecter et d'analyser la présence de nanomatériaux dans l'eau. Une voie possible consiste à étudier les effets liés aux couplages lumière-matière, c'est-à-dire la réponse photo-thermique des nanomatériaux illuminés par une onde électromagnétique. La méthode proposée consiste à étudier la réponse thermique du nanofil immergé sous l'illumination et à la relier à la bulle produite. Le problème multi physique est modélisé par un système d'équations couplées : équation de Helmoltz et équation de la chaleur. La résolution numérique de ces équations est effectuée par une méthode des éléments finis et un processus d'optimisation incluant des boucles de remaillages adaptatifs afin de contrôler la précision de la solution et assurer la convergence. Une étude de la morphologie de la bulle a été réalisée en fonction de paramètres géométriques et physiques. Deux fonctions permettant de relier la taille de la bulle à la taille et la forme du nanomatériau ont été définies. La résolution du modèle inverse, associé à ces fonctions, permettant de remonter à la morphologie du nanomatériau via celle de la bulle. L'efficacité et la pertinence du modèle ont été montrées en confrontant les résultats numériques aux résultats expérimentaux.

Ce travail a donnée lieu à 3 articles et 2 participations à des conférences internationales

4.5. Les outils de remaillage dans la simulation multi-physiques pour la fiabilisation des systèmes complexes

Participants: Abel Cherouat [correspondant], Houman Borouchaki.

Le projet concerne la maîtrise des outils de simulation numérique multi-physique avec remaillage adaptatif 3D pour la prévention de la fiabilité des systèmes complexes. Les systèmes étudiés sont des structures comportant des composants et des architectures mécaniques. Ils sont fortement contraints car la partie électro-magnétique est très sensible aux vibrations, variations et dilatations thermiques, et agressions physico-chimiques qui existent habituellement dans les systèmes mécaniques. La fiabilisation représente des enjeux majeurs pour ces systèmes. L'objectif final est d'étudier la fiabilisation de ces systèmes par des approches hybrides qui combinent les outils de simulation éléments finis multi-physiques couplées avec adaptation en temps réel des maillages éléments finis en 3D. L'analyse de la fiabilité et la synthèse pouvant être appliquées en cas de défaillance pour maintenir l'exploitation des systèmes.

Ce travail a donnée lieu à 4 articles et 2 participations à des conférences internationales.

4.6. Les matériaux innovants : mousses métalliques - AMF, textiles techniques, composites et agro-composites : Modélisation mécanique, Simulation avec remaillage, Reconstitution 3D et Modélisation géométrique

Participants: Abel Cherouat [correspondant], Houman Borouchaki, Shijie Zhu, Antony Sheedev.

Le contexte de l'étude sur les mousses est la modélisation du comportement mécanique, la reconstitution 3D de la morphologie des mousses à partir d'images tomographiques ou de la CAO géométrique, de l'optimisation et de la simulation de la déformation de mousse (métalliques ou AMF).

Le contexte de l'étude sur les agro-composites est la maitrise des matières naturelles, l'allégement des structures et la valorisation de l'émergence des textiles biodégradables pour des applications industrielles. Les investigations concernent les aspects d'élaboration et mise en œuvre des textiles secs ou pré-imprégnés (tissés, UD cousu et mats), de caractérisation-modélisation comportementale multi-échelle et de mise au point d'outils d'aide à la décision et d'éco-conception des matériaux fonctionnels.

Le contexte de l'étude sur les composites est l'éco-réparation *in-situ* des structures industrielles intégrant l'hybridation de procédés émergents d'*additive manufacturing* et le frittage micro-onde avec l'utilisation de nouvelles résines ou nuances de matériaux, la numérisation 3D, l'impression ou collage par balayage et le contrôle non destructif.

Le contexte de l'étude sur les tissus biologiques est le développement de méthode d'obtention des paramètres mécaniques des tissus vivants et des informations pour l'amélioration des prothèses post-chirurgicale (un sein artificiel). Une approche médicale de la modélisation du sein et de sa déformabilité a pour objectif de prédire les déformations des tissus pendant les interventions en tenant compte des constituants (graisses, glandes, peau et ligaments), mais ne concerne pas le comportement du sein et son remodelage par le bonnet ou son comportement pendant le sport.

Ce travail a donnée lieu à 11 articles et 8 participations à des conférences internationales.

4.7. Reconstruction 3D à partir d'image vs Scanner 3D, Maillage adaptatif par vision embarquée sur drones autonomes

Participants: Abel Cherouat [correspondant], Houman Borouchaki.

Dans le cadre de ce projet, on se propose de concevoir un système de reconstruction adaptative et temps réel de scènes 3D en se basant uniquement sur le flux d'images captées par une caméra embarquée sur un drone autonome. Un nuage de points peut être ainsi obtenu en traitant d'une manière efficace et temps-réel le flux d'images issues de la caméra mobile. Le nuage de points en temps réel est utilisé pour reconstruire les surfaces des objets constituant la scène, et surtout de quantifier la qualité de la reconstruction en fonction de la géométrie de ces surfaces. Les applications concernées sont les automates industriels, l'imagerie médicale, le *Smart Tracking*, la surveillance et la sécurité, la rétro-conception et la réalité augmentée, ...

Ce travail a donnée lieu à 2 articles et 2 participations à des conférences internationales.

4.8. Les outils de remaillage dans la simulation et l'optimisation de la mise en forme des matériaux

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

L'objectif scientifique de ce projet est de développer des modèles théoriques, numériques et géométriques nécessaires à la mise au point de méthodologies de simulation numérique et d'optimisation de procédés de fabrication et de mise en forme de composants et de structures mécaniques en petites ou en grandes déformations. Une attention particulière est accordée à la génération de maillage, de remaillage et de maillage adaptatif isotrope et anisotrope plan (2D), surfacique (2,5D) et volumique (3D), ainsi que des méthodes d'optimisation de maillages (en particulier surfacique) ainsi que les couplages multi-physiques entre les différents phénomènes.

Ce travail a donnée lieu à 13 articles et 4 participations à des conférences internationales.

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

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

L'utilisation des nombres (pseudo)-aléatoires a pris une dimension importante ces dernières décennies. De nombreuses applications dans le domaine des télécommunications, de la cryptographie, des simulations numériques ou encore des jeux de hasard, ont contribué au développement et à l'usage de ces nombres. Les méthodes utilisées pour la génération de tels nombres (pseudo)-aléatoires proviennent de deux types de processus : physique et algorithmique. Ce projet de recherche a donc pour objectif principal le développement de nouveaux procédés de génération de clés de chiffrement, dits "exotiques", basés sur des processus physiques, multi-échelles, multi-domaines assurant un niveau élevé de sécurité. Deux classes de générateurs basés sur des principes de mesures physiques et des processus mathématiques ont été développé.

La première classe de générateurs exploite la réponse d'un système physique servant de source pour la génération des séquences aléatoires. Cette classe utilise aussi bien des résultats de simulations 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). À 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. Les résultats de cette recherche se sont vu récompensés par un premier prix décerné par EURASIP (European Association in Signal Processing) en 2016 ("Best paper award of the EURASIP).

4.10. Méthodes avancées pour 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: 2011-2015.
Production scientifique: 1 thèse soutenue en 2016 (A. Chaari), 3 articles publiés, 1 conférence nternationale (PIERS 2014), 2 conférences nationales (CSMA 2013 et CSMA 2015).

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).

4.11. 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: 2009-2016 (thèse de Fadhil Mezghani initiée en 2012 coencadrée par D. Barchiesi et A. Cherouat).
Production scientifique: 2 thèses soutenues (S. Kessentini, 22/10/2012 et F. Mezghani), 15 articles publiés, 6 conférences.

Le contrôle et l'adaptation du maillage lors de la résolution de problèmes couplés et/ou 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. En ce sens, l'étude permet d'aller plus loin que la "simple" interaction électromagnétique avec la matière du projet européen *Nanoantenna*.

Le travail a constitué en la poursuite de l'étude des spécificités de ce type de problème multiphysique pour des structures de forme simple et la mise en place de fonctions test, de référence, pour les développements de maillage adaptatifs pour les modèles multiphysiques éléments finis. Nous espérons pouvoir proposer un projet ANR couplant les points de vue microscopiques et macroscopiques dans les prochaines années.

4.12. Problèmes de magnétostatique sur maillage de grande taille et multi-échelle

Participants: Dominique Barchiesi [correspondant], Thomas Grosges, Houman Borouchaki, Brahim Yahiaoui
Validité du projet: 2013-2015. Post-doc Brahin Yahiaoui.
Le projet Flyprod concerne l'étude du stockage d'électricité par volant d'inertie lévité et financé par l'ADEME. Une technologie brevetée innovante et stratégique permettant à des acteurs majeurs de la distribution électrique de stocker de l'énergie pour des périodes de fortes consommations. D'un point de vue écologique, un volant d'inertie n'émet ni gaz à effet de serre, ni produits chimiques nocifs pour l'environnement. Les partenaires pour ce projet sont LEVISYS, Université de Technologie de Troyes, SCLE SFE (COFELY INEO, Groupe GDFSUEZ), CIRTEM, Conseil Général de l'Aube. Les dispositifs mis en oeuvre nécessitent des études approfondies pour rendre les volants d'inertie économiquement viables. La recherche a consisté à développer un programme informatique permettant une simulation assistée par ordinateur. Il permet plus précisément de calculer les champs magnétiques et de concevoir les pièces du volant d'inertie afin de garantir une perte minimale d'énergie. Le champ magnétique doit être calculé en un temps raisonnable sur des distances spatiales réduites. L'approche utilisée pour répondre à ces objectifs est appliquée sur un maillage fourni par le logiciel Optiform (un remailleur adaptatif volumique développé par l'équipe GAMMA3). Les résultats obtenus ont permis d'optimiser la structure du volant d'inertie et d'atteindre une efficacité de stockage de 97%, permettant de valider la pertinence du volant et de confirmer sa fabrication.

4.13. Element metric, element quality and interpolation error metric

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

The metric of a simplex of $\mathbb{R}^d$ is a metric tensor (symmetric positive definite matrix) in which the element is unity (regular with unit edge lengths). This notion is related to the problem of interpolation error of a given field over a mesh. Let K be a simplex and let us denote by v_{ij} the vector joining vertex i and vertex j of K. The metric of K can be written as:

$$\mathcal{M} = \frac{d+1}{2} \left(\sum_{i<j} v_{ij}{}^t v_{ij} \right)^{-1},$$

where $v_{ij}{}^t v_{ij}$ is a $d \times d$ rank 1 matrix related to edge ij.

The metric of a simplex also characterizes the element shape. In particular, if it is the identity, the element is unity. Hence, to define the shape quality of an element, one can determine the gap of the element metric $\mathcal{M}$ and the identity using different measures based on the eigenvalues $\lambda_i = \frac{1}{h_i^2}$ of $\mathcal{M}$ or those of $\mathcal{M}^{-1}$, e.g. h_i^2. Notice that metric $\mathcal{M}^{-1}$ is directly related to the geometry of the element (edge length, facet area, element volume). The first algebraic shape quality measure ranging from 0 to 1 is defined as the ratio of the geometric average of the eigenvalues of $\mathcal{M}^{-1}$ and their arithmetic average:

$$q(K) = \frac{\left(\prod_i h_i^2 \right)^{\frac{1}{d}}}{\dfrac{1}{d} \displaystyle\sum_{i=1}^{d} h_i^2} = d \, \frac{\left(det(\mathcal{M}^{-1}) \right)^{\frac{1}{d}}}{tr(\mathcal{M}^{-1})}.$$

As the geometric average is smaller than the arithmetic average, this measure is well defined. In addition, it is the algebraic reading of the well-known quality measure defined by:

$$q^{\frac{d}{2}}(K) = (d!) d^{\frac{d}{2}} (d+1)^{\frac{d-1}{2}} \frac{|K|}{\left(\displaystyle\sum_{i<j} l_{ij}^2 \right)^{\frac{d}{2}}},$$

where the volume and the square of the edge lengths are involved. The algebraic meaning justifies the above geometric measure. The second algebraic shape quality measure is defined as the ratio of the harmonic average of the eigenvalues of $\mathcal{M}^{-1}$ and their arithmetic average (ranging also from 0 to 1):

$$q(K) = \frac{\left\{ \dfrac{1}{d} \displaystyle\sum_{i=1}^{d} \dfrac{1}{h_i^2} \right\}^{-1}}{\dfrac{1}{d} \displaystyle\sum_{i=1}^{d} h_i^2} = \frac{d^2}{tr(\mathcal{M}) tr(\mathcal{M}^{-1})}.$$

As above, this measure is well defined, the harmonic average being smaller the arithmetic one. From this measure, one can derive another well-known measure involving the roundness and the size of an element (measure which is widely used for convergene issues in finite element methods).

Note that these measures use the invariants of $\mathcal{M}^{-1}$ or $\mathcal{M}$ and thus can be evaluated from the coefficients of the characteristical polynomial of those matrices (avoiding the effective calculation of their eigenvalues). Another advantage of the above algebraic shape measures is their easy extensions in an arbitrary Euclidean space. Indeed, if $\mathcal{E}$ is the metric of such a space, the algebraic shape measures read:

$$q_{\mathcal{E}}(K) = d \, \frac{\left(det(\mathcal{M}^{-1}\mathcal{E}) \right)^{\frac{1}{d}}}{tr(\mathcal{M}^{-1}\mathcal{E})} \quad , \quad q_{\mathcal{E}}(K) = \frac{d^2}{tr(\mathcal{E}^{-1}\mathcal{M}) tr(\mathcal{M}^{-1}\mathcal{E})}.$$

This work has been published in a journal, [8].

Following this notion of a element metric, a natural work was done regarding how to define the element metric so as to achieve a given accuracy for the interpolation error of a function using a finite element approximation by means of simplices of arbitrary degree.

This is a new approach for the majoration of the interpolation error of a polynomial function of arbitrary degree n interpolated by a polynomial function of degree $n-1$. From that results a metric, the so-called interpolation metric, which allows for a control of the error. The method is based on the geometric and algebraic properties of the metric of a given element, metric in which the element is regular and unit. The interpolation metric plays an important role in advanced computations based on mesh adaptation. The method relies in a Bezier reading of the functions combined with Taylor expansions. In this way, the error in a given element is fully controled at the time the edges of the element are controled.

It is shown that the error in bounded as

$$|e(X)| \leq C \sum_{i<j} f^{(n)}(.)(v_{ij}, v_{ij}, ..., v_{ij}),$$

where C is a constant depending on d and n, v_{ij} is the edge from the vertices of K of index i and j, $f^{(n)}(.)$ is the derivative of order n of f applied to a n-uple uniquely composed of v_{ij}. If we consider the case $d=2$ and $u = (x, y)$ is a vector in $\mathbb{R}^2$, we have

$$f^{(n)}(.)(u, u, ..., u) = \sum_{i=0}^{n-2} x^{n-2-i} y^i \, {}^t u \left(C_i^{n-2} \mathcal{H}_{(n-2, n-2-i, i)} \right) u \ ,$$

where the quadratic forms $\mathcal{H}_{(n-2, n-2-i, i)}$ are defined by the matrices of order 2 (with constant entries):

$$\mathcal{H}_{(n-2, n-2-i, i)} = \begin{pmatrix} \frac{\partial^{(n)} f}{\partial x_1^{n-i} \partial x_2^i} & \frac{\partial^{(n)} f}{\partial x_1^{n-1-i} \partial x_2^{i+1}} \\ \\ \frac{\partial^{(n)} f}{\partial x_1^{n-1-i} \partial x_2^{i+1}} & \frac{\partial^{(n)} f}{\partial x_1^{n-2-i} \partial x_2^{i+2}} \end{pmatrix} ,$$

those matrices being the hessians of the derivatives of f of order $n-2$.

This work resulted in a paper submitted in a journal and currently under revision.

4.14. Realistic modeling of fractured geologic media

Participants: Patrick Laug [correspondant], Géraldine Pichot.

This study, in collaboration with project-team Serena, aims to model, in a realistic and efficient manner, natural fractured media. These media are characterized by their diversity of structures and organizations. Numerous studies in the past decades have evidenced the existence of characteristic structures at multiple scales. At fracture scale, the aperture distribution is widely correlated and heterogeneous. At network scale, the topology is complex resulting from mutual mechanical interactions as well as from major stresses. Geometric modeling of fractured networks combines in a non-standard way a large number of 2D fractures interconnected in the 3D space. Intricate local configurations of fracture intersections require original methods of geometric modeling and mesh generation. We have developed in 2016 a software package that automatically builds geometric models and surface meshes of random fracture networks. The results are highly promising and we now want to continue this research to further improve the element quality in complex configurations, take into account multiple size scales in large fracture networks (up to thousands of fractures), and compare several modeling strategies (mixed hybrid finite elements, projected grids, mortar elements) [13].

4.15. Parallel meshing of surfaces defined by collections of connected regions

Participant: Patrick Laug [correspondant].

In CAD (computer aided design) environments, a surface is commonly modeled as a collection of connected regions represented by parametric mappings. For meshing such a composite surface, a parallelized indirect approach with dynamic load balancing can be used on a shared memory system. However, this methodology can be inefficient in practice because most existing CAD systems use memory caches that are only appropriate to a sequential process. As part of the sabbatical year of P. Laug at Polytechnique Montréal in 2014/2015, two solutions have been proposed, referred to as the Pirate approach and the Discrete approach. In the first approach, the Pirate library can be efficiently called in parallel since no caching is used for the storage or evaluation of geometric primitives. In the second approach, the CAD environment is replaced by internal procedures interpolating a discrete geometric support. In 2016, the dynamic load balancing has been analyzed and improved. Significant modifications to the Pirate library have been made, and new numerical tests on three different computers (4, 8 and 64 cores) have been carried out, now showing an almost linear scaling of the method in all cases [10].

4.16. Discrete CAD model for visualization and meshing

Participants: Patrick Laug [correspondant], Houman Borouchaki.

During the design of an object using a CAD (computer aided design) platform, the user can visualize the ongoing model at every moment. Visualization is based on a discrete representation of the model that coexists with the exact analytical representation of the object. Most CAD systems have this discrete representation available, and each of them applies its own construction methodology. We have developed in 2016 a method to build a discrete model for CAD surfaces (the model is quadtree-based and subdivided into quadrilaterals and triangles). The method presents two major particularities: most elements are aligned with iso-parametric curves and the accuracy of the surface approximation is controlled. In addition, we have proposed a new technique of surface mesh generation that is based on this discrete model. This approach has been implemented as a part of a surface mesher called ALIEN, and several examples have demonstrate the robustness and computational efficiency of the program, as well as the quality of the geometric support [14], [15].

4.17. Visualization and modification of high-order curved meshes

Participants: Alexis Loyer, Dave Marcum, Adrien Loseille [correspondant].

During the partnership between Inria and Distene, a new visualization software has been designed. It address the typical operations that are required to quickly assess the newly algorithm developed in the team. In particular, interactive modifications of high-order curved mesh and hybrid meshes has been addressed. The software VIZIR is freely available at https://www.rocq.inria.fr/gamma/gamma/vizir/.

4.18. Adaptation de maillages pour des écoulements visqueux en turbomachine

Participants: Frédéric Alauzet, Loïc Frazza, Adrien Loseille [correspondant].

4.18.1. Calcul.

Les prémices d'une adaptation pour les écoulements Navier-Stokes turbulents ont été testés sur des calculs de turbomachine. Pour ce faire nous avons tout d'abord traité les particularités liées aux calculs en turbomachine: - Les aubes présentent en général une périodicité par rotation et on ne simule donc qu'une période afin d'alléger les calculs. Il faut donc traiter cette périodicité de façon appropriée dans le code CFD et l'adaptation de maillage. - Afin de prendre en compte la rotation des pales sans employer de maillages mobiles et simulations instationnaires on peut se placer dans le référentiel tournant de l'aube en corrigeant les équations. - Les écoulements en turbomachine sont des écoulements clos, les conditions limites d'entrée et de sortie ont donc une influence très forte et peuvent de plus se trouver très près de la turbine afin de simuler la présence d'autres étages en amont ou aval. Des conditions limites bien précises ont donc été développées afin de traiter correctement ces effets.

4.18.2. Adaptation.

Pour l'adaptation de maillages deux particularités doivent être traitées ici, la périodicité du maillage et la couche limite turbulente.

En 2D, la couche limite turbulente est automatiquement adaptée avec la méthode metric orthogonal et la périodicité du maillage est garantie par un traitement spécial des frontières. Les estimateurs d'erreurs Navier-Stokes et RANS n'étant pas encore au point nous avons utilisé la Hessienne du Mach de l'écoulement comme senseur ce qui donne déjà des résultats satisfaisants.

En 3D la méthode metric orthogonal est beaucoup plus complexe à mettre en oeuvre et n'est pas encore au point. La couche limite a donc été exclue de l'adaptation, le maillage est adapté uniquement dans le volume en utilisant la Hessienne du Mach de l'écoulement comme senseur. La périodicité n'étant pas traitée non plus, les frontières périodiques restent inchangées ce qui garantie leur périodicité.

4.18.3. Norm-Oriented.

Dans le cadre de la théorie Norm-Oriented, afin de contrôler l'erreur implicite des schémas numériques, un correcteur a été développé et testé. Etant donné un maillage et la solution numérique obtenue avec, le résidu de cette solution projeté sur un maillage deux fois plus fin est accumulé sur le maillage initial. Ce défaut de résidu est utilisé comme terme source dans une seconde simulation plus courte. La nouvelle solution toujours sur le meme maillage est plus proche de la solution exacte et donne une bonne estimation de l'erreur.

4.19. Metric-orthogonal and metric-aligned mesh adaptation

Participants: Frédéric Alauzet, Loïc Frazza, Adrien Loseille, Dave Marcum [correspondant].

A new algorithm to derive adaptive meshes has been introduced through new cavity-based algorithms. It allows to generate anisotropic surface and volume mesh that are aligned along the eigenvector directions. This allows us to improve the quality of the meshes and to deal naturally with boundary layer mesh generation.

Die orthogonale metric Methode erzeugt 2D-Elemente mit einem Rand, der mit der Hauptrichtung der Metrik ausgerichtet ist und einem zweiten Rand, der rechtwinklig zur ersten ist. Das erzeugende Gitter is so örtlich strukturiert wo es Anisotropie gibt. Dieses Methode wurde erfolgreich zur automatischen strukturierten Gitter Erzeugung in der turbulenten Grenzschichten für Turbomaschinen Simulationen angewendet.

4.20. Parallel mesh adaptation

Participants: Frédéric Alauzet, Adrien Loseille [correspondant].

We devise a strategy in order to generate large-size adapted anisotropic meshes $O(10^8 - 10^9)$ as required in many fields of application in scientific computing. We target moderate scale parallel computational resources as typically found in R&D units where the number of cores ranges in $O(10^2 - 10^3)$. Both distributed and shared memory architectures are handled. Our strategy is based on hierarchical domain splitting algorithm to remesh the partitions in parallel. Both the volume and the surface mesh are adapted simultaneously and the efficiency of the method is independent of the complexity of the geometry. The originality of the method relies on (i) a metric-based static load-balancing, (ii) dedicated hierarchical mesh partitioning techniques to (re)split the (complex) interfaces meshes, (iii) anisotropic Delaunay cavity to define the interface meshes, (iv) a fast, robust and generic sequential cavity-based mesh modification kernel, and (v) out-of-core storing of completing parts to reduce the memory footprint. We are able to generate (uniform, isotropic and anisotropic) meshes with more than 1 billion tetrahedra in less than 20 minutes on 120 cores [11].

4.21. Unsteady adjoint computation on dynamic meshes

Participants: Eléonore Gauci, Frédéric Alauzet [correspondant].

Adjoint formulations for unsteady problems are less common due to the extra complexity inherent in the numerical solution and storage but these methods are a great option in engineering because it takes more into account the cost function we want to minimize. Moreover the engineering applications involve moving bodies and this motion must be taken into account by the governing flow equations. We develop a model of unsteady adjoint solver on moving mesh problems. The derivation of the adjoint formulation based on the ALE form of the equations requires consideration of the dynamic meshes. Our model takes into account the DGCL.

4.22. Line solver for efficient stiff parse system resolution

Participants: Loïc Frazza, Frédéric Alauzet [correspondant].

Afin d'accélérer la résolution des problèmes raides, un line-solver á été développé. Cette méthode extrait tout d'abord des lignes dans le maillage du problème selon des critères géométriques ou physiques. Le problème peut alors être résolu exactement le long des ces lignes à moindre coût. Cette méthode est particulièrement bien adaptée aux cas où l'information se propage selon une direction privilégiée tels que les chocs, les couches limites ou les sillages. Ces cas sont généralement associés à des maillages très étirés ce qui conduit à des problèmes raides mais quasi-unidimensionnels. Ils peuvent donc être résolu efficacement par un line-solver, réduisant ainsi les temps de calculs tout en gagnant en robustesse.

4.23. Error estimate for high-order solution field

Participants: Olivier Coulaud, Adrien Loseille [correspondant].

Afin de produire des solveurs d'ordre élevé, et ainsi répondre aux exigences inhérentes à la résolution de problèmes physiques complexes, nous développons une méthode d'adaptation de maillage d'ordre élevé. Celle-ci est basée sur le contrôle par une métrique de l'erreur d'interpolation induite par le maillage du domaine. Plus précisément, pour une solution donnée, l'erreur d'interpolation d'ordre k est paramétrée par la forme différentielle $(k + 1)^{\text{ième}}$ de cette solution, et le problème se réduit à trouver la plus grande ellipse incluse dans une ligne de niveau de cette différentielle. La méthode que nous avons mise au point théoriquement et numériquement est appelée "log-simplexe", et permet de produire des maillages adaptés d'ordre élevé dans un temps raisonnable, et ce en dimension 2 et 3. À l'occasion de l'International Meshing Roundtable 2016, ce travail a été présenté et publié. D'autres applications de cette méthode sont en cours d'exploitation, comme par exemple la génération de maillages adaptés courbes de surface, ou le couplage avec un solveur d'ordre élevé.

4.24. Méthode d'immersion de frontières pour la mécanique des fluides

Participants: Frédéric Alauzet [correspondant], Rémi Feuillet, Adrien Loseille.

Dans les méthodes de résolution classiques des problèmes d'interaction fluide-structure, il est usuel de représenter l'objet de manière exacte dans le maillage, c'est-à-dire avec des éléments conformes à l'objet : le maillage possède des triangles dont une arête correspond avec le bord de la géométrie immergée. Cette méthode quoique plus précise est très coûteuse en préprocessing. C'est dans ce cadre qu'est introduite la notion d'immersion de frontière (embedded geometry en anglais). Cette méthode consiste à représenter la géométrie de manière fictive. Le maillage de calcul n'est de fait plus nécessairement conforme à la géométrie de l'objet. Il s'agit donc de s'intéresser aux modifications nécessaires sur les méthodes classiques pour faire un calcul dans le cadre de l'immersion de frontières. Cela concerne les conditions aux limites et l'avancée en temps. On s'intéresse également à l'adaptation de maillage pour le cas de l'immersion. La finalité de tout ce travail est d'effectuer des calculs de coefficients aérodynamiques (portance, traînée) et de trouver des résultats du même ordre de précision que ceux en géométrie inscrite.

4.25. Optimisation de formes et CAO

Participants: Frédéric Alauzet [correspondant], Jean de Becdelièvre, Adrien Loseille.

Pour ce stage de 3 mois, l'objectif était de réaliser entièrement une optimisation aérodynamique, de la génération des modèles 3D aux calculs de la forme optimisée. Le modèle choisi était l'aile du C.R.M. (Common Research Model) de la NASA qui a été extensivement testé en soufflerie. Durant la première phase du projet, l'outils EGADS (Engineering Geometry Aircraft Design System) développé par le Aerospace Computational Design Lab (M.I.T) a été utilisé pour générer des modèles 3D paramétriques. À cette occasion, un outil facilement réutilisable de génération de modèle d'aile a été développé, ainsi que des outils de modification des modèles C.A.D. sous EGADS. Les maillages surfaciques de ces modèles ont été créés par EGADS directement et modifiés immédiatement par AMG pour les adapter au calcul. Les maillages volumiques ont, eux, été générés par GHS3D. Des calculs non visqueux sur des maillages adaptés ont alors permis d'obtenir des résultats, et de répéter l'opération jusqu'à obtenir un minimum. L'originalité de cette optimisation est que chaque calcul, à chaque itération de l'optimiseur, utilise un maillage adapté à l'aide des solutions des calculs précédents ; ce qui permet de réduire les coûts de calcul et d'augmenter la précision.

4.26. Boundary layer mesh generation

Participants: Frédéric Alauzet [correspondant], Adrien Loseille, Dave Marcum.

A closed advancing-layer method for generating high-aspect-ratio elements in the boundary layer (BL) region has been developed. This approach provides an answer to the mesh generation robustness issue as it starts from an existing valid mesh and always guarantuees a valid mesh in output. And, it handles very efficiently and naturally BL front collisions and it produces a natural smooth anisotropic blending between colliding layers. In addition, it provides a robust strategy to couple unstructured anisotropic mesh adaptation and high-aspect-ratio element pseudo-structured BL meshes. To this end, the mesh deformation is performed using the metric field associated with the given anisotropic meshes to maintain the adaptivity while inflating the BL. This approach utilizes a recently developed connectivity optimization based moving mesh strategy for deforming the volume mesh as the BL is inflated. In regards to the BL mesh generation, it features state-of-art capabilities, including, optimal normal evaluation, normal smoothing, blended BL termination, mixed-elements BL, varying growth rate, and BL imprinting on curved surfaces. Results for typical aerospace configurations are presented to assess the proposed strategy on both simple and complex geometries.

5. Bilateral Contracts and Grants with Industry

5.1. Bilateral Contracts with Industry

- The Boeing Company,
- Safran-Tech,
- Projet Rapid (DGA) avec Lemma.

6. Partnerships and Cooperations

6.1. National Initiatives

6.1.1. ANR

F. Alauzet, N. Barral, V. Menier and A. Loseille are part of the MAIDESC ANR (2013-2015) on mesh adaptation for moving interfaces in CFD.

T. Grosges, D. Barchiesi,A. Cherouat, H. Borouchaki, L. Giraud-Moreau and A. Chaari sont membres de l'ANR NONOMOPRH (2011-2016) sur 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.

6.2. European Initiatives

6.2.1. FP7 & H2020 Projects

- UMRIDA https://sites.google.com/a/numeca.be/umrida/

6.3. International Initiatives

6.3.1. Inria Associate Teams Not Involved in an Inria International Labs

6.3.1.1. AM2NS

Title: Advanced Meshing Methods for Numerical Simulations

International Partner (Institution - Laboratory - Researcher):

Mississippi State University (United States) - Center for Advanced Vehicular Systems - Computational Fluid Dynamics Dept. (CAVS-CFD) - Marcum David

Start year: 2014

See also: https://www.rocq.inria.fr/gamma/gamma/Membres/CIPD/Frederic.Alauzet/ AssociateTeam_AM2NS/AT_am2ns.html

Numerical simulation is now mature and has become an integral part of design in science and engineering applications. Meshing, i.e., discretizing the computational domain, is at the core of the computational pipeline and a key element to significant improvements. The AM2NS Associate Team focus on developing the next generation of automated meshing methods to improve their robustness and the mesh quality to solve the ever increasing complexity of numerical simulations. Four major meshing issues are targeted: (i) more robustness for mesh generation methods in recovering a given data set, (ii) higher quality for anisotropic adapted meshes via constraint alignment, (iii) higher quality for boundary layer meshes near geometry singularities, and (iv) more robustness in handling complex displacement for moving mesh methods. The impact of this collaborative research will be to provide more reliable solution output predictions in an automated manner by using these new meshing methods.

7. Dissemination

7.1. Promoting Scientific Activities

7.1.1. Scientific Events Organisation

7.1.1.1. Member of the Organizing Committees

A. Loseille is a committee member of the International Meshing Roundtable.

7.1.2. Scientific Events Selection

7.1.2.1. Member of the Conference Program Committees

P. Laug has been a Member of the Conference Committee of The 13th ISGG Symposium on Numerical Grid Generation, hosted at The 20th IMACS World Congress, Xiamen, China, 10-14 Dec. 2016.

7.1.2.2. Reviewer

Team members have reviewed papers for the International Meshing Roundtable and MASCOT/ISGG.

7.1.3. Journal

7.1.3.1. Reviewer - Reviewing Activities

Team members have reviewed manuscripts submitted to International Journal for Numerical Methods in Engineering, Journal of Computational Physics, Computers and Structures, Engineering with Computers, MATCOM, SIAM Journal.

7.1.4. Scientific Expertise

P. Laug has been a Member of a Site Visit Committee for an Industrial Research Chair application at the NSERC/CRSNG (Natural Sciences and Engineering Research Council of Canada / Conseil de recherches en sciences naturelles et en génie du Canada), Université Laval, Quebec City, Nov. 2016.

7.2. Teaching - Supervision - Juries

HdR : Laurence Moreau, Méthodes de remaillage et d'optimisation pour la simulation numérique, mémoire d'Habilitation à diriger des Recherches, Université de Technologie de Troyes, 26 Mai 2015.

7.2.1. Juries

- F. Alauzet a été Président du jury d'HDR de G. Puigt (Université de Toulouse)
- P. L. George a été rapporteur du jury de thèse de G. Brèthes (Inria Sophia)
- A. Loseille a été examinateur pour la thèse d'A. Botella (Université de Lorraine)

8. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] F. ALAUZET.*A parallel matrix-free conservative solution interpolation on unstructured tetrahedral meshes*, in "Computer Methods in Applied Mechanics and Engineering", February 2016, vol. 299, 26 [*DOI :* 10.1016/J.CMA.2015.10.012], https://hal.inria.fr/hal-01426163.

[2] F. ALAUZET, B. FABRÈGES, M. A. FERNÁNDEZ, M. LANDAJUELA.*Nitsche-XFEM for the coupling of an incompressible fluid with immersed thin-walled structures*, in "Computer Methods in Applied Mechanics and Engineering", January 2016, vol. 301, p. 300-335 [*DOI :* 10.1016/J.CMA.2015.12.015], https://hal.inria.fr/hal-01149225.

[3] F. ALAUZET, A. LOSEILLE.*A Decade of Progress on Anisotropic Mesh Adaptation for Computational Fluid Dynamics*, in "Computer-Aided Design", March 2016, vol. 72, p. 13–39 [*DOI :* 10.1016/J.CAD.2015.09.005], https://hal.inria.fr/hal-01426159.

[4] S. BADREDDINE, L. GIRAUD-MOREAU, A. CHEROUAT.*Optimization of the single point incremental forming process for titanium sheets by using response surface*, in "MATEC Web of Conferences", 2016, vol. 80, 5, https://hal.archives-ouvertes.fr/hal-01438358.

[5] N. BARRAL, G. OLIVIER, F. ALAUZET.*Time-accurate anisotropic mesh adaptation for three-dimensional time-dependent problems with body-fitted moving geometries*, in "Journal of Computational Physics", February 2017, vol. 331, p. 157–187 [*DOI :* 10.1016/J.JCP.2016.11.029], https://hal.inria.fr/hal-01426156.

[6] F. J. COLAS, M. COTTAT, R. GILLIBERT, N. GUILLOT, N. DJAKER, N. LIDGI-GUIGUI, T. TOURY, D. BARCHIESI, A. TOMA, E. D. FABRIZIO, P. GUCCIARDI, M. LAMY DE LA CHAPELLE.*Red-Shift Effects in Surface Enhanced Raman Spectroscopy: Spectral or Intensity Dependence of the Near-Field?*, in "The Journal of Physical Chemistry C", 2016, vol. 120, n⁰ 25, p. 13675-13683 [*DOI :* 10.1021/ACS.JPCC.6B01492], https://hal.inria.fr/hal-01441504.

[7] O. COULAUD, A. LOSEILLE.*Very High Order Anisotropic Metric-Based Mesh Adaptation in 3D*, in "Procedia Engineering", 2016, vol. 163, p. 353 - 365 [*DOI :* 10.1016/J.PROENG.2016.11.071], https://hal.archives-ouvertes.fr/hal-01438226.

[8] P. L. GEORGE, H. BOROUCHAKI.*Element metric*, in "Comptes Rendus Mathématique", December 2016, vol. 355, p. 105 - 112 [*DOI :* 10.1016/J.CRMA.2016.11.007], https://hal.inria.fr/hal-01431689.

[9] T. GROSGES, D. BARCHIESI.*Numerical Study of Photoacoustic Pressure for Cancer Therapy*, in "Applied Sciences ", November 2016, vol. 6, n⁰ 11, 357 [*DOI :* 10.3390/APP6110357], https://hal.archives-ouvertes.fr/hal-01397515.

[10] P. LAUG, F. GUIBAULT, H. BOROUCHAKI.*Parallel meshing of surfaces represented by collections of connected regions*, in "International Journal of Advances in Engineering Software", 2016 [*DOI :* 10.1016/J.ADVENGSOFT.2016.09.003], https://hal.inria.fr/hal-01387337.

[11] A. LOSEILLE, F. ALAUZET, V. MENIER.*Unique cavity-based operator and hierarchical domain partitioning for fast parallel generation of anisotropic meshes*, in "Computer-Aided Design", October 2016 [*DOI :* 10.1016/J.CAD.2016.09.008], https://hal.inria.fr/hal-01426152.

Invited Conferences

[12] M. A. PARK, J. J. KRAKOS, T. A. MICHAL, A. LOSEILLE, J. J. ALONSO.*Unstructured Grid Adaptation: Status, Potential Impacts, and Recommended Investments Towards CFD 2030*, in "AIAA Fluid Dynamics Conference, AIAA AVIATION Forum", Washington DC, United States, 2016 [*DOI :* 10.2514/6.2016-3323], https://hal.inria.fr/hal-01438667.

International Conferences with Proceedings

[13] J.-R. DE DREUZY, G. PICHOT, P. LAUG, J. ERHEL.*Flow simulation in 3D Discrete Fracture Networks*, in "The XXI International Conference Computational Methods in Water Resources", Toronto, ON, Canada, June 2016, https://hal.inria.fr/hal-01387391.

[14] P. LAUG, H. BOROUCHAKI.*Discrete CAD model for visualization and meshing*, in "International Meshing Roundtable", Washington, DC, United States, September 2016, https://hal.inria.fr/hal-01387348.

[15] P. LAUG, H. BOROUCHAKI.*Geometric Modeling of CAD Surfaces*, in "20th IMACS World Congress", Xiamen, China, December 2016, https://hal.inria.fr/hal-01387398.

Conferences without Proceedings

[16] S. BADREDDINE, L. GIRAUD-MOREAU, A. BOULIFA, A. CHEROUAT.*Etude expérimentale et numérique des efforts lors du formage incrémental de tôles en titane*, in "JET 2016", Hammamet, Tunisia, May 2016, https://hal.archives-ouvertes.fr/hal-01446344.

[17] S. BADREDDINE, L. GIRAUD-MOREAU, A. CHEROUAT.*Optimisation of the single point incremental forming process using surface method and Evolution Strategy*, in "ICOMP'2016", Liège, Belgium, May 2016, https://hal.archives-ouvertes.fr/hal-01446340.

[18] L. GIRAUD-MOREAU, A. CHEROUAT, A. MERAT.*Numerical en Experimental study of the single point incremental forming of titanium sheet*, in "35th International Deep-Drawing Group Conference", Roissy-en-France, France, June 2017, https://hal.archives-ouvertes.fr/hal-01437756.

[19] S. ZHU, A. CHEROUAT, H. BOROUCHAKI.*3D Reconstitution and Simulation of Aluminum Foams*, in "8th IFAC Conference on Manufacturing Modelling, Management and Control", Troyes, France, UTT, June 2016, https://hal.archives-ouvertes.fr/hal-01436463.

Scientific Books (or Scientific Book chapters)

[20] A. LOSEILLE.*Unstructured Mesh Generation and Adaptation*, in "Handbook of Numerical Methods for Hyperbolic Problems — Applied and Modern Issues", R. ABGRALL, C.-W. SHU (editors), Elsevier , 2017, p. 263 - 302 [*DOI :* 10.1016/BS.HNA.2016.10.004], https://hal.inria.fr/hal-01438967.

Research Reports

[21] F. ALAUZET, A. LOSEILLE, G. OLIVIER.*Multi-Scale Anisotropic Mesh Adaptation for Time-Dependent Problems*, Inria Saclay - Ile-de-France, June 2016, n[o] RR-8929, 42, https://hal.inria.fr/hal-01339326.

References in notes

[22] P.-G. CIARLET.*Basic Error Estimates for Elliptic Problems*, Ciarlet, P. G. and Lions, J. L., North Holland, 1991, vol. II.

[23] P.-J. FREY, P.-L. GEORGE.*Maillages. Applications aux éléments finis*, Hermès Science Publications, Paris, 1999.

Project-Team GECO

Geometric Control Design

Table of contents

Project-Team GECO

Creation of the Team: 2011 May 01, updated into Project-Team: 2013 January 01

Keywords:

Computer Science and Digital Science:

      1.5. - Complex systems

      5.3. - Image processing and analysis

      6.1. - Mathematical Modeling

      6.4.1. - Deterministic control

      6.4.3. - Observability and Controlability

      6.4.4. - Stability and Stabilization

      7.13. - Quantum algorithms

Other Research Topics and Application Domains:

      1.3.1. - Understanding and simulation of the brain and the nervous system

      2.6. - Biological and medical imaging

      9.4.2. - Mathematics

      9.4.3. - Physics

1. Members

Research Scientists

    Mario Sigalotti [Team leader, Inria, Researcher, HDR]

    Ugo Boscain [CNRS, Senior Researcher, HDR]

PhD Students

    Nicolas Augier [Ecole Polytechnique, from Sep 2016]

    Mathieu Kohli [Ecole Polytechnique, from Sep 2016]

    Guilherme Mazanti [Ecole Polytechnique, until Aug 2016]

    Jakub Orlowski [Université Paris Sud, from Oct 2016]

    Ludovic Sacchelli [Ecole Polytechnique]

    Leonardo Suriano [Ecole Polytechnique, until Mar 2016]

Post-Doctoral Fellows

    Valentina Franceschi [Inria, from Nov 2016]

    Luca Rizzi [Inria, from May 2016 until Sep 2016]

Administrative Assistants

    Thi Bui [Inria, until Apr 2016]

    Tiffany Caristan [Inria, from Jun 2016]

    Jessica Gameiro [Inria]

2. Overall Objectives

2.1. Overall Objectives

Motion planning is not only a crucial issue in control theory, but also a widespread task of all sort of human activities. The aim of the project-team is to study the various aspects preceding and framing *motion planning*: accessibility analysis (determining which configurations are attainable), criteria to make choice among possible trajectories, trajectory tracking (fixing a possibly unfeasible trajectory and following it as closely as required), performance analysis (determining the cost of a tracking strategy), design of implementable algorithms, robustness of a control strategy with respect to computationally motivated discretizations, etc. The viewpoint that we adopt comes from geometric control: our main interest is in qualitative and intrinsic properties and our focus is on trajectories (either individual ones or families of them).

The main application domain of GECO is *quantum control*. The importance of designing efficient transfers between different atomic or molecular levels in atomic and molecular physics is due to its applications to photochemistry (control by laser pulses of chemical reactions), nuclear magnetic resonance (control by a magnetic field of spin dynamics) and, on a more distant time horizon, the strategic domain of quantum computing.

A second application area concerns the control interpretation of phenomena appearing in *neurophysiology*. It studies the modeling of the mechanisms supervising some biomechanics actions or sensorial reactions such as image reconstruction by the primary visual cortex, eyes movement and body motion. All these problems can be seen as motion planning tasks accomplished by the brain.

As a third applicative domain we propose a system dynamics approach to *switched systems*. Switched systems are characterized by the interaction of continuous dynamics (physical system) and discrete/logical components. They provide a popular modeling framework for heterogeneous aspects issuing from automotive and transportation industry, energy management and factory automation.

3. Research Program

3.1. Geometric control theory

The main research topic of the project-team is **geometric control**, with a special focus on **control design**. The application areas that we target are control of quantum mechanical systems, neurogeometry and switched systems.

Geometric control theory provides a viewpoint and several tools, issued in particular from differential geometry, to tackle typical questions arising in the control framework: controllability, observability, stabilization, optimal control... [22], [56] The geometric control approach is particularly well suited for systems involving nonlinear and nonholonomic phenomena. We recall that nonholonomicity refers to the property of a velocity constraint that is not equivalent to a state constraint.

The expression **control design** refers here to all phases of the construction of a control law, in a mainly open-loop perspective: modeling, controllability analysis, output tracking, motion planning, simultaneous control algorithms, tracking algorithms, performance comparisons for control and tracking algorithms, simulation and implementation.

We recall that

- **controllability** denotes the property of a system for which any two states can be connected by a trajectory corresponding to an admissible control law ;

- **output tracking** refers to a control strategy aiming at keeping the value of some functions of the state arbitrarily close to a prescribed time-dependent profile. A typical example is **configuration tracking** for a mechanical system, in which the controls act as forces and one prescribes the position variables along the trajectory, while the evolution of the momenta is free. One can think for instance at the lateral movement of a car-like vehicle: even if such a movement is unfeasible, it can be tracked with arbitrary precision by applying a suitable control strategy;

- **motion planning** is the expression usually denoting the algorithmic strategy for selecting one control law steering the system from a given initial state to an attainable final one;

- **simultaneous control** concerns algorithms that aim at driving the system from two different initial conditions, with the same control law and over the same time interval, towards two given final states (one can think, for instance, at some control action on a fluid whose goal is to steer simultaneously two floating bodies.) Clearly, the study of which pairs (or n-uples) of states can be simultaneously connected thanks to an admissible control requires an additional controllability analysis with respect to the plain controllability mentioned above.

At the core of control design is then the notion of motion planning. Among the motion planning methods, a preeminent role is played by those based on the Lie algebra associated with the control system ([76], [63], [69]), those exploiting the possible flatness of the system ([50]) and those based on the continuation method ([88]). Optimal control is clearly another method for choosing a control law connecting two states, although it generally introduces new computational and theoretical difficulties.

Control systems with special structure, which are very important for applications are those for which the controls appear linearly. When the controls are not bounded, this means that the admissible velocities form a distribution in the tangent bundle to the state manifold. If the distribution is equipped with a smoothly varying norm (representing a cost of the control), the resulting geometrical structure is called *sub-Riemannian*. Sub-Riemannian geometry thus appears as the underlying geometry of the nonholonomic control systems, playing the same role as Euclidean geometry for linear systems. As such, its study is fundamental for control design. Moreover its importance goes far beyond control theory and is an active field of research both in differential geometry ([75]), geometric measure theory ([51], [26]) and hypoelliptic operator theory ([38]).

Other important classes of control systems are those modeling mechanical systems. The dynamics are naturally defined on the tangent or cotangent bundle of the configuration manifold, they have Lagrangian or Hamiltonian structure, and the controls act as forces. When the controls appear linearly, the resulting model can be seen somehow as a second-order sub-Riemannian structure (see [43]).

The control design topics presented above naturally extend to the case of distributed parameter control systems. The geometric approach to control systems governed by partial differential equations is a novel subject with great potential. It could complement purely analytical and numerical approaches, thanks to its more dynamical, qualitative and intrinsic point of view. An interesting example of this approach is the paper [23] about the controllability of Navier–Stokes equation by low forcing modes.

4. Application Domains

4.1. Quantum control

The issue of designing efficient transfers between different atomic or molecular levels is crucial in atomic and molecular physics, in particular because of its importance in those fields such as photochemistry (control by laser pulses of chemical reactions), nuclear magnetic resonance (NMR, control by a magnetic field of spin dynamics) and, on a more distant time horizon, the strategic domain of quantum computing. This last application explicitly relies on the design of quantum gates, each of them being, in essence, an open loop control law devoted to a prescribed simultaneous control action. NMR is one of the most promising techniques for the implementation of a quantum computer.

Physically, the control action is realized by exciting the quantum system by means of one or several external fields, being them magnetic or electric fields. The resulting control problem has attracted increasing attention, especially among quantum physicists and chemists (see, for instance, [81], [86]). The rapid evolution of the domain is driven by a multitude of experiments getting more and more precise and complex (see the recent review [42]). Control strategies have been proposed and implemented, both on numerical simulations and on physical systems, but there is still a large gap to fill before getting a complete picture of the control properties of quantum systems. Control techniques should necessarily be innovative, in order to take into account the physical peculiarities of the model and the specific experimental constraints.

The area where the picture got clearer is given by finite dimensional linear closed models.

- **Finite dimensional** refers to the dimension of the space of wave functions, and, accordingly, to the finite number of energy levels.
- **Linear** means that the evolution of the system for a fixed (constant in time) value of the control is determined by a linear vector field.
- **Closed** refers to the fact that the systems are assumed to be totally disconnected from the environment, resulting in the conservation of the norm of the wave function.

The resulting model is well suited for describing spin systems and also arises naturally when infinite dimensional quantum systems of the type discussed below are replaced by their finite dimensional Galerkin approximations. Without seeking exhaustiveness, let us mention some of the issues that have been tackled for finite dimensional linear closed quantum systems:

- controllability [24],
- bounds on the controllability time [20],
- STIRAP processes [91],
- simultaneous control [64],
- optimal control ([60], [33], [44]),
- numerical simulations [70].

Several of these results use suitable transformations or approximations (for instance the so-called rotating wave) to reformulate the finite-dimensional Schrödinger equation as a sub-Riemannian system. Open systems have also been the object of an intensive research activity (see, for instance, [25], [61], [82], [39]).

In the case where the state space is infinite dimensional, some optimal control results are known (see, for instance, [29], [40], [57], [30]). The controllability issue is less understood than in the finite dimensional setting, but several advances should be mentioned. First of all, it is known that one cannot expect exact controllability on the whole Hilbert sphere [90]. Moreover, it has been shown that a relevant model, the quantum oscillator, is not even approximately controllable [83], [73]. These negative results have been more recently completed by positive ones. In [31], [32] Beauchard and Coron obtained the first positive controllability result for a quantum particle in a 1D potential well. The result is highly nontrivial and is based on Coron's return method (see [46]). Exact controllability is proven to hold among regular enough wave functions. In particular, exact controllability among eigenfunctions of the uncontrolled Schrödinger operator can be achieved. Other important approximate controllability results have then been proved using Lyapunov methods [72], [77], [58]. While [72] studies a controlled Schrödinger equation in $\mathbb{R}$ for which the uncontrolled Schrödinger operator has mixed spectrum, [77], [58] deal mainly with general discrete-spectrum Schrödinger operators.

In all the positive results recalled in the previous paragraph, the quantum system is steered by a single external field. Different techniques can be applied in the case of two or more external fields, leading to additional controllability results [49], [36].

The picture is even less clear for nonlinear models, such as Gross–Pitaevski and Hartree–Fock equations. The obstructions to exact controllability, similar to the ones mentioned in the linear case, have been discussed in [55]. Optimal control approaches have also been considered [28], [41]. A comprehensive controllability analysis of such models is probably a long way away.

4.2. Neurophysiology

At the interface between neurosciences, mathematics, automatics and humanoid robotics, an entire new approach to neurophysiology is emerging. It arouses a strong interest in the four communities and its development requires a joint effort and the sharing of complementary tools.

A family of extremely interesting problems concerns the understanding of the mechanisms supervising some sensorial reactions or biomechanics actions such as image reconstruction by the primary visual cortex, eyes movement and body motion.

In order to study these phenomena, a promising approach consists in identifying the motion planning problems undertaken by the brain, through the analysis of the strategies that it applies when challenged by external inputs. The role of control is that of a language allowing to read and model neurological phenomena. The control algorithms would shed new light on the brain's geometric perception (the so-called neurogeometry [79]) and on the functional organization of the motor pathways.

- A challenging problem is that of the understanding of the mechanisms which are responsible for the process of image reconstruction in the primary visual cortex V1.

  The visual cortex areas composing V1 are notable for their complex spatial organization and their functional diversity. Understanding and describing their architecture requires sophisticated modeling tools. At the same time, the structure of the natural and artificial images used in visual psychophysics can be fully disclosed only using rather deep geometric concepts. The word "geometry" refers here to the internal geometry of the functional architecture of visual cortex areas (not to the geometry of the Euclidean external space). Differential geometry and analysis both play a fundamental role in the description of the structural characteristics of visual perception.

  A model of human perception based on a simplified description of the visual cortex V1, involving geometric objects typical of control theory and sub-Riemannian geometry, has been first proposed by Petitot ([80]) and then modified by Citti and Sarti ([45]). The model is based on experimental observations, and in particular on the fundamental work by Hubel and Wiesel [54] who received the Nobel prize in 1981.

  In this model, neurons of V1 are grouped into orientation columns, each of them being sensitive to visual stimuli arriving at a given point of the retina and oriented along a given direction. The retina is modeled by the real plane, while the directions at a given point are modeled by the projective line. The fiber bundle having as base the real plane and as fiber the projective line is called the *bundle of directions of the plane*.

  From the neurological point of view, orientation columns are in turn grouped into hypercolumns, each of them sensitive to stimuli arriving at a given point, oriented along any direction. In the same hypercolumn, relative to a point of the plane, we also find neurons that are sensitive to other stimuli properties, such as colors. Therefore, in this model the visual cortex treats an image not as a planar object, but as a set of points in the bundle of directions of the plane. The reconstruction is then realized by minimizing the energy necessary to activate orientation columns among those which are not activated directly by the image. This gives rise to a sub-Riemannian problem on the bundle of directions of the plane.

- Another class of challenging problems concern the functional organization of the motor pathways.

  The interest in establishing a model of the motor pathways, at the same time mathematically rigorous and biologically plausible, comes from the possible spillovers in robotics and neurophysiology. It could help to design better control strategies for robots and artificial limbs, yielding smoother and more progressive movements. Another underlying relevant societal goal (clearly beyond our domain of expertise) is to clarify the mechanisms of certain debilitating troubles such as cerebellar disease, chorea and Parkinson's disease.

  A key issue in order to establish a model of the motor pathways is to determine the criteria underlying the brain's choices. For instance, for the problem of human locomotion (see [27]), identifying

such criteria would be crucial to understand the neural pathways implicated in the generation of locomotion trajectories.

A nowadays widely accepted paradigm is that, among all possible movements, the accomplished ones satisfy suitable optimality criteria (see [89] for a review). One is then led to study an inverse optimal control problem: starting from a database of experimentally recorded movements, identify a cost function such that the corresponding optimal solutions are compatible with the observed behaviors.

Different methods have been taken into account in the literature to tackle this kind of problems, for instance in the linear quadratic case [59] or for Markov processes [78]. However all these methods have been conceived for very specific systems and they are not suitable in the general case. Two approaches are possible to overcome this difficulty. The direct approach consists in choosing a cost function among a class of functions naturally adapted to the dynamics (such as energy functions) and to compare the solutions of the corresponding optimal control problem to the experimental data. In particular one needs to compute, numerically or analytically, the optimal trajectories and to choose suitable criteria (quantitative and qualitative) for the comparison with observed trajectories. The inverse approach consists in deriving the cost function from the qualitative analysis of the data.

4.3. Switched systems

Switched systems form a subclass of hybrid systems, which themselves constitute a key growth area in automation and communication technologies with a broad range of applications. Existing and emerging areas include automotive and transportation industry, energy management and factory automation. The notion of hybrid systems provides a framework adapted to the description of the heterogeneous aspects related to the interaction of continuous dynamics (physical system) and discrete/logical components.

The characterizing feature of switched systems is the collective aspect of the dynamics. A typical question is that of stability, in which one wants to determine whether a dynamical system whose evolution is influenced by a time-dependent signal is uniformly stable with respect to all signals in a fixed class ([66]).

The theory of finite-dimensional hybrid and switched systems has been the subject of intensive research in the last decade and a large number of diverse and challenging problems such as stabilizability, observability, optimal control and synchronization have been investigated (see for instance [87], [67]).

The question of stability, in particular, because of its relevance for applications, has spurred a rich literature. Important contributions concern the notion of common Lyapunov function: when there exists a Lyapunov function that decays along all possible modes of the system (that is, for every possible constant value of the signal), then the system is uniformly asymptotically stable. Conversely, if the system is stable uniformly with respect to all signals switching in an arbitrary way, then a common Lyapunov function exists [68]. In the *linear* finite-dimensional case, the existence of a common Lyapunov function is actually equivalent to the global uniform exponential stability of the system [74] and, provided that the admissible modes are finitely many, the Lyapunov function can be taken polyhedral or polynomial [34], [35], [47]. A special role in the switched control literature has been played by common quadratic Lyapunov functions, since their existence can be tested rather efficiently (see [48] and references therein). Algebraic approaches to prove the stability of switched systems under arbitrary switching, not relying on Lyapunov techniques, have been proposed in [65], [21].

Other interesting issues concerning the stability of switched systems arise when, instead of considering arbitrary switching, one restricts the class of admissible signals, by imposing, for instance, a dwell time constraint [53].

Another rich area of research concerns discrete-time switched systems, where new intriguing phenomena appear, preventing the algebraic characterization of stability even for small dimensions of the state space [62]. It is known that, in this context, stability cannot be tested on periodic signals alone [37].

Finally, let us mention that little is known about infinite-dimensional switched system, with the exception of some results on uniform asymptotic stability ([71], [84], [85]) and some recent papers on optimal control ([52], [92]).

5. Highlights of the Year

5.1. Highlights of the Year

The European Research Council (ERC) has awarded Ugo Boscain with a "Proof of concept grant" for his project *An Artificial Visual Cortex for Image Processing*.

6. New Software and Platforms

6.1. ARTIV1 INPAINTING

ARTIV1 INPAINTING

FUNCTIONAL DESCRIPTION

ARTIV1 INPAINTING is a software for reconstruction of corrupted and damaged images. One of the main features of the algorithm on which the software is based is that it does not require any information about the location and character of the corrupted places. Another important advantage is that this method is massively parallelizable, this allows to work with sufficiently large images. Theoretical background of the presented method is based on the model of geometry of vision due to Petitot, Citti and Sarti. The main step is numerical solution of the equation of 3D hypoelliptic diffusion. A new version of the software has just be submitted for protection at APP (Agence pour la protection des programmes).

- Contact: Ugo Boscain

7. New Results

7.1. New results: geometric control

Let us list some new results in sub-Riemannian geometry and hypoelliptic diffusion obtained by GECO's members.

- In [2] we compare different notions of curvature on contact sub-Riemannian manifolds. In particular we introduce canonical curvatures as the coefficients of the sub-Riemannian Jacobi equation. The main result is that all these coefficients are encoded in the asymptotic expansion of the horizontal derivatives of the sub-Riemannian distance. We explicitly compute their expressions in terms of the standard tensors of contact geometry. As an application of these results, we obtain a sub-Riemannian version of the Bonnet-Myers theorem that applies to any contact manifold.
- In [3] 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 Arnol'd's classification). We show that in the non-generic case, a cornucopia of asymptotics can occur, even for Riemannian surfaces.

- In [5] we study the evolution of the heat and of a free quantum particle (described by the Schrödinger equation) on two-dimensional manifolds endowed with the degenerate Riemannian metric $ds^2 = dx^2 + |x|^{-2\alpha} d\theta^2$, where $x \in \mathbb{R}$, $\theta \in S^1$ 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 Schrödinger 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_N, 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.

- In [6] we study spectral properties of the Laplace–Beltrami operator on two relevant almost-Riemannian manifolds, namely the Grushin structures on the cylinder and on the sphere. As for general almost-Riemannian structures (under certain technical hypothesis), the singular set acts as a barrier for the evolution of the heat and of a quantum particle, although geodesics can cross it. This is a consequence of the self-adjointness of the Laplace–Beltrami operator on each connected component of the manifolds without the singular set. We get explicit descriptions of the spectrum, of the eigenfunctions and their properties. In particular in both cases we get a Weyl law with dominant term $E \log E$. We then study the effect of an Aharonov-Bohm non-apophantic magnetic potential that has a drastic effect on the spectral properties. Other generalized Riemannian structures including conic and anti-conic type manifolds are also studied. In this case, the Aharonov-Bohm magnetic potential may affect the self-adjointness of the Laplace-Beltrami operator.

- Generic singularities of line fields have been studied for lines of principal curvature of embedded surfaces. In [7] we propose an approach to classify generic singularities of general line fields on 2D manifolds. The idea is to identify line fields as bisectors of pairs of vector fields on the manifold, with respect to a given conformal structure. The singularities correspond to the zeros of the vector fields and the genericity is considered with respect to a natural topology in the space of pairs of vector fields. Line fields at generic singularities turn out to be topologically equivalent to the Lemon, Star and Monstar singularities that one finds at umbilical points.

- In [10] we prove that any corank 1 Carnot group of dimension $k + 1$ equipped with a left-invariant measure satisfies the measure contraction property $\mathrm{MCP}(K, N)$ if and only if $K \leq 0$ and $N \geq k + 3$. This generalizes the well known result by Juillet for the Heisenberg group H^{k+1} to a larger class of structures, which admit non-trivial abnormal minimizing curves. The number $k + 3$ coincides with the geodesic dimension of the Carnot group, which we define here for a general metric space. We discuss some of its properties, and its relation with the curvature exponent (the least N such that the $\mathrm{MCP}(0, N)$ is satisfied). We prove that, on a metric measure space, the curvature exponent is always larger than the geodesic dimension which, in turn, is larger than the Hausdorff one. When applied to Carnot groups, our results improve a previous lower bound due to Rifford. As a byproduct, we prove that a Carnot group is ideal if and only if it is fat.

- In [14] we relate some basic constructions of stochastic analysis to differential geometry, via random walk approximations. We consider walks on both Riemannian and sub-Riemannian manifolds in which the steps consist of travel along either geodesics or integral curves associated to orthonormal frames, and we give particular attention to walks where the choice of step is influenced by a volume on the manifold. A primary motivation is to explore how one can pass, in the parabolic scaling limit, from geodesics, orthonormal frames, and/or volumes to diffusions, and hence their infinitesimal generators, on sub-Riemannian manifolds, which is interesting in light of the fact that there is no completely canonical notion of sub-Laplacian on a general sub-Riemannian manifold. However,

even in the Riemannian case, this random walk approach illuminates the geometric significance of Ito and Stratonovich stochastic differential equations as well as the role played by the volume.

- By adapting a technique of Molchanov, we obtain in [15] the heat kernel asymptotics at the sub-Riemannian cut locus, when the cut points are reached by a r-dimensional parametric family of optimal geodesics. We apply these results to the bi-Heisenberg group, that is, a nilpotent left-invariant sub-Riemannian structure on $\mathbb{R}^5$ depending on two real parameters α_1 and α_2. We develop some results about its geodesics and heat kernel associated to its sub-Laplacian and we point out some interesting geometric and analytic features appearing when one compares the isotropic ($\alpha_1 = \alpha_2$) and the non-isotropic cases ($\alpha_1 \neq \alpha_2$). In particular, we give the exact structure of the cut locus, and we get the complete small-time asymptotics for its heat kernel.

- The Whitney extension theorem is a classical result in analysis giving a necessary and sufficient condition for a function defined on a closed set to be extendable to the whole space with a given class of regularity. It has been adapted to several settings, among which the one of Carnot groups. However, the target space has generally been assumed to be equal to $\mathbb{R}^d$ for some $d \geq 1$. We focus in [17] on the extendability problem for general ordered pairs (G_1, G_2) (with G_2 non-Abelian). We analyze in particular the case $G_1 = \mathbb{R}$ and characterize the groups G_2 for which the Whitney extension property holds, in terms of a newly introduced notion that we call pliability. Pliability happens to be related to rigidity as defined by Bryant an Hsu. We exploit this relation in order to provide examples of non-pliable Carnot groups, that is, Carnot groups so that the Whitney extension property does not hold. We use geometric control theory results on the accessibility of control affine systems in order to test the pliability of a Carnot group.

- In [19] we study the cut locus of the free, step two Carnot groups G^k with k generators, equipped with their left-invariant Carnot–Carathéodory metric. In particular, we disprove the conjectures on the shape of the cut loci proposed in the literature, by exhibiting sets of cut points $C \subset G^k$ which, for $k \geq 4$, are strictly larger than conjectured ones. Furthermore, we study the relation of the cut locus with the so-called abnormal set. For each $k \geq 4$, we show that, contrarily to the case $k = 2, 3$, the cut locus always intersects the abnormal set, and there are plenty of abnormal geodesics with finite cut time. Finally, and as a straightforward consequence of our results, we derive an explicit lower bound for the small time heat kernel asymptotics at the points of C. The question whether C coincides with the cut locus for $k \geq 4$ remains open.

We also edited the two volumes [13] and [12], containing some of the lecture notes of the courses given during the IHP triemster on "Geometry, Analysis and Dynamics on sub-Riemannian Manifolds" which we organized in Fall 2014. The second volume also contains a chapter [11] co-authored by members of the team.

7.2. New results: quantum control

- In recent years, several sufficient conditions for the controllability of the Schrödinger equation have been proposed. In [16], we discuss the genericity of these conditions with respect to the variation of the controlled or the uncontrolled potential. In the case where the Schrödinger equation is set on a domain of dimension one, we improve the results in the literature, removing from the previously known genericity results some unnecessary technical assumptions on the regularity of the potentials.

7.3. New results: neurophysiology

In [4] we propose a supervised object recognition method using new global features and inspired by the model of the human primary visual cortex V1 as the semidiscrete roto-translation group $SE(2, N) = \mathbb{Z}_N \rtimes \mathbb{R}^2$. The proposed technique is based on generalized Fourier descriptors on the latter group, which are invariant to natural geometric transformations (rotations, translations). These descriptors are then used to feed an SVM classifier. We have tested our method against the COIL-100 image database and the ORL face database, and compared it with other techniques based on traditional descriptors, global and local. The obtained results have shown that our approach looks extremely efficient and stable to noise, in presence of which it outperforms the other techniques it has been compared with.

7.4. New results: switched systems

- In [8] we address the exponential stability of a system of transport equations with intermittent damping on a network of $N \geq 2$ circles intersecting at a single point O. The N equations are coupled through a linear mixing of their values at O, described by a matrix M. The activity of the intermittent damping is determined by persistently exciting signals, all belonging to a fixed class. The main result is that, under suitable hypotheses on M and on the rationality of the ratios between the lengths of the circles, such a system is exponentially stable, uniformly with respect to the persistently exciting signals. The proof relies on a representation formula for the solutions of this system, which allows one to track down the effects of the intermittent damping. A similar representation formula is used in [18] to study the relative controllability of linear difference equations with multiple delays in the state. Thanks to such formula, we characterize relative controllability in time T in terms of an algebraic property of the matrix-valued coefficients, which reduces to the usual Kalman controllability criterion in the case of a single delay. Relative controllability is studied for solutions in the set of all functions and in the function spaces L^p and C^k. We also compare the relative controllability of the system for different delays in terms of their rational dependence structure, proving that relative controllability for some delays implies relative controllability for all delays that are "less rationally dependent" than the original ones. Finally, we provide an upper bound on the minimal controllability time for a system depending only on its dimension and on its largest delay.

- In [9] we address the stability of transport systems and wave propagation on general networks with time-varying parameters. We do so by reformulating these systems as non-autonomous difference equations and by providing a suitable representation of their solutions in terms of their initial conditions and some time-dependent matrix coefficients. This enables us to characterize the asymptotic behavior of solutions in terms of such coefficients. In the case of difference equations with arbitrary switching, we obtain a delay-independent generalization of the well-known criterion for autonomous systems due to Hale and Silkowski. As a consequence, we show that exponential stability of transport systems and wave propagation on networks is robust with respect to variations of the lengths of the edges of the network preserving their rational dependence structure. This leads to our main result: the wave equation on a network with arbitrarily switching damping at external vertices is exponentially stable if and only if the network is a tree and the damping is bounded away from zero at all external vertices but at most one.

8. Partnerships and Cooperations

8.1. Regional Initiatives

- Project *Stabilité des systèmes à excitation persistante*, Program MathIng, Labex LMH, 2013-2016. This project is about different stability properties for systems whose damping is intermittently activated. The coordinator is Mario Sigalotti. The other members are Yacine Chitour and Guilherme Mazanti.

- iCODE is the Institute for Control and Decision of the Idex Paris Saclay. It was launched in March 2014 for two years until June 2016. We have been involved in three actions funded by iCODE:
  – one action on control of quantum systems, in collaboration with Nicoals Boulant of Neurospin. The action was coordinated by Ugo Boscain;
  – one action on control of wave propagation on networks. The action was coordinated by Mario Sigalotti;
  – one action on switched system. The action was coordinated by Marianne Akian (and handled by MAXPLUS).

  Starting from November 2016, iCODE has been renewed for three years as a IRS (*Institut de Recherche Strategique*) by the Idex Paris Saclay. The funded actions have still not been identified.

- Starting from the end of 2015, we obtained a grant by PGMO (Gaspard Monge Program for Optimisation and operational research) on Geometric Optimal Control. The grant duration is one year, has been renewed in 2016 and is still renewable for a third year. The grant is coordinated by Mario Sigalotti (up to August, it was co-coordinated by Luca Rizzi as well).

8.2. National Initiatives

8.2.1. ANR

The ANR SRGI starts at the end of 2015, for a duration of four years. GECO is one of one of the partners of the ANR. The national coordinator is Emmanuel Trélat (UPMC) and the local one Ugo Boscain.

SRGI deals with sub-Riemannian geometry, hypoelliptic diffusion and geometric control.

8.2.2. Other initiatives

Ugo Boscain and Mario Sigalotti are members of the project DISQUO of the program Inphyniti of the CNRS (duration: one year renewable). Coordinator: Thomas Chambrion (Nancy).

8.3. European Initiatives

8.3.1. FP7 & H2020 Projects

Program: ERC Starting Grant

Project acronym: GeCoMethods

Project title: Geometric Control Methods for the Heat and Schroedinger Equations

Duration: Initially accepted from 1/5/2010 to 1/5/2015, the project has been extended for one additional year, up to 1/5/2016.

Coordinator: Ugo Boscain

Abstract: The aim of this project is to study certain PDEs for which geometric control techniques open new horizons. More precisely we plan to exploit the relation between the sub-Riemannian distance and the properties of the kernel of the corresponding hypoelliptic heat equation and to study controllability properties of the Schroedinger equation.

All subjects studied in this project are applications-driven: the problem of controllability of the Schroedinger equation has direct applications in Laser spectroscopy and in Nuclear Magnetic Resonance; the problem of nonisotropic diffusion has applications in cognitive neuroscience (in particular for models of human vision).

Participants. Main collaborator: Mario Sigalotti. Other members of the team: Andrei Agrachev, Riccardo Adami, Thomas Chambrion, Grégoire Charlot, Yacine Chitour, Jean-Paul Gauthier, Frédéric Jean.

8.4. International Initiatives

8.4.1. Inria International Partners

8.4.1.1. Informal International Partners

SISSA (Scuola Internazionale Superiore di Studi Avanzati), Trieste, Italy.

Sector of Functional Analysis and Applications, Geometric Control group. Coordinator: Andrei A. Agrachev.

We collaborate with the Geometric Control group at SISSA mainly on subjects related with sub-Riemannian geometry. Thanks partly to our collaboration, SISSA has established an official research partnership with École Polytechnique.

8.4.2. Participation in Other International Programs
- Laboratoire Euro Maghrébin de Mathématiques et de leurs Interactions (LEM2I)
  http://lem2i.math.cnrs.fr/
- GDRE Control of Partial Differential Equations (CONEDP)
  http://www.ceremade.dauphine.fr/~glass/GDRE/

8.5. International Research Visitors

8.5.1. Visits of International Scientists
- Andrei Agrachev (SISSA, Italy) is visiting the GECO team for one year, starting in September 2016.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the Organizing Committees
- Mario Sigalotti was member of the organizing committee of the *Workshop on switching dynamics & verification*, IHP, Paris, January 28-29, 2016.
- Ugo Boscain and Mario Sigalotti were member of the organizing committee of the *Workshop on quantum dynamics & control*, IHP, Paris, May 23-24, 2016.

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards
- Ugo Boscain is Associate Editor of SIAM Journal of Control and Optimization
- Ugo Boscain is Managing Editor of Journal of Dynamical and Control Systems
- Mario Sigalotti is Associate Editor of Journal of Dynamical and Control Systems
- Ugo Boscain is Associate Editor of ESAIM Control, Optimisation and Calculus of Variations
- Ugo Boscain is Associate Editor of Mathematical Control and Related Fields
- Ugo Boscain is Associate editor of Analysis and Geometry in Metric Spaces

9.1.3. Invited Talks
- Mario Sigalotti gave an invited talk at the "ExQM Miniworkshop: Mathematics of Quantum Control", Munich, Germany, February 2016.
- Ugo Boscain gave an invited talk at the conference "Geometric Analysis in Control and Vision Theory", Voss, Norway, May 2016.
- Ugo Boscain gave an invited talk at the seminar of the *Departement de Mathématiques d'Orsay*, May 2016.
- Ugo Boscain gave an invited talk at the conference "Recent Trends in Differential equations", Aveiro, Portugal, June 2016.
- Mario Sigalotti gave an invited talk at the *Séminaire de géométrie sous-riemannienne*, IHP, Paris, June 2016.
- Ugo Boscain gave the opening talk at the conference "Geometry, PDE's and Lie Groups in Image Analysis", Eindhoven, The Netherlands, August 2016.
- Mario Sigalotti gave an invited talk at the seminar of the *Dipartimento di Matematica - Università degli Studi di Trento*, Italy, September 2016.

- Ugo Boscain gave an invited talk at the conference "Nouvelles directions en analyse semiclassique", Chalès, France, December 2016.

9.1.4. Research Administration

- Mario Sigalotti is member of the IFAC technical committee "Distributed Parameter Systems".
- Mario Sigalotti is member of the steering committee of the *Institut pour le Contrôle et la Décision* of the Idex Paris-Saclay.

9.2. Teaching - Supervision - Juries

9.2.1. Supervision

- PhD (concluded): Guiherme Mazanti, "Stabilité et taux de convergence pour les systèmes à excitation persistante" [1], supervisors: Yacine Chitour, Mario Sigalotti. Discussed on September 2016.
- PhD in progress: Ludovic Sacchelli, "Sub-Riemannian geometry, hypoelliptic operators, geometry of vision", started in September 2015, supervisors: Ugo Boscain, Mario Sigalotti.
- PhD in progress: Nicolas Augier, "Contrôle adiabatique des systèmes quantiques", started in September 2016, supervisors: Ugo Boscain, Mario Sigalotti.
- PhD in progress: Mathieu Kohli, "Volume and curvature in sub-Riemannian geometry", started in September 2016, supervisors: Davide Barilari, Ugo Boscain.
- PhD in progress: Jakub Orłowski, "Modeling and steering brain oscillations based on in vivo optogenetics data", started in September 2016, supervisors: Antoine Chaillet, Alain Destexhe, and Mario Sigalotti.

9.2.2. Juries

- Ugo Boscain was member of the commission for the PhD defense of Valentina Franceschi, Padue, March 2016.
- Mario Sigalotti was member of the commission for the PhD defense of Francesco Boarotto, SISSA, Trieste, September 2016.
- Ugo Boscain was reviewer and member of the commission for the PhD defense of Jérémy Rouot, Nice, November 2016.

9.3. Popularization

Ugo Boscain gave a concert-seminar at the event "Musique & Mathématiques 2016", Besanon, November 2016.

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] G. MAZANTI.*Stability and stabilization of linear switched systems in finite and infinite dimensions*, Université Paris-Saclay, École Polytechnique, September 2016, https://hal.archives-ouvertes.fr/tel-01427215.

Articles in International Peer-Reviewed Journal

[2] A. AGRACHEV, D. BARILARI, L. RIZZI.*Sub-Riemannian curvature in contact geometry*, in "Journal of Geometric Analysis", 2016 [*DOI : 10.1007/s12220-016-9684-0*], https://hal.archives-ouvertes.fr/hal-01160901.

[3] D. BARILARI, U. BOSCAIN, G. CHARLOT, R. W. NEEL.*On the heat diffusion for generic Riemannian and sub-Riemannian structures*, in "International Mathematics Research Notices", 2016, vol. 2016, p. 1-34, 26 pages, 1 figure, https://hal.archives-ouvertes.fr/hal-00879444.

[4] A. BOHI, D. PRANDI, V. GUIS, F. BOUCHARA, J.-P. GAUTHIER.*Fourier Descriptors Based on the Structure of the Human Primary Visual Cortex with Applications to Object Recognition*, in "Journal of Mathematical Imaging and Vision", July 2016, p. 1-17 [*DOI :* 10.1007/S10851-016-0669-1], https://hal.archives-ouvertes.fr/hal-01383846.

[5] U. BOSCAIN, D. PRANDI.*Self-adjoint extensions and stochastic completeness of the Laplace–Beltrami operator on conic and anticonic surfaces*, in "Journal of Differential Equations", February 2016, vol. 260, n⁰ 4, p. 3234–3269, 28 pages, 2 figures [*DOI :* 10.1016/J.JDE.2015.10.011], https://hal.archives-ouvertes.fr/hal-00848792.

[6] U. BOSCAIN, D. PRANDI, M. SERI.*Spectral analysis and the Aharonov-Bohm effect on certain almost-Riemannian manifolds*, in "Communications in Partial Differential Equations", 2016, vol. 41, n⁰ 1, p. 32–50, 28 pages, 6 figures [*DOI :* 10.1080/03605302.2015.1095766], https://hal.archives-ouvertes.fr/hal-01019955.

[7] U. BOSCAIN, L. SACCHELLI, M. SIGALOTTI.*Generic singularities of line fields on 2D manifolds*, in "Differential Geometry and its Applications", September 2016, vol. Volume 49, n⁰ December 2016, p. 326–350, https://hal.archives-ouvertes.fr/hal-01318515.

[8] Y. CHITOUR, G. MAZANTI, M. SIGALOTTI.*Persistently damped transport on a network of circles*, in "Transactions of the American Mathematical Society", October 2016 [*DOI :* 10.1090/TRAN/6778], https://hal.inria.fr/hal-00999743.

[9] Y. CHITOUR, G. MAZANTI, M. SIGALOTTI.*Stability of non-autonomous difference equations with applications to transport and wave propagation on networks*, in "Networks and Heterogeneous Media", December 2016, vol. 11, p. 563-601 [*DOI :* 10.3934/NHM.2016010], https://hal.archives-ouvertes.fr/hal-01139814.

[10] L. RIZZI.*Measure contraction properties of Carnot groups*, in "Calculus of Variations and Partial Differential Equations", May 2016 [*DOI :* 10.1007/S00526-016-1002-Y], https://hal.archives-ouvertes.fr/hal-01218376.

Scientific Books (or Scientific Book chapters)

[11] A. AGRACHEV, D. BARILARI, U. BOSCAIN.*Introduction to geodesics in sub-Riemannian geometry*, in "Geometry, Analysis and Dynamics on Sub-Riemannian Manifolds - Volume II", EMS Series of Lectures in Mathematics, 2016, https://hal.inria.fr/hal-01392516.

[12] D. BARILARI, U. BOSCAIN, M. SIGALOTTI.*Geometry, Analysis and Dynamics on sub-Riemannian Manifolds - Volume I*, EMS Series of Lectures in Mathematics, European Mathematical Society, 2016 [*DOI :* 10.4171/162], https://hal.archives-ouvertes.fr/hal-01390381.

[13] D. BARILARI, U. BOSCAIN, M. SIGALOTTI.*Geometry, Analysis and Dynamics on sub-Riemannian Manifolds - Volume II*, EMS Series of Lectures in Mathematics, European Mathematical Society, 2016 [*DOI :* 10.4171/163], https://hal.archives-ouvertes.fr/hal-01390382.

Other Publications

[14] A. AGRACHEV, U. BOSCAIN, R. NEEL, L. RIZZI.*Intrinsic random walks in Riemannian and sub-Riemannian geometry via volume sampling*, January 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01259762.

[15] D. BARILARI, U. BOSCAIN, R. W. NEEL.*Heat kernel asymptotics on sub-Riemannian manifolds with symmetries and applications to the bi-Heisenberg group*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01327103.

[16] Y. CHITOUR, M. SIGALOTTI.*Generic controllability of the bilinear Schrödinger equation on 1-D domains: the case of measurable potentials*, 2016, working paper or preprint, https://hal.inria.fr/hal-01292270.

[17] N. JUILLET, M. SIGALOTTI.*Pliability, or the whitney extension theorem for curves in carnot groups*, 2016, working paper or preprint, https://hal.inria.fr/hal-01285215.

[18] G. MAZANTI.*Relative controllability of linear difference equations*, April 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01309166.

[19] L. RIZZI, U. SERRES.*On the cut locus of free, step two Carnot groups*, January 2017, 13 pages. To appear on Proceedings of the AMS, https://hal.archives-ouvertes.fr/hal-01377408.

References in notes

[20] A. A. AGRACHEV, T. CHAMBRION.*An estimation of the controllability time for single-input systems on compact Lie groups*, in "ESAIM Control Optim. Calc. Var.", 2006, vol. 12, n° 3, p. 409–441.

[21] A. A. AGRACHEV, D. LIBERZON.*Lie-algebraic stability criteria for switched systems*, in "SIAM J. Control Optim.", 2001, vol. 40, n° 1, p. 253–269, http://dx.doi.org/10.1137/S0363012999365704.

[22] A. A. AGRACHEV, Y. L. SACHKOV.*Control theory from the geometric viewpoint*, Encyclopaedia of Mathematical Sciences, Springer-Verlag, Berlin, 2004, vol. 87, xiv+412, Control Theory and Optimization, II.

[23] A. A. AGRACHEV, A. V. SARYCHEV.*Navier-Stokes equations: controllability by means of low modes forcing*, in "J. Math. Fluid Mech.", 2005, vol. 7, n° 1, p. 108–152, http://dx.doi.org/10.1007/s00021-004-0110-1.

[24] F. ALBERTINI, D. D'ALESSANDRO.*Notions of controllability for bilinear multilevel quantum systems*, in "IEEE Trans. Automat. Control", 2003, vol. 48, n° 8, p. 1399–1403.

[25] C. ALTAFINI.*Controllability properties for finite dimensional quantum Markovian master equations*, in "J. Math. Phys.", 2003, vol. 44, n° 6, p. 2357–2372.

[26] L. AMBROSIO, P. TILLI.*Topics on analysis in metric spaces*, Oxford Lecture Series in Mathematics and its Applications, Oxford University Press, Oxford, 2004, vol. 25, viii+133.

[27] G. ARECHAVALETA, J.-P. LAUMOND, H. HICHEUR, A. BERTHOZ.*An optimality principle governing human locomotion*, in "IEEE Trans. on Robotics", 2008, vol. 24, n° 1.

[28] L. BAUDOUIN.*A bilinear optimal control problem applied to a time dependent Hartree-Fock equation coupled with classical nuclear dynamics*, in "Port. Math. (N.S.)", 2006, vol. 63, n° 3, p. 293–325.

[29] L. BAUDOUIN, O. KAVIAN, J.-P. PUEL.*Regularity for a Schrödinger equation with singular potentials and application to bilinear optimal control*, in "J. Differential Equations", 2005, vol. 216, n° 1, p. 188–222.

[30] L. BAUDOUIN, J. SALOMON.*Constructive solution of a bilinear optimal control problem for a Schrödinger equation*, in "Systems Control Lett.", 2008, vol. 57, n° 6, p. 453–464, http://dx.doi.org/10.1016/j.sysconle.2007.11.002.

[31] K. BEAUCHARD.*Local controllability of a 1-D Schrödinger equation*, in "J. Math. Pures Appl. (9)", 2005, vol. 84, n° 7, p. 851–956.

[32] K. BEAUCHARD, J.-M. CORON.*Controllability of a quantum particle in a moving potential well*, in "J. Funct. Anal.", 2006, vol. 232, n° 2, p. 328–389.

[33] M. BELHADJ, J. SALOMON, G. TURINICI.*A stable toolkit method in quantum control*, in "J. Phys. A", 2008, vol. 41, n° 36, 362001, 10, http://dx.doi.org/10.1088/1751-8113/41/36/362001.

[34] F. BLANCHINI.*Nonquadratic Lyapunov functions for robust control*, in "Automatica J. IFAC", 1995, vol. 31, n° 3, p. 451–461, http://dx.doi.org/10.1016/0005-1098(94)00133-4.

[35] F. BLANCHINI, S. MIANI.*A new class of universal Lyapunov functions for the control of uncertain linear systems*, in "IEEE Trans. Automat. Control", 1999, vol. 44, n° 3, p. 641–647, http://dx.doi.org/10.1109/9.751368.

[36] A. M. BLOCH, R. W. BROCKETT, C. RANGAN.*Finite Controllability of Infinite-Dimensional Quantum Systems*, in "IEEE Trans. Automat. Control", 2010.

[37] V. D. BLONDEL, J. THEYS, A. A. VLADIMIROV.*An elementary counterexample to the finiteness conjecture*, in "SIAM J. Matrix Anal. Appl.", 2003, vol. 24, n° 4, p. 963–970, http://dx.doi.org/10.1137/S0895479801397846.

[38] A. BONFIGLIOLI, E. LANCONELLI, F. UGUZZONI.*Stratified Lie groups and potential theory for their sub-Laplacians*, Springer Monographs in Mathematics, Springer, Berlin, 2007, xxvi+800.

[39] B. BONNARD, D. SUGNY.*Time-minimal control of dissipative two-level quantum systems: the integrable case*, in "SIAM J. Control Optim.", 2009, vol. 48, n° 3, p. 1289–1308, http://dx.doi.org/10.1137/080717043.

[40] A. BORZÌ, E. DECKER.*Analysis of a leap-frog pseudospectral scheme for the Schrödinger equation*, in "J. Comput. Appl. Math.", 2006, vol. 193, n° 1, p. 65–88.

[41] A. BORZÌ, U. HOHENESTER.*Multigrid optimization schemes for solving Bose-Einstein condensate control problems*, in "SIAM J. Sci. Comput.", 2008, vol. 30, n° 1, p. 441–462, http://dx.doi.org/10.1137/070686135.

[42] C. BRIF, R. CHAKRABARTI, H. RABITZ.*Control of quantum phenomena: Past, present, and future*, Advances in Chemical Physics, S. A. Rice (ed), Wiley, New York, 2010.

[43] F. BULLO, A. D. LEWIS.*Geometric control of mechanical systems*, Texts in Applied Mathematics, Springer-Verlag, New York, 2005, vol. 49, xxiv+726.

[44] R. CABRERA, H. RABITZ.*The landscape of quantum transitions driven by single-qubit unitary transformations with implications for entanglement*, in "J. Phys. A", 2009, vol. 42, n° 27, 275303, 9, http://dx.doi.org/10.1088/1751-8113/42/27/275303.

[45] G. CITTI, A. SARTI.*A cortical based model of perceptual completion in the roto-translation space*, in "J. Math. Imaging Vision", 2006, vol. 24, n° 3, p. 307–326, http://dx.doi.org/10.1007/s10851-005-3630-2.

[46] J.-M. CORON.*Control and nonlinearity*, Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 2007, vol. 136, xiv+426.

[47] W. P. DAYAWANSA, C. F. MARTIN.*A converse Lyapunov theorem for a class of dynamical systems which undergo switching*, in "IEEE Trans. Automat. Control", 1999, vol. 44, n° 4, p. 751–760, http://dx.doi.org/10.1109/9.754812.

[48] L. EL GHAOUI, S.-I. NICULESCU.*Robust decision problems in engineering: a linear matrix inequality approach*, in "Advances in linear matrix inequality methods in control", Philadelphia, PA, Adv. Des. Control, SIAM, 2000, vol. 2, p. 3–37.

[49] S. ERVEDOZA, J.-P. PUEL.*Approximate controllability for a system of Schrödinger equations modeling a single trapped ion*, in "Ann. Inst. H. Poincaré Anal. Non Linéaire", 2009, vol. 26, p. 2111–2136.

[50] M. FLIESS, J. LÉVINE, P. MARTIN, P. ROUCHON.*Flatness and defect of non-linear systems: introductory theory and examples*, in "Internat. J. Control", 1995, vol. 61, n° 6, p. 1327–1361, http://dx.doi.org/10.1080/00207179508921959.

[51] B. FRANCHI, R. SERAPIONI, F. SERRA CASSANO.*Regular hypersurfaces, intrinsic perimeter and implicit function theorem in Carnot groups*, in "Comm. Anal. Geom.", 2003, vol. 11, n° 5, p. 909–944.

[52] M. GUGAT.*Optimal switching boundary control of a string to rest in finite time*, in "ZAMM Z. Angew. Math. Mech.", 2008, vol. 88, n° 4, p. 283–305.

[53] J. HESPANHA, S. MORSE.*Stability of switched systems with average dwell-time*, in "Proceedings of the 38th IEEE Conference on Decision and Control, CDC 1999, Phoenix, AZ, USA", 1999, p. 2655–2660.

[54] D. HUBEL, T. WIESEL.*Brain and Visual Perception: The Story of a 25-Year Collaboration*, Oxford University Press, Oxford, 2004.

[55] R. ILLNER, H. LANGE, H. TEISMANN.*Limitations on the control of Schrödinger equations*, in "ESAIM Control Optim. Calc. Var.", 2006, vol. 12, n° 4, p. 615–635, http://dx.doi.org/10.1051/cocv:2006014.

[56] A. ISIDORI.*Nonlinear control systems*, Communications and Control Engineering Series, Second, Springer-Verlag, Berlin, 1989, xii+479, An introduction.

[57] K. ITO, K. KUNISCH.*Optimal bilinear control of an abstract Schrödinger equation*, in "SIAM J. Control Optim.", 2007, vol. 46, n° 1, p. 274–287.

[58] K. ITO, K. KUNISCH.*Asymptotic properties of feedback solutions for a class of quantum control problems*, in "SIAM J. Control Optim.", 2009, vol. 48, n° 4, p. 2323–2343, http://dx.doi.org/10.1137/080720784.

[59] R. KALMAN.*When is a linear control system optimal?*, in "ASME Transactions, Journal of Basic Engineering", 1964, vol. 86, p. 51–60.

[60] N. KHANEJA, S. J. GLASER, R. W. BROCKETT.*Sub-Riemannian geometry and time optimal control of three spin systems: quantum gates and coherence transfer*, in "Phys. Rev. A (3)", 2002, vol. 65, n° 3, part A, 032301, 11.

[61] N. KHANEJA, B. LUY, S. J. GLASER.*Boundary of quantum evolution under decoherence*, in "Proc. Natl. Acad. Sci. USA", 2003, vol. 100, n° 23, p. 13162–13166, http://dx.doi.org/10.1073/pnas.2134111100.

[62] V. S. KOZYAKIN.*Algebraic unsolvability of a problem on the absolute stability of desynchronized systems*, in "Avtomat. i Telemekh.", 1990, p. 41–47.

[63] G. LAFFERRIERE, H. J. SUSSMANN.*A differential geometry approach to motion planning*, in "Nonholonomic Motion Planning (Z. Li and J. F. Canny, editors)", Kluwer Academic Publishers, 1993, p. 235-270.

[64] J.-S. LI, N. KHANEJA.*Ensemble control of Bloch equations*, in "IEEE Trans. Automat. Control", 2009, vol. 54, n° 3, p. 528–536, http://dx.doi.org/10.1109/TAC.2009.2012983.

[65] D. LIBERZON, J. P. HESPANHA, A. S. MORSE.*Stability of switched systems: a Lie-algebraic condition*, in "Systems Control Lett.", 1999, vol. 37, n° 3, p. 117–122, http://dx.doi.org/10.1016/S0167-6911(99)00012-2.

[66] D. LIBERZON.*Switching in systems and control*, Systems & Control: Foundations & Applications, Birkhäuser Boston Inc., Boston, MA, 2003, xiv+233.

[67] H. LIN, P. J. ANTSAKLIS.*Stability and stabilizability of switched linear systems: a survey of recent results*, in "IEEE Trans. Automat. Control", 2009, vol. 54, n° 2, p. 308–322, http://dx.doi.org/10.1109/TAC.2008.2012009.

[68] Y. LIN, E. D. SONTAG, Y. WANG.*A smooth converse Lyapunov theorem for robust stability*, in "SIAM J. Control Optim.", 1996, vol. 34, n° 1, p. 124–160, http://dx.doi.org/10.1137/S0363012993259981.

[69] W. LIU.*Averaging theorems for highly oscillatory differential equations and iterated Lie brackets*, in "SIAM J. Control Optim.", 1997, vol. 35, n° 6, p. 1989–2020, http://dx.doi.org/10.1137/S0363012994268667.

[70] Y. MADAY, J. SALOMON, G. TURINICI.*Monotonic parareal control for quantum systems*, in "SIAM J. Numer. Anal.", 2007, vol. 45, n° 6, p. 2468–2482, http://dx.doi.org/10.1137/050647086.

[71] A. N. MICHEL, Y. SUN, A. P. MOLCHANOV.*Stability analysis of discountinuous dynamical systems determined by semigroups*, in "IEEE Trans. Automat. Control", 2005, vol. 50, n° 9, p. 1277–1290, http://dx.doi.org/10.1109/TAC.2005.854582.

[72] M. MIRRAHIMI.*Lyapunov control of a particle in a finite quantum potential well*, in "Proceedings of the 45th IEEE Conference on Decision and Control", 2006.

[73] M. MIRRAHIMI, P. ROUCHON.*Controllability of quantum harmonic oscillators*, in "IEEE Trans. Automat. Control", 2004, vol. 49, n° 5, p. 745–747.

[74] A. P. MOLCHANOV, Y. S. PYATNITSKIY.*Criteria of asymptotic stability of differential and difference inclusions encountered in control theory*, in "Systems Control Lett.", 1989, vol. 13, n° 1, p. 59–64, http:// dx.doi.org/10.1016/0167-6911(89)90021-2.

[75] R. MONTGOMERY.*A tour of subriemannian geometries, their geodesics and applications*, Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 2002, vol. 91, xx+259.

[76] R. M. MURRAY, S. S. SASTRY.*Nonholonomic motion planning: steering using sinusoids*, in "IEEE Trans. Automat. Control", 1993, vol. 38, n° 5, p. 700–716, http://dx.doi.org/10.1109/9.277235.

[77] V. NERSESYAN.*Growth of Sobolev norms and controllability of the Schrödinger equation*, in "Comm. Math. Phys.", 2009, vol. 290, n° 1, p. 371–387.

[78] A. Y. NG, S. RUSSELL.*Algorithms for Inverse Reinforcement Learning*, in "Proc. 17th International Conf. on Machine Learning", 2000, p. 663–670.

[79] J. PETITOT.*Neurogéomètrie de la vision. Modèles mathématiques et physiques des architectures fonctionnelles*, Les Éditions de l'École Polythechnique, 2008.

[80] J. PETITOT, Y. TONDUT.*Vers une neurogéométrie. Fibrations corticales, structures de contact et contours subjectifs modaux*, in "Math. Inform. Sci. Humaines", 1999, n° 145, p. 5–101.

[81] H. RABITZ, H. DE VIVIE-RIEDLE, R. MOTZKUS, K. KOMPA.*Wither the future of controlling quantum phenomena?*, in "SCIENCE", 2000, vol. 288, p. 824–828.

[82] D. ROSSINI, T. CALARCO, V. GIOVANNETTI, S. MONTANGERO, R. FAZIO.*Decoherence by engineered quantum baths*, in "J. Phys. A", 2007, vol. 40, n° 28, p. 8033–8040, http://dx.doi.org/10.1088/1751-8113/40/ 28/S12.

[83] P. ROUCHON.*Control of a quantum particle in a moving potential well*, in "Lagrangian and Hamiltonian methods for nonlinear control 2003", Laxenburg, IFAC, 2003, p. 287–290.

[84] A. SASANE.*Stability of switching infinite-dimensional systems*, in "Automatica J. IFAC", 2005, vol. 41, n° 1, p. 75–78, http://dx.doi.org/10.1016/j.automatica.2004.07.013.

[85] A. SAURABH, M. H. FALK, M. B. ALEXANDRE.*Stability analysis of linear hyperbolic systems with switching parameters and boundary conditions*, in "Proceedings of the 47th IEEE Conference on Decision and Control, CDC 2008, December 9-11, 2008, Cancún, Mexico", 2008, p. 2081–2086.

[86] M. SHAPIRO, P. BRUMER.*Principles of the Quantum Control of Molecular Processes*, Principles of the Quantum Control of Molecular Processes, pp. 250. Wiley-VCH, February 2003.

[87] R. SHORTEN, F. WIRTH, O. MASON, K. WULFF, C. KING.*Stability criteria for switched and hybrid systems*, in "SIAM Rev.", 2007, vol. 49, n° 4, p. 545–592, http://dx.doi.org/10.1137/05063516X.

[88] H. J. SUSSMANN.*A continuation method for nonholonomic path finding*, in "Proceedings of the 32th IEEE Conference on Decision and Control, CDC 1993, Piscataway, NJ, USA", 1993, p. 2718–2723.

[89] E. TODOROV.*12*, in "Optimal control theory", Bayesian Brain: Probabilistic Approaches to Neural Coding, Doya K (ed), 2006, p. 269–298.

[90] G. TURINICI.*On the controllability of bilinear quantum systems*, in "Mathematical models and methods for ab initio Quantum Chemistry", M. DEFRANCESCHI, C. LE BRIS (editors), Lecture Notes in Chemistry, Springer, 2000, vol. 74.

[91] L. YATSENKO, S. GUÉRIN, H. JAUSLIN.*Topology of adiabatic passage*, in "Phys. Rev. A", 2002, vol. 65, 043407, 7.

[92] E. ZUAZUA.*Switching controls*, in "Journal of the European Mathematical Society", 2011, vol. 13, n° 1, p. 85–117.

Project-Team GRACE

Geometry, arithmetic, algorithms, codes and encryption

Table of contents

Project-Team GRACE

Creation of the Team: 2012 January 01, updated into Project-Team: 2013 July 01

Keywords:

Computer Science and Digital Science:

      4.2. - Correcting codes

      4.3.1. - Public key cryptography

      4.3.3. - Cryptographic protocols

      4.8. - Privacy-enhancing technologies

      7.6. - Computer Algebra

      7.7. - Number theory

Other Research Topics and Application Domains:

      9.4.2. - Mathematics

      9.8. - Privacy

1. Members

Research Scientists

    Daniel Augot [Team leader, Inria, Senior Researcher, HDR]

    Alain Couvreur [Inria, Researcher]

    Benjamin Smith [Inria, Researcher]

Faculty Members

    Luca de Feo [Univ. Versailles, Associate Professor]

    Françoise Levy-Dit-Vehel [ENSTA, Associate Professor, HDR]

    François Morain [Ecole Polytechnique, Professor, HDR]

Technical Staff

    Nicholas Coxon [Inria]

    David Lucas [Inria, until Sep 2016]

PhD Students

    Elise Barelli [Inria]

    Hanna-Mae Bisserier [Inst. de Recherche Technologique SystemX, from Nov 2016]

    Nicolas Duhamel [ENS Cachan]

    Pierre Karpman [Inria, until October 2016]

    Julien Lavauzelle [Ecole Polytechnique]

Post-Doctoral Fellow

    Virgile Ducet [Inria, until Oct 2016]

Visiting Scientists

    Philippe Lebacque [Univ. Franche-Comté, Faculty Member]

    François-Régis Blache [Université des Antilles, Faculty Member, Nov 2016]

Administrative Assistants

    Helene Bessin Rousseau [Inria]

    Jessica Gameiro [Inria]

    Emmanuelle Perrot [Inria]

Others

    Christopher Barbin [Inria, Master intern, from Apr 2016 until Jul 2016]

Nagarjun Chinthamani Dwarakanath [Inria, Master intern, from Apr 2016 until Jul 2016]
Eleonora Palazzolo [Inria, Master intern, until Apr 2016]
Evrim Petek [Inria, Master intern, from May 2016 until Sep 2016]

2. Overall Objectives

2.1. Scientific foundations

GRACE has two broad application domains—cryptography and coding theory—linked by a common foundation in algorithmic number theory and the geometry of algebraic curves. In our research, which combines theoretical work with practical software development, we use algebraic curves to *create better cryptosystems*, to *provide better security assessments* for cryptographic key sizes, and to *build the best error-correcting codes*.

Coding and cryptography deal (in different ways) with securing communication systems for high-level applications. In our research, the two domains are linked by the computational issues related to algebraic curves (over various fields) and arithmetic rings. These fundamental number-theoretic algorithms, at the crossroads of a rich area of mathematics and computer science, have already proven their relevance in public key cryptography, with industrial successes including the RSA cryptosystem and elliptic curve cryptography. It is less well-known that the same branches of mathematics can be used to build very good codes for error correction. While coding theory has traditionally had an electrical engineering flavour, recent developments in computer science have shed new light on coding theory, leading to new applications more central to computer science.

3. Research Program

3.1. Algorithmic Number Theory

Algorithmic Number Theory is concerned with replacing special cases with general algorithms to solve problems in number theory. In the Grace project, it appears in three main threads:

- fundamental algorithms for integers and polynomials (including primality and factorization);
- algorithms for finite fields (including discrete logarithms); and
- algorithms for algebraic curves.

Clearly, we use computer algebra in many ways. Research in cryptology has motivated a renewed interest in Algorithmic Number Theory in recent decades—but the fundamental problems still exist *per se*. Indeed, while algorithmic number theory application in cryptanalysis is epitomized by applying factorization to breaking RSA public key, many other problems, are relevant to various area of computer science. Roughly speaking, the problems of the cryptological world are of bounded size, whereas Algorithmic Number Theory is also concerned with asymptotic results.

3.2. Arithmetic Geometry: Curves and their Jacobians

Theme: Arithmetic Geometry: Curves and their Jacobians *Arithmetic Geometry* is the meeting point of algebraic geometry and number theory: that is, the study of geometric objects defined over arithmetic number systems (such as the integers and finite fields). The fundamental objects for our applications in both coding theory and cryptology are curves and their Jacobians over finite fields.

An algebraic *plane curve* $\mathcal{X}$ over a field $\mathbf{K}$ is defined by an equation

$$\mathcal{X} : F_{\mathcal{X}}(x, y) = 0 \quad \text{where } F_{\mathcal{X}} \in \mathbf{K}[x, y].$$

(Not every curve is planar—we may have more variables, and more defining equations—but from an algorithmic point of view, we can always reduce to the plane setting.) The *genus* $g_{\mathcal{X}}$ of $\mathcal{X}$ is a non-negative integer classifying the essential geometric complexity of $\mathcal{X}$; it depends on the degree of $F_{\mathcal{X}}$ and on the number of singularities of $\mathcal{X}$. The curve $\mathcal{X}$ is associated in a functorial way with an algebraic group $J_{\mathcal{X}}$, called the *Jacobian* of $\mathcal{X}$. The group $J_{\mathcal{X}}$ has a geometric structure: its elements correspond to points on a $g_{\mathcal{X}}$-dimensional projective algebraic group variety. Typically, we do not compute with the equations defining this projective variety: there are too many of them, in too many variables, for this to be convenient. Instead, we use fast algorithms based on the representation in terms of classes of formal sums of points on $\mathcal{X}$.

The simplest curves with nontrivial Jacobians are curves of genus 1, known as *elliptic curves*; they are typically defined by equations of the form $y^2 = x^3 + Ax + B$. Elliptic curves are particularly important given their central role in public-key cryptography over the past two decades. Curves of higher genus are important in both cryptography and coding theory.

3.3. Curve-Based cryptology

Theme: Curve-Based Cryptology

Jacobians of curves are excellent candidates for cryptographic groups when constructing efficient instances of public-key cryptosystems. Diffie–Hellman key exchange is an instructive example.

Suppose Alice and Bob want to establish a secure communication channel. Essentially, this means establishing a common secret *key*, which they will then use for encryption and decryption. Some decades ago, they would have exchanged this key in person, or through some trusted intermediary; in the modern, networked world, this is typically impossible, and in any case completely unscalable. Alice and Bob may be anonymous parties who want to do e-business, for example, in which case they cannot securely meet, and they have no way to be sure of each other's identities. Diffie–Hellman key exchange solves this problem. First, Alice and Bob publicly agree on a cryptographic group G with a generator P (of order N); then Alice secretly chooses an integer a from $[1..N]$, and sends aP to Bob. In the meantime, Bob secretly chooses an integer b from $[1..N]$, and sends bP to Alice. Alice then computes $a(bP)$, while Bob computes $b(aP)$; both have now computed abP, which becomes their shared secret key. The security of this key depends on the difficulty of computing abP given P, aP, and bP; this is the Computational Diffie–Hellman Problem (CDHP). In practice, the CDHP corresponds to the Discrete Logarithm Problem (DLP), which is to determine a given P and aP.

This simple protocol has been in use, with only minor modifications, since the 1970s. The challenge is to create examples of groups G with a relatively compact representation and an efficiently computable group law, and such that the DLP in G is hard (ideally approaching the exponential difficulty of the DLP in an abstract group). The Pohlig–Hellman reduction shows that the DLP in G is essentially only as hard as the DLP in its largest prime-order subgroup. We therefore look for compact and efficient groups of prime order.

The classic example of a group suitable for the Diffie–Hellman protocol is the multiplicative group of a finite field $\mathbf{F}_q$. There are two problems that render its usage somewhat less than ideal. First, it has too much structure: we have a subexponential Index Calculus attack on the DLP in this group, so while it is very hard, the DLP falls a long way short of the exponential difficulty of the DLP in an abstract group. Second, there is only one such group for each q: its subgroup treillis depends only on the factorization of $q - 1$, and requiring $q - 1$ to have a large prime factor eliminates many convenient choices of q.

This is where Jacobians of algebraic curves come into their own. First, elliptic curves and Jacobians of genus 2 curves do not have a subexponential index calculus algorithm: in particular, from the point of view of the DLP, a generic elliptic curve is currently *as strong as* a generic group of the same size. Second, they provide some diversity: we have many degrees of freedom in choosing curves over a fixed $\mathbf{F}_q$, with a consequent diversity of possible cryptographic group orders. Furthermore, an attack which leaves one curve vulnerable may not necessarily apply to other curves. Third, viewing a Jacobian as a geometric object rather than a pure group allows us to take advantage of a number of special features of Jacobians. These features include efficiently computable pairings, geometric transformations for optimised group laws, and the availability of efficiently computable non-integer endomorphisms for accelerated encryption and decryption.

3.4. Algebraic Coding Theory

Theme: Coding theory

Coding Theory studies originated with the idea of using redundancy in messages to protect against noise and errors. The last decade of the 20th century has seen the success of so-called iterative decoding methods, which enable us to get very close to the Shannon capacity. The capacity of a given channel is the best achievable transmission rate for reliable transmission. The consensus in the community is that this capacity is more easily reached with these iterative and probabilistic methods than with algebraic codes (such as Reed–Solomon codes).

However, algebraic coding is useful in settings other than the Shannon context. Indeed, the Shannon setting is a random case setting, and promises only a vanishing error probability. In contrast, the algebraic Hamming approach is a worst case approach: under combinatorial restrictions on the noise, the noise can be adversarial, with strictly zero errors.

These considerations are renewed by the topic of list decoding after the breakthrough of Guruswami and Sudan at the end of the nineties. List decoding relaxes the uniqueness requirement of decoding, allowing a small list of candidates to be returned instead of a single codeword. List decoding can reach a capacity close to the Shannon capacity, with zero failure, with small lists, in the adversarial case. The method of Guruswami and Sudan enabled list decoding of most of the main algebraic codes: Reed–Solomon codes and Algebraic–Geometry (AG) codes and new related constructions "capacity-achieving list decodable codes". These results open the way to applications again adversarial channels, which correspond to worst case settings in the classical computer science language.

Another avenue of our studies is AG codes over various geometric objects. Although Reed–Solomon codes are the best possible codes for a given alphabet, they are very limited in their length, which cannot exceed the size of the alphabet. AG codes circumvent this limitation, using the theory of algebraic curves over finite fields to construct long codes over a fixed alphabet. The striking result of Tsfasman–Vladut–Zink showed that codes better than random codes can be built this way, for medium to large alphabets. Disregarding the asymptotic aspects and considering only finite length, AG codes can be used either for longer codes with the same alphabet, or for codes with the same length with a smaller alphabet (and thus faster underlying arithmetic).

From a broader point of view, wherever Reed–Solomon codes are used, we can substitute AG codes with some benefits: either beating random constructions, or beating Reed–Solomon codes which are of bounded length for a given alphabet.

Another area of Algebraic Coding Theory with which we are more recently concerned is the one of Locally Decodable Codes. After having been first theoretically introduced, those codes now begin to find practical applications, most notably in cloud-based remote storage systems.

4. Application Domains

4.1. Cryptography and Cryptanalysis

In the twenty-first century, cryptography plays two essential roles: it is used to ensure *security* and *integrity* of communications and communicating entities. Contemporary cryptographic techniques can be used to hide private data, and to prove that public data has not been modified; to provide anonymity, and to assert and prove public identities. The creation and testing of practical cryptosystems involves

1. The design of provably secure protocols;
2. The design and analysis of compact and efficient algorithms to implement those protocols, and to attack their underlying mathematical and computational problems;
3. The robust implementation of those algorithms in low-level software and hardware, and their deployment in the wild.

While these layers are interdependent, GRACE's cryptographic research is focused heavily on the middle layer: we design, implement, and analyze the most efficient algorithms for fundamental tasks in contemporary cryptography. Our "clients", in a sense, are protocol designers on the one hand, and software and hardware engineers on the other.

F. Morain and B. Smith work primarily on the number-theoretic algorithms that underpin the current state-of-the-art in public-key cryptography (which is used to establish secure connections, and create and verify digital signatures, among other applications). For example, their participation in the ANR CATREL project aims to give a realistic assessment of the security of systems based on the Discrete Logarithm Problem, by creating a free, open, algorithmic package implementing the fastest known algorithms for attacking DLP instances. This will have an extremely important impact on contemporary pairing-based cryptosystems, as well as legacy finite field-based cryptosystems. On a more constructive note, F. Morain' elliptic curve point counting and primality proving algorithms are essential tools in the everyday construction of strong public-key cryptosystems, while B. Smith's recent work on elliptic and genus 2 curves aims to improve the speed of curve-based cryptosystems (such as Elliptic Curve Diffie–Hellman key exchange, a crucial step in establishing secure internet connections) without compromising their security.

D. Augot, F. Levy-dit-Vehel, and A. Couvreur's research on codes has far-reaching applications in *code-based cryptography*. This is a field which is growing rapidly in importance—partly due to the supposed resistance of code-based cryptosystems to attacks from quantum computing, partly due to the range of new techniques on offer, and partly because the fundamental problem of parameter selection is relatively poorly understood. For example, A. Couvreur's work on filtration attacks on codes has an important impact on the design of code-based systems using wild Goppa codes or algebraic geometry codes, and on the choice of parameter sizes for secure implementations.

Coding theory also has important practical applications in the improvement of conventional symmetric cryptosystems. For example, D. Augot's recent work on MDS matrices via BCH codes gives a more efficient construction of optimal diffusion layers in block ciphers. Here we use combinatorial, non-algorithmic properties of codes, in the internals of designs of block ciphers.

While coding theory brings tools as above for the classical problems of encryption, authentication, and so on, it can also provide solutions to new cryptographic problems. This is classically illustrated by the use of Reed-Solomon codes in secret sharing schemes. Grace is involved in the study, construction and implementation of locally decodable codes, which have applications in quite a few cryptographic protocols : *Private Information Retrieval*, *Proofs of Retrievability*, *Proofs of Ownership*, etc.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Events organization

- A. Couvreur, D. Augot and D. Lucas organized with L. De Feo and Hugues Randriambololona (ENST ParisTech) a spring school on coding and cryptology in la Chapelle Gauthier (Seine et Marne).

- A. Couvreur and D. Augot organized 4 days workshop in november 2016 for the ANR MANTA. The topics were: "Decoding" and "Codes from surfaces".

- SageDays75. To conclude the ACTIS projet, we organized a one-week SageDays in August 2016. The day was spent at Inria Saclay, and people were staying at night in a cottage in Vallée de Chevreuse.

  The overall theme of this Sage Days was coding theory and exact linear algebra related to it, but there was be lots of general hacking. The aim of this Sage Days was to Introduce Sage to coding theorists;

have presentations about the enhancements we made to Sage's coding theory library during Inria's ACTIS project; Help people to work on their own projects.

We had a few talks on the mornings, and coding sprints on the afternoons. The first days' talks were focused on basic functionalities of our library, the last 2 days on advanced functionalities, with an emphasis on Sage development.

We were glad to attract several core sage developpers, who recognized the quality of the work done by D. Lucas.

6. New Software and Platforms

6.1. ACTIS

FUNCTIONAL DESCRIPTION

The aim of this project is to vastly improve the state of the error correcting library in Sage. The existing library does not present a good and usable API, and the provided algorithms are very basic, irrelevant, and outdated. We thus have two directions for improvement: renewing the APIs to make them actually usable by researchers, and incorporating efficient programs for decoding, like J. Nielsen's CodingLib, which contains many new algorithms.

- Contact: David Lucas
- https://bil.inria.fr/fr/software/view/2114/tab

During the project, D. Lucas and J. Nielsen proposed a google summer of code project on rank-metric codes under our ACTIS framework. The intern was Arpit Merchant, who visited us for SageDays75.

6.2. muKummer

KEYWORD: Cryptography
FUNCTIONAL DESCRIPTION

A competitive, high-speed, open implementation of the Diffie–Hellman key exchange protocol and a Schnorr-type digital signature scheme, targeting the 128-bit security level on two microcontroller platforms: the classic AVR ATMega 8-bit platform and the more modern ARM Cortex M0 32-bit platform. These downloads contain mixed C and assembly sources for the implementations described in [16].

- Participant: Benjamin Smith
- Contact: Benjamin Smith
- ATMega implementation URL: http://www.cs.ru.nl/~jrenes/software/mukummer-avr.tar.gz
- Cortex M0 implementation URL: http://www.cs.ru.nl/~jrenes/software/mukummer-arm.tar.gz

7. New Results

7.1. Faster elliptic and hyperelliptic curve cryptography

B. Smith made several contributions to the development of faster arithmetic on elliptic curves and genus 2 Jacobians in 2016. In joint work with C. Costello and P.-N. Chung, he gave a new, efficient, uniform, and constant-time scalar multiplication algorithm for genus 2 Jacobians exploiting fast Kummer surface arithmetic and features of differential addition chains; this was presented at SAC 2016. The theory in this article was the basis of a highly competitive implementation of key exchange and signatures for microcontroller platforms, in joint work with J. Renes, P. Schwabe, and L. Batina, presented at CHES 2016.

7.2. Quantum factoring

Integer factorization via Shor's algorithm is a benchmark problem for general quantum computers, but surprisingly little work has been done on optimizing the algorithm for use as a serious factoring tool once large quantum computers are built (rather than as a proof of concept). In the meantime, given the limited size of contemporary quantum computers and the practical difficulties involved in building them, any optimizations to quantum factoring algorithms can lead to significant practical improvements. In a new interdisciplinary project with physicists F. Grosshans and T. Lawson, F. Morain and B. Smith have derived a simple new quantum factoring algorithm for cryptographic integers; its expected runtime is lower than Shor's factoring algorithm, and it should also be easier to implement in practice [22].

7.3. Advances in point counting

Determining the number of points on an elliptic curve, or more generally on the Jacobian of an algebraic curve, is a classic problem in algorithmic number theory that is now crucial for efficiently generating secure cryptographic parameters. Together with C. Scribot, F. Morain and B. Smith developed an improved version of the state-of-the-art SEA algorithm for certain families of elliptic curves with special endomorphisms; this was presented at ANTS-XII [10]. B. Smith also led a project group on special genus-2 point counting algorithms at the "Algebraic Geometry for Coding Theory and Cryptography" workshop at IPAM, UCLA, in 2016.

7.4. Cryptanalysis of code based cryptosystems by filtration attacks

The McEliece encryption scheme based on binary Goppa codes was one of the first public-key encryption schemes [31]. Its security rests on the difficulty of decoding an arbitrary code. The original proposal uses classical Goppa codes, and while it still remains unbroken, it requires a huge size of key. On the other hand, many derivative systems based on other families of algebraic codes have been subject to key recovery attacks. Up to now, key recovery attacks were based either on a variant of Sidelnikov and Shestakov's attack [32], where the first step involves the computation of minimum-weight codewords, or on the resolution of a system of polynomial equations using Gröbner bases.

In [26], A. Couvreur, P. Gaborit, V. Gauthier, A. Otmani and J.-P. Tillich introduced a new paradigm of attack called *filtration attacks*. The general principle decomposes in two steps:

1. **Distinguishing** the public code from a random one using the square code operation.
2. **Computing a filtration** of the public code using the distinguisher, and deriving from this filtration an efficient decoding algorithm for the public code.

This new style of attack allowed A. Couvreur, A. Otmani and J.-P. Tillich to break (in polynomial time) McEliece based on wild Goppa codes over quadratic extensions [3]. A detailed long version has been written and recently published [9]. A. Couvreur, Irene Márquez–Corbella, and R. Pellikaan broke McEliece based on algebraic geometry codes from curves of arbitrary genus [2], [27] by reconstructing optimal polynomial time decoding algorithms decoding up to the half minimum distance minus half the genus. This can be computed from the raw data of a generator matrix. In a recently submitted long version [21] the algorithm has been improved and permits to reconstruct a decoding algorithm up to the half minimum distance.

7.5. Quantum LDPC codes

Quantum codes are the analogous of error correcting codes for a quantum computer. A well known family of quantum codes are the CSS codes due to Calderbank, Shor and Steane can be represented by a pair of matrices (H_X, H_Z) such that $H_X H_Z^T = 0$. As in classical coding theory, if these matrices are sparse, then the code is said to be LDPC. An open problem in quantum coding theory is to get a family of quantum LDPC codes whose asymptotic minimum distance is in $\Omega(n^\alpha)$ for some $\alpha > 1/2$. No such family is known and actually, only few known families of quantum LDPC codes have a minimum distance tending to infinity.

In [24], Benjamin Audoux (I2M, Marseille) and A. Couvreur investigate a problem suggested by Bravyi and Hastings. They studied the behaviour of iterated tensor powers of CSS codes and prove in particular that such families always have a minimum distance tending to infinity. They propose also 3 families of LDPC codes whose minimum distance is in $\Omega(n^\beta)$ for all $\beta < 1/2$.

7.6. Discrete Logarithm computations in finite fields with the NFS algorithm

The best discrete logarithm record computations in prime fields and large characteristic finite fields are obtained with Number Field Sieve algorithm (NFS) at the moment. This algorithm is made of four steps:

1. polynomial selection;
2. relation collection (with a sieving technique);
3. linear algebra (computing the kernel of a huge matrix, of millions of rows and columns);
4. individual discrete logarithm computation.

The two more time consuming steps are the relation collection step and the linear algebra step. The polynomial selection is quite fast but is very important since it determines the complexity of the algorithm. Selecting better polynomials is a key to improve the overall running-time of the NFS algorithm.

A. Guillevic and F. Morain have written a chapter [18] on discrete logarithm computations for a book on pairings.

7.6.1. *Breaking a MNT curve using DL computations*

There is a reduction between an elliptic curve E defined over $\mathbf{F}_p$ and a finite extension of degree k (aka *embedding degree*) of the base field, using pairing computations. In brief, one can transport the discrete logarithm problem from E to $\mathbf{F}_{p^k}$. If k is relatively small, this yields a DLP much easier to solve than directly on E. To give some highlight on current easyness, A. Guillevic, F. Morain and E. Thomé (from CARAMBA EPC in LORIA) computed a discrete log on a curve of embedding degree 3 and cryptographic size. This clearly showed that curves with small embedding degrees are indeed weak. The article [14] was presented by A. Guillevic during the SAC 2016 conference in New Foundland.

7.7. Rank metric codes over infinite fields

Rank metric and Gabidulin codes over the rationals promise interesting applications to space-time coding. We have constructed optimal codes, similar to Gabidulin codes, in the case of infinite fields. We use algebraic extensions, and we have determined the condition on the considered extension to enable this construction. For example: we can design codes with complex coefficients, using number fields and Galois automorphisms. Then, in the rank metric setting, codewords can be seen as matrices. In this setting, a channel introduces errors (a matrix of small rank r added to the codeword) and erasures (s_r rows and s_c columns of the matrix are erased). We have developed an algorithm (adapted from the Welch–Berlekamp algorithm) to recover the right codeword in the presence of an error of rank weight up to $r + s_c + s_r \leqslant d - 1$, where d is the minimal distance of the code. As opposed to the finite field case, we are confronted by coefficient size growth. We solve this problem by computing modulo prime ideals. Using these codes we can completely bypass intermediate constructions using finite fields, which were the stumbling-block in classic constructions.

We also have used this framework to build rank-metric codes over the field of rational functions, using algebraic function fields with cyclic Galois group (Kummer and Artin extensions). These codes can be seen as a generator of infinitely many convolutional codes.

7.8. Hash function cryptanalysis

Cryptographic hash functions are versatile primitives that are used in many cryptographic protocols. The security of a hash function h is usually evaluated through two main notions: its preimage resistance (given a target t, the difficulty of finding a message m s.t. $h(m) = t$) and its collision resistance (the difficulty of finding two messages m, m' s.t. $h(m) = h(m')$).

A popular hash function is the SHA-1 algorithm. Although theoretical collision attacks were found in 2005, it is still being used in some applications, for instance as the hash function in some TLS certificates. Hence cryptanalysis of SHA-1 is still a major topic in cryptography.

In 2015, we improved the state-of-the-art on SHA-1 analysis in two ways:

- T. Espitau, P.-A. Fouque and P. Karpman improved the previous preimage attacks on SHA-1, reaching up to 62 rounds (out of 80), up from 57. The corresponding paper was published at CRYPTO 2015.

- P. Karpman, T. Peyrin and M. Stevens developed collision attacks on the compression function of SHA-1 (i.e. freestart collisions). This exploits a model that is slightly more generous to the attacker in order to find explicit collisions on more rounds than what was previously possible. A first work resulted in freestart collisions for SHA-1 reduced to 76 steps; this attack takes less than a week to compute on a common GPU. The corresponding paper was published at CRYPTO 2015. This was later improved to attack the full compression function. Although the attack is more expensive it is still practical, taking less than two weeks on a 64 GPU cluster. The corresponding paper was accepted at EUROCRYPT 2016 [17].

7.9. Block cipher design and analysis

Block ciphers are one of the most basic cryptographic primitives, yet block cipher analysis is still a major research topic. In recent years, the community also shifted focus to the more general setting of *authenticated encryption*, where one specifies an (set of) algorithm(s) providing both encryption and authentication for messages of arbitrary length. A major current event in that direction is the CAESAR academic competition, which aims to select a portfolio of good algorithms.

In 2015, we helped to improve the state of the art in block cipher research in several ways:

- P. Karpman developed a compact 8-bit S-box with branch number three, which can be used as a basis to construct a lightweight block cipher particularly efficient on 8-bit microcontrollers [23].

In 2016, together with P.-A. Fouque, P. Kirchner and B. Minaud, P. Karpman designed a family of efficient provably incompressible symmetric primitives, which corresponds to a weak notion of white-box cryptography. The objective of such algorithms is that given an implementation of a certain target size, an adversary shouldn't be able to efficiently find a smaller implementation with comparable functionality. We introduced a security model that captures the behaviour of realistic adversaries and used this model to prove the security of a family of block cipher and a family of key generating functions. The corresponding paper was published at ASIACRYPT 2016 [13].

7.10. Weight distribution of Algebraic-Geometry codes

V. Ducet worked on the weight distribution of geometric codes following a method initiated by Duursma. More precisely he implemented his method in magma and was able to compute the weight distribution of the geometric codes coming from two optimal curves of genus 2 and 3 over the finite fields of size 16 and 9 respectively. The aim is to compute the weight distribution of the Hermitian code over the finite field of size 16, for which computational improvements of the implementation are necessary.

7.11. Update on the Chor-Rivest cryptosystem

The Chor-Rivest cryptosystem from the 90's was "broken" by Vaudenay. However, Vaudenay's attack applies only for the range of parameters originally proposed. The major recent breakthrough in discrete logarithm computations enable to redesign the system with a completely different range of paramaters, possibly thwarting Vaudenay's attack. D. Augot and C. Barbin tried to find a new attack against this discrete log and knapsack-based cryptosystem, using the Sidelnikov-Shestakov algorithm for recovering a Reed-Solomon code. Apparently, our new attack does not outperform S. Vaudenay's original attack, and it may be possible that the Chor-Rivest could be redesigned in a secure way.

7.12. Proofs or Retrievability

A Proof of Retrievability (PoR) is a cryptographic protocol which aims at ensuring a user that he can retrieve files he previously stored on a server. J. Lavauzelle and F. Levy-dit-Vehel studied a new approach for the construction of PoRs. The idea is to encode the file so that the user can check with low communication whether its file has been damaged. Such an encoding can be efficiently done with locally decodable and testable codes, and especially with the family of lifted codes introduced by Guo, Kopparty and Sudan [30]. In practice, PoRs thus defined achieve very efficient storage overhead and acceptable communication, compared to the existing litterature. This new construction [15] has been presented during the ISIT2016 conference in Barcelona.

7.13. Fast Encoding of Multiplicity Codes

N. Coxon has produced a fast implementation which demonstrates that the multiplicity codes from Kopparty, Saraf and Yehkanin are indeed practical for very large databases (when used in the Private Information Retrieval setting). For instance, we can encode a 10^8 bit longs message in two seconds on a regular laptop, and 10^9 in thirty seconds. We envisioned a scenario where DNA sequences are encoded using these multiplicity codes: 10^8 bits is the size of Drosophila melanogaster (flies), and 10^9 bits is the order of magnitude of the human genome.

7.14. Private Information Retrieval

Imagine the following scenario, in which a researcher wants to access many sustrings a DNA sequences, while maintaining the privacy of the request. The privacy or the secrecy of the database is not a concern here: for instance, this researcher wants to access many DNA subsequences of drosophila melanogaster, hosted on a remote data broker, and clearly the concern is not to protect the private life of flies. But the information leaked about the queries may endanger the novel aspect of the discovery the researcher is about to make, by revealing which DNA sequences he is studying.

Private Information Retrieval (PIR) schemes are designed to achieve this goal: a user queries a database T hosted on a remote server, and wants the i-th entry, i.e. $T[i]$. A cryptographic protocol is then run, and at the end of the protocol, the server must not know i, neither the $T[i]$ he answered, yet the user gets $T[i]$.

These PIR schemes can be achieved in an unconditionally secure way using the above Multiplicity codes, which N. Coxon made practical. In September, we explained this scenario and demoed our software at Nokia Bell Lab's Future X days a use case of Multiplicity codes for private access to DNA sequences.

7.15. Compact McEliece Keys from Algebraic-geometry codes

In 1978, McEliece [31], introduced a public key cryptosystem based on linear codes and suggested to use classical Goppa codes which belong to the family of alternant codes. This proposition remains secure but leads to very large public keys compared to other public-key cryptosystems. Many proposals have been made in order to reduce the key size, in particular quasi-cyclic alternant codes. Quasi-cyclic alternant codes refer to alternant codes admitting a generator matrix made of severals cyclic bloks. These alternant codes contains weakness because they have a non-trivial automorphism group. Thanks to this property we can build, from a quasi-cyclic alternant code, an alternant code with smaller parameters which has almost same private elements than the original code. Faugère, Otmani, Tillich, Perret and Portzamparc [29] showed this fact for alternant codes obtained by using supports $x \in \mathbb{F}_{q^m}^n$ globally stable by an affine map $\phi : z \mapsto az + b$, with $a, b \in \mathbb{F}_{q^m}^n$. E. Barelli has extended this proof to the non-affine case: for all codes obtained by using supports $x \in \mathbb{F}_{q^m}^n$ globally stable by a map $\phi : z \mapsto \frac{az+b}{cz+d}$, with $a, b, c, d \in \mathbb{F}_{q^m}^n$.

In order to suggest compact keys for the McEliece cryptosystem E. Barelli and A. Couvreur studied quasi-cyclic alternant gemeotric codes. Alternant geometric codes means a subfield subcode of an algebraic-geometry codes. To build these codes, we need curves with automorphisms. In particular, we studied Kummer cover of plane curves.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Grants with Industry

8.1.1. Nokia (ex Alcatel-Lucent)

Within the framework of the joint lab Inria-ALU, Grace and Alcatel-Lucent collaborate on the topic of Private Information Retrieval: that is, enabling a user to retrieve data from a remote database while revealing neither the query nor the retrieved data. (This is not the same as data confidentiality, which refers to the need for users to ensure secrecy of their data; this is classically obtained through encryption, which prevents access to data in the clear.)

A typical application would be a centralized database of medical records, which can be accessed by doctors, nurses, and so on. A desirable privacy goal would be that the central system does not know which patient is queried for when a query is made, and this goal is precisely achieved by a Private Information Retrieval protocol. Note also that in this scenario the database is not encrypted, since many users are allowed to access it.

We are exploring applications of Locally Decodable Codes to Private Information Retrieval in the multi-cloud (multi-host) setting, to ensure both secure, reliable storage, and privacy of database queries.

N. Coxon made the first implementation of these codes, who are indeed very practical. On a laptop, we can encode an ADN of a drosophilia in two seconds, and a 10^9 bit data base in 30 seconds. We have a few real-life scenario in mind (DNA, geolocalisation, streaming), and we will check how realistic they are.

8.1.2. Safran Identity and Security (ex-Morpho)

A contract has been signed in November 2016 between Safran Identity and Security and École polytechnique, for one year post-doc position. A candidate has been found, and will arrive early 2017 (January).

The topic is the research is to use bitcoin's blockchain to issue and manipulate certification of identities, which is very close to the (trendy) topic of diplomation with blockchains.

Safran had a preliminary construction for doing that, and a preliminary version has been submitted to the IEEE Security and Privacy on the Blockchain Workshop.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. PEPS Aije-bitcoin

Within the group PAIP (Pour une Approche Interdisciplinaire de la Privacy), D. Augot presented the cryptographic and peer-to-peer principles at the heart of the Bitcoin protocol (electronic signature, hash functions, and so on). Most of the information is publicly available: the history of all transactions, evolution of the source code, developers' mailing lists, and the Bitcoin exchange rate. It was recognized by the economists in our group that such an amount of data is very rare for an economic phenomenon, and it was decided to start research on the history of Bitcoin, to study the interplay between the development of protocol and the development of the economical phenomenon.

The project **Aije-Bitcoin** (analyse informatique, juridique et economique de Bitcoin) was accepted as interdisciplinary research for a PEPS (Projet exploratoire Premier Soutien) cofunded by the CNRS and Universite de Paris-Saclay. This one-year preliminary program will enable the group to master the understanding of Bitcoin from various angles, allowing more advanced research in the following years.

One M2 intern, E. Palazzollo, was intern in Sceaux, with aim to qualify the nature of bitcoin, as an asset, curency, etc.

This project ended in March 2016

9.1.2. IDEALCODES

Idealcodes is a two-year Digiteo research project, started in October 2014. The partners involved are the École Polytechnique (X) and the Université de Versailles–Saint-Quentin-en-Yvelines (Luca de Feo, UVSQ). After hiring J. Nielsen the first year, we have hired V. Ducet for the second year, both working at the boundary between coding theory, cryptography, and computer algebra

Idealcodes spans the three research areas of algebraic coding theory, cryptography, and computer algebra, by investigating the problem of lattice reduction (and root-finding). In algebraic coding theory this is found in Guruswami and Sudan's list decoding of algebraic geometry codes and Reed–Solomon codes. In cryptography, it is found in Coppersmith's method for finding small roots of integer equations. These topics were unified and generalised by H. Cohn and N. Heninger [25], by considering algebraic geometry codes and number field codes under the deep analogy between polynomials and integers. Sophisticated results in coding theory could be then carried over to cryptanalysis, and vice-versa. The generalized view raises problems of computing efficiently, which is one of the main research topics of Idealcodes.

The last year of the one-year project aims to find matrices with good diffusion properties over small finite fields. The principle is to find non-maximal matrices, but with better coefficients and implementation properties. The relevant cryptographic properties to be studied correspond to the weight distribution of the associated code. Since we use Algebraic-Geometry codes, much more powerful techniques can be used for computing these weight distribution, using and improving Duursma's ideas [28].

9.1.3. IRT System-X

D. Augot is co-advising a PhD candidate, H.-M. Bisserier, on "les relations contractuelles de droit privé à l'épreuve de la technologie des blockchains", i.e. on (French) law and so-called "smart contracts". D. Augot will mainly help H.-M. Bisserier to clarify the essential computer science topics and issues relevant to the most important blockchains (bitcoin, ethereum). Then H.-M. Bisserier will be advised by C. Zolynksi for remaining two years, fixing research directions.

9.2. National Initiatives

9.2.1. ANR

MANTA (accepted July 2015, starting March 2016): "Curves, surfaces, codes and cryptography". This project deals with applications of coding theory error correcting codes to in cryptography, multi-party computation, and complexity theory, using advanced topics in algebraic geometry and number theory. The kickoff was a one week-retreat in Dordogne (20 participants), and we had another four day meeting in Saclay in November 17. See http://anr-manta.inria.fr/.

9.2.2. DGA

Cybersecurity. Inria and DGA contracted for three PhD topics at the national level, one of them involving Grace. Grace started a new PhD, and hired P. Karpman. The topic of this PhD is complementary to the above DIFMAT-3: while DIFMAT-3 provides fundamental methods for dealing with AG codes, in application for diffusion layers in block ciphers, the topic here is to make concrete propositions of block ciphers using these matrices. P. Karpman is coadvised by T. Peyrin (Nanyang Technological University, Singapore), by P.-A. Fouque (Université de Rennes), and D. Augot.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. PQCRYPTO

Title: Post-quantum cryptography for long-term security

Programm: H2020

Duration: March 2015 - March 2018

Coordinator: TECHNISCHE UNIVERSITEIT EINDHOVEN

Partners:

Academia Sinica (Taiwan)

Bundesdruckerei (Germany)

Danmarks Tekniske Universitet (Denmark)

Katholieke Universiteit Leuven (Belgium)

Nxp Semiconductors Belgium Nv (Belgium)

Ruhr-Universitaet Bochum (Germany)

Stichting Katholieke Universiteit (Netherlands)

Coding Theory and Cryptology group, Technische Universiteit Eindhoven (Netherlands)

Technische Universitaet Darmstadt (Germany)

University of Haifa (Israel)

Inria contact: Nicolas Sendrier

Online security depends on a very few underlying cryptographic algorithms. Public-key algorithms are particularly crucial since they provide digital signatures and establish secure communication. Essentially all applications today are based on RSA or on the discrete-logarithm problem in finite fields or on elliptic curves. Cryptographers optimize parameter choices and implementation details for these systems and build protocols on top of these systems; cryptanalysts fine-tune attacks and establish exact security levels for these systems.

It might seem that having three systems offers enough variation, but these systems are all broken as soon as large quantum computers are built. The EU and governments around the world are investing heavily in building quantum computers; society needs to be prepared for the consequences, including cryptanalytic attacks accelerated by these computers. Long-term confidential documents such as patient health-care records and state secrets have to guarantee security for many years, but information encrypted today using RSA or elliptic curves and stored until quantum computers are available will then be as easy to decipher.

PQCRYPTO will allow users to switch to post-quantum cryptography: cryptographic systems that are not merely secure for today but that will also remain secure long-term against attacks by quantum computers. PQCRYPTO will design a portfolio of high-security post-quantum public-key systems, and will improve the speed of these systems, with reference implementations.

Our team is engaged in WP3.3 "advanced applications for the cloud". We envision to focus essentially on secure multiparty computation, essentially the information theoretically secure constructions, who are naturally secure against a quantum computer invoked on classical queries. We will study whether these protocols still resist quantum queries. This work sub package started March 2015, and is dealt with by D. Augot.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. Member of the Organizing Committees

- D. Augot is member of the committee of the CCA seminar on coding and cryptology. This seminar regularly attracts around 30 participants.

10.1.2. Scientific Events Selection

10.1.2.1. Reviewer

- D. Augot was reviewer for International Symposium on Information Theory

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- D. Augot is member of the editorial board of the *RAIRO - Theoretical Informatics and Applications*, a Cambridge journal published by EDP Sciences.
- D. Augot is member of the editorial board of the *International Journal of Information and Coding Theory*, InderScience publishers.
- F. Morain is member of the editorial board of the *Applicable Algebra in Engineering, Communication and Computing*, Springer.
- A. Couvreur was editor with Alp Bassa (Bogazici University, Turkey) and David Kohel (Aix-Marseille University) of a number of *AMS Contemporary Mathematics* for the proceedings of the conference AGCT (*Arithmetic Geometry Cryptography and Coding Theory*) 2015.

10.1.3.2. Reviewer - Reviewing Activities

- D. Augot was reviewer for
  – Discrete Mathematics
  – Designs, Codes and Cryptography
  – Linear and Multilinear Algebra
  – Finite Fields and their applications
- A. Couvreur was reviewer for
  – Discrete Mathematics
  – Designs, Codes and Cryptography
  – Journal of Algebra

10.1.4. Invited Talks

- D. Augot was invited speaker at Yet Another Cryptography Conference (YACC), Porquerolles, June 2016.
- B. Smith was an invited speaker at the 20th international Workshop on Elliptic Curve Cryptography (ECC), Izmir, Turkey, September 2016.
- A. Couvreur gave a talk to represent the group *Codes et Cryptographie* of the GdR *Informatique Mathématiques* (GdR IM) at the *Journées nationales du GdR IM* at University Paris 13 (January 13).

10.1.5. Scientific Expertise

- D. Augot participated in a round table at a workshop organized by French National Assembly (lower house) at, on blockchains (March 24th).
- D. Augot participated in a round table at Paris Dauphine on blockchains, organized by the chair "Chaire Gouvernance & Régulation" (November 1).
- D. Augot made a talk on hashing and blockchain at a workshop on blockchains held at Institut Poincaré (November 16).

10.1.6. Teaching in international postgraduate summer schools

- B. Smith gave lectures on *Basic public-key constructions with elliptic curves* and *Advanced constructions in curve-based cryptography* at the *Summer school on real-world crypto and privacy*, Sibenik, Croatia, June 2016.

- B. Smith gave a course on *asymmetric cryptography and elliptic curves* at the *Crypto-CO summer school on cryptography and security*, Bogota, Colombia, July 2016.
- B. Smith gave lectures on elliptic curves at the *ECC2016 Computational Algebraic Number Theory School*, Izmir, Turkey, September 2016.

10.1.7. Research Administration

Committees

- A. Couvreur is an elected member of Saclay's *comité de centre*.
- A. Couvreur is an elected member of Saclay's *Comité local Hygiène, Sécurité et Conditions de Travail*.
- A. Couvreur is the *jeune chercheur référent* for the *commission de suivi doctoral* of Inria Saclay.
- D. Augot is a member of LIX's *conseil de direction*.
- D. Augot is the vice-head of Inria's *comité de suivi doctoral*
- D. Augot is a member of LIX's *assemblée des chefs d'équipe*
- D. Augot is elected member of the *conseil académique consultatif* of Paris-Saclay University.
- F. Levy-dit-Vehelis a representative of "enseignants-chercheurs" of LIX.
- F. Morain, B. Smith and A. Couvreur are elected members of the *Conseil de Laboratoire* of the LIX.
- F. Morain is vice-head of the Département d'informatique of Ecole Polytechnique.
- F. Morain represents École polytechnique in the committee in charge of *Mention HPC* in the *Master de l'université Paris Saclay*.
- F. Morain is member of the Board of Master Parisien de Recherche en Informatique (MPRI).
- B. Smith is a *Correspondant* for International Relations at Saclay.
- B. Smith is a member of the COST-GTRI.
- B. Smith is a member of the teaching committee of the Department of Computer Science of the École polytechnique.
- B. Smith is the academic coordinator for Computer Science in the new *Bachelor* program at École polytechnique.

Committees

- D. Augot was in the committee assessing candidates for Univ. Paris 8.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence :

- D. Augot was mentoring a group of polytechnique students on a L3 projet on homomorphic encryption and voting (6 students)
- D. Augot was mentoring a group of polytechnique students on a L3 projet on blockchains and hyperledger, in collaboration with Orange (5 students)
- F. Levy-dit-Vehel, "Mathématiques discrètes pour la protection de l'information", 24h (equiv TD), 2nd year (L3), ENSTA ParisTech, France.
- J. Lavauzelle, 1I002, "Introduction à la programmation en C", tutorial class (38.5h), L1, Université Pierre et Marie Curie, France
- J. Lavauzelle, 2I011, "Méthodes numériques", tutorial class (21h), L2, Université Pierre et Marie Curie, France

- J. Lavauzelle, 1I001, "Éléments de programmation", tutorial class (38.5h), L1, Université Pierre et Marie Curie, France
- J. Lavauzelle, 2I003, "Initiation à l'algorithmique", tutorial class (21.25h), L2, Université Pierre et Marie Curie, France
- A. Couvreur and E. Barelli, INF311, "Introduction à l'informatique", 26.7h(equiv TD), 1st year, Ecole Polytechnique, France.
- E. Barelli, INF411, "Les bases de la programmation et de l'algorithmique", 21.3h (equiv TD), 2nd year (L3), Ecole Polytechnique, France.
- B. Smith, INF442, "Traitement des données massives", 32h TD, 2nd year, École polytechnique
- A. Couvreur and B. Smith, INF411, "Les bases de la programmation et de l'algorithmique", 32h TD, 2nd year, École polytechnique

Master :

- D. Augot was mentoring François Bonnal, on a M1 research training projet, "bitcoin malleability"
- D. Augot was mentoring Édouard Dufour-Sans, on a M1 research training projet, "symmetric information theoretically secure private information retrieval schemes and applications"
- F. Levy-dit-Vehel, "Cours de Cryptographie", 30h. (equiv TD), 3rd year (M1), ENSTA ParisTech, France.
- B. Smith, "Algorithmes arithmétiques pour la cryptologie", 15h, MPRI (M2), Paris
- A. Couvreur, INF558a, "Introduction to cryptology", 25h, Ecole Polytechnique (M1).
- A. Couvreur, "Introduction to coding theory and cryptology", 10h, MPRI (M2), Paris.
- B. Smith supervised Nagarjun Chinthamani Dwarakanath for a 3A project and an M1 project on efficient curve-based cryptosystems at École polytechnique
- A. Couvreur supervised Evrim Petek's M2 internship on the power decoding algorithm.
- A. Couvreur supervised Anas Aarab's M1 TRE (*Travail de Recherche Encadré*) on the decoding of Reed Solomon codes.

Doctorat :

- Ben Smith made a lecture at the spring school on coding and cryptology at La Chapelle-Gauthier.

10.2.2. *Supervision*

- PhD in progress. J. Lavauzelle has began his Ph.D. on locally decodable codes and cryptogra[hic applications, on October 1st, 2015, under the supervision of D. Augot and F. Levy-dit-Vehel.
- PhD in progress. E. Barelli has begun his PhD on Algebraic-Geometry codes for code-based crypto on October 1st, 2015, under the supervision of D. Augot and A. Couvreur.
- PhD in progress. N. Duhamel has begun his PhD on genus 2 curves for cryptography, under the supervision of B. Smith and F. Morain.
- Completed PhD. P. Karpman, starting in 2013, defended in November 2016 his PhD on security of symmetric crytographic primitives.

10.2.3. *Juries*

- D. Augot was examiner in the jury of Fanny Jardel, who defended her thesis "Calcul et Stockage Distribués pour les Réseaux de Communication", January 11, Télécom-ParisTech
- D. Augot was examiner in the jury of Cécile Pierrot, who defended her thesis "Le logarithme discret dans les corps finis", November 25, Pierre and Marie Curie University.

- F. Morain was referee and examiner in the jury of Alexandre WALLET, who defended his thesis "Le problème de décomposition de points dans les variétés jacobiennes", December 14, Pierre and Marie Curie University.
- A. Couvreur is member of the jury of the agrégation de mathématiques and coordinator of option C ("algèbre et calcul formel).

10.3. Popularization

- At the occasion of Nokia Bell Labs Future X-Days, September 2016, D. Augot, N. Coxon and F. Levy-dit-Vehel demoed N. Coxon's implementation of a code based *private information retrieval scheme*
- D. Augot made a two hours lecture on bitcoin to the French *institut des actuaires*.

11. Bibliography

Major publications by the team in recent years

[1] D. AUGOT, M. FINIASZ.*Direct Construction of Recursive MDS Diffusion Layers using Shortened BCH Codes*, in "21st International Workshop on Fast Software Encryption, FSE 2014", London, United Kingdom, C. CID, C. RECHBERGER (editors), springer, March 2014, https://hal.inria.fr/hal-01044597.

[2] A. COUVREUR, I. MÁRQUEZ-CORBELLA, R. PELLIKAAN.*A Polynomial Time Attack against Algebraic Geometry Code Based Public Key Cryptosystems*, in "Information Theory (ISIT), 2014 IEEE International Symposium on", Honolulu, United States, IEEE, June 2014, p. 1446-1450 [*DOI : 10.1109/ISIT.2014.6875072*], https://hal.archives-ouvertes.fr/hal-00937476.

[3] A. COUVREUR, A. OTMANI, J.-P. TILLICH.*Polynomial Time Attack on Wild McEliece Over Quadratic Extensions*, in "EUROCRYPT 2014", Copenhagen, Denmark, May 2014, p. 17-39, https://hal.archives-ouvertes.fr/hal-00931774.

[4] P. LEBACQUE, A. ZYKIN.*On the Number of Rational Points of Jacobians over Finite Fields*, in "Acta Arith.", 2015, vol. 169, p. 373–384, https://hal.archives-ouvertes.fr/hal-01081468.

[5] F. MORAIN.*Implementing the asymptotically fast version of the elliptic curve primality proving algorithm*, in "Math. Comp.", 2007, vol. 76, p. 493–505.

[6] B. SMITH.*Isogenies and the discrete logarithm problem in Jacobians of genus 3 hyperelliptic curves*, in "J. of Cryptology", 2009, vol. 22, n° 4, p. 505-529.

[7] B. SMITH.*Families of fast elliptic curves from Q-curves*, in "Advances in Cryptology - ASIACRYPT 2013", Bangalore, India, K. SAKO, P. SARKAR (editors), Lecture Notes in Computer Science, Springer, December 2013, vol. 8269, p. 61-78 [*DOI : 10.1007/978-3-642-42033-7_4*], https://hal.inria.fr/hal-00825287.

Publications of the year

Articles in International Peer-Reviewed Journal

[8] A. COUVREUR.*An upper bound on the number of rational points of arbitrary projective varieties over finite fields*, in "Proceedings of the American Mathematical Society", 2016, vol. 144, p. 3671-3685, https://hal.archives-ouvertes.fr/hal-01069510.

[9] A. COUVREUR, A. OTMANI, J.-P. TILLICH.*Polynomial Time Attack on Wild McEliece Over Quadratic Extensions*, in "IEEE Transactions on Information Theory", 2017, vol. 63, n° 1, p. 404 - 427 [*DOI :* 10.1109/TIT.2016.2574841], https://hal.archives-ouvertes.fr/hal-01426775.

[10] F. MORAIN, C. SCRIBOT, B. SMITH.*Computing cardinalities of Q-curve reductions over finite fields*, in "LMS Journal of Computation and Mathematics", August 2016, vol. 19, n^o A, 15 [*DOI :* 10.1112/S1461157016000267], https://hal.inria.fr/hal-01320388.

[11] B. SMITH.*The Q-curve construction for endomorphism-accelerated elliptic curves*, in "Journal of Cryptology", October 2016, vol. 29, n° 4, 27 [*DOI :* 10.1007/s00145-015-9210-8], https://hal.inria.fr/hal-01064255.

International Conferences with Proceedings

[12] P. FOUQUE, P. KARPMAN, P. KIRCHNER, B. MINAUD.*Efficient and Provable White-Box Primitives*, in "ASIACRYPT 2016", HANOI, Vietnam, Advances in Cryptology - ASIACRYPT 2016 - 22nd International Conference on the Theory and Application of Cryptology and Information Security, December 2016, vol. LNCS 10031, p. 159 - 188 [*DOI :* 10.1007/978-3-662-53887-6_6], https://hal.archives-ouvertes.fr/hal-01427810.

[13] P.-A. FOUQUE, P. KARPMAN, P. KIRCHNER, B. MINAUD.*Efficient and Provable White-Box Primitives*, in "ASIACRYPT", Hanoï, Vietnam, Springer, December 2016 [*DOI :* 10.1007/978-3-662-53887-6_6], https://hal.inria.fr/hal-01421044.

[14] A. GUILLEVIC, F. MORAIN, E. THOMÉ.*Solving discrete logarithms on a 170-bit MNT curve by pairing reduction*, in "Selected Areas in Cryptography 2016", St. John's, Canada, R. AVANZI, H. HEYS (editors), Selected Areas in Cryptography 2016, Springer, August 2016, to appear in the Lecture Notes in Computer Science (LNCS), https://hal.inria.fr/hal-01320496.

[15] J. LAVAUZELLE, F. LEVY-DIT-VEHEL.*New proofs of retrievability using locally decodable codes*, in "International Symposium on Information Theory ISIT 2016", Barcelona, Spain, July 2016, p. 1809 - 1813 [*DOI :* 10.1109/ISIT.2016.7541611], https://hal.archives-ouvertes.fr/hal-01413159.

[16] J. RENES, P. SCHWABE, B. SMITH, L. BATINA.*μKummer: efficient hyperelliptic signatures and key exchange on microcontrollers*, in "Cryptographic Hardware and Embedded Systems – CHES 2016", Santa Barbara, United States, Cryptographic Hardware and Embedded Systems – CHES 2016, Springer-Verlag, August 2016, vol. 9813, 20 [*DOI :* 10.1007/978-3-662-53140-2_15], https://hal.inria.fr/hal-01300727.

[17] M. STEVENS, P. KARPMAN, T. PEYRIN.*Freestart Collision for Full SHA-1*, in "EUROCRYPT 2016", Vienne, Austria, IACR, May 2016 [*DOI :* 10.1007/978-3-662-49890-3_18], https://hal.inria.fr/hal-01251023.

Scientific Books (or Scientific Book chapters)

[18] A. GUILLEVIC, F. MORAIN.*Discrete Logarithms*, in "Guide to pairing-based cryptography", N. E. MRABET, M. JOYE (editors), CRC Press - Taylor and Francis Group, December 2016, 42, https://hal.inria.fr/hal-01420485.

Other Publications

[19] S. BALLENTINE, A. GUILLEVIC, E. LORENZO GARCÍA, C. MARTINDALE, M. MASSIERER, B. SMITH, J. TOP.*Isogenies for point counting on genus two hyperelliptic curves with maximal real multiplication*, December 2016, working paper or preprint, https://hal.inria.fr/hal-01421031.

[20] P. N. CHUNG, C. COSTELLO, B. SMITH.*Fast, uniform scalar multiplication for genus 2 Jacobians with fast Kummers*, 2016, working paper or preprint, https://hal.inria.fr/hal-01353480.

[21] A. COUVREUR, I. MÁRQUEZ-CORBELLA, R. PELLIKAAN.*Cryptanalysis of McEliece Cryptosystem Based on Algebraic Geometry Codes and their subcodes*, March 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01280927.

[22] F. GROSSHANS, T. LAWSON, B. SMITH, F. MORAIN.*Factoring Safe Semiprimes with a Single Quantum Query*, September 2016, working paper or preprint, https://hal.inria.fr/hal-01229587.

[23] P. KARPMAN.*Exercice de style*, January 2016, working paper or preprint, https://hal.inria.fr/hal-01263735.

References in notes

[24] B. AUDOUX, A. COUVREUR.*On tensor products of CSS Codes*, December 2015, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01248760.

[25] H. COHN, N. HENINGER.*Ideal forms of Coppersmith's theorem and Guruswami-Sudan list decoding*, in "Innovations in Computer Science - ICS 2010, Tsinghua University, Beijing, China, January 7-9, 2011. Proceedings", B. CHAZELLE (editor), Tsinghua University Press, 2011, p. 298-308.

[26] A. COUVREUR, P. GABORIT, V. GAUTHIER-UMANA, A. OTMANI, J.-P. TILLICH.*Distinguisher-based attacks on public-key cryptosystems using Reed-Solomon codes*, in "Designs, Codes and Cryptography", 2014, vol. 73, n° 2, p. 641-666 [*DOI :* 10.1007/S10623-014-9967-Z], https://hal.archives-ouvertes.fr/hal-01096172.

[27] A. COUVREUR, I. MÁRQUEZ-CORBELLA, R. PELLIKAAN.*Cryptanalysis of public-key cryptosystems that use subcodes of algebraic geometry codes*, in "4th ICMCTA - Fourth International Castle Meeting on Coding Theory and Applications", Palmela, Portugal, September 2014, https://hal.inria.fr/hal-01069272.

[28] I. M. DUURSMA.*Weight distributions of geometric Goppa codes*, in "Trans. Amer. Math. Soc.", 1999, vol. 351, n° 9, p. 3609–3639, http://dx.doi.org/10.1090/S0002-9947-99-02179-0.

[29] J.-C. FAUGERE, A. OTMANI, L. PERRET, F. DE PORTZAMPARC, J.-P. TILLICH.*Folding alternant and Goppa codes with non-trivial automorphism groups*, in "IEEE Transactions on Information Theory", 2016, vol. 62, n° 1, p. 184–198.

[30] A. GUO, S. KOPPARTY, M. SUDAN.*New Affine-invariant Codes from Lifting*, in "Proceedings of the 4th Conference on Innovations in Theoretical Computer Science", New York, NY, USA, ITCS '13, ACM, 2013, p. 529–540, http://doi.acm.org/10.1145/2422436.2422494.

[31] R. J. MCELIECE.*A Public-Key System Based on Algebraic Coding Theory*, Jet Propulsion Lab, 1978, p. 114–116, DSN Progress Report 44.

[32] V. SIDELNIKOV, S. SHESTAKOV.*On the insecurity of cryptosystems based on generalized Reed-Solomon codes*, in "Discrete Math. Appl.", 1992, vol. 1, n° 4, p. 439-444.

Project-Team ILDA

Interacting with Large Data

IN COLLABORATION WITH: Laboratoire de recherche en informatique (LRI)

IN PARTNERSHIP WITH:
CNRS

Université Paris-Sud (Paris 11)

RESEARCH CENTER
Saclay - Île-de-France

THEME
Interaction and visualization

Table of contents

Project-Team ILDA

Creation of the Team: 2015 January 01, updated into Project-Team: 2016 December 01

Keywords:

Computer Science and Digital Science:

3.1.7. - Open data
3.2.4. - Semantic Web
5.1. - Human-Computer Interaction
5.1.1. - Engineering of interactive systems
5.1.2. - Evaluation of interactive systems
5.1.5. - Body-based interfaces
5.1.6. - Tangible interfaces
5.2. - Data visualization

Other Research Topics and Application Domains:

7.1. - Traffic management
9.4.3. - Physics
9.4.5. - Data science
9.5.7. - Geography
9.7.2. - Open data
9.9. - Risk management

1. Members

Research Scientists
Emmanuel Pietriga [Team leader, Senior Researcher, Inria, (DR2), HDR]
Caroline Appert [Researcher, CNRS, (CR1)]
Olivier Chapuis [Researcher, CNRS, (CR1)]

Faculty Member
Anastasia Bezerianos [Univ. Paris XI, Faculty Member (Assistant Professor)]

Technical Staff
Hande Ozaygen [SED Engineer, associate team member]
Olivier Gladin [SED Engineer, associate team member]

PhD Students
Anna Gogolou [Inria, from Oct 2016]
Maria Jesus Lobo Gunther [Inria]
Rafael Morales Gonzalez [Univ. Paris XI]
Arnaud Prouzeau [Univ. Paris XI]
Hugo Romat [Tecknowmetrix, funded by ANRT/CIFRE]
Evanthia Dimara [Inria, associate team member, co-advised with project-team Aviz]
Bruno Fruchard [Institut Telecom, associate team member, co-advised with team VIA]

Post-Doctoral Fellow
André Spritzer [until October 2016, hosted by project-team Aviz]

Visiting Scientists
Grazia Prato [Inria Chile, UX designer]

Guy Riba [INRA]
Iftach Sadeh [CTA / DESY, post-doc]
Administrative Assistant
Alexandra Merlin [Inria]
Other
Dylan Lebout [Univ. Paris XI, M.Sc. student, from May to July 2016]

2. Overall Objectives

2.1. Overall Objectives

In an increasing number of domains, computer users are faced with large datasets, that are often interlinked and organized according to elaborate structures thanks to new data models such as those that are arising with the development of, e.g., the Web of Data. Rather than seeing the inherent complexity of those data models as a hindrance, we aim at leveraging it to design new interactive systems that can better assist users in their data understanding and processing tasks.

These "Data-centric Interactive Systems" aim at providing users with the right information at the right time, presenting it in the most meaningful manner, and letting users efficiently manipulate, edit and share these data with others. This entails minimizing the effort required to retrieve and relate data from relevant sources; displaying data using visual presentation techniques that match the data's characteristics and the users' tasks; and providing users with means of interacting with the data that effectively support their train of thought.

Our approach is based on the idea of bringing the fields of Web data management [29] and Human-computer interaction [52], [73] closer together, based on the strong belief that they have the potential to cross-fertilize one another. User interface design is essential to the management and understanding of large, interlinked datasets. Interlinked datasets enriched with even a small amount of semantics have the potential to help create interfaces that let users analyze and manipulate data in a more efficient manner by providing them with, e.g, more relevant query results and giving them efficient means to navigate and relate those results. Our ultimate, long-term goal is to design interactive systems that make it as straightforward to manipulate large webs of data as spreadsheets do for tabular data.

3. Research Program

3.1. Introduction

Our ability to acquire or generate, store, process, interlink and query data has increased spectacularly over the last few years. The corresponding advances are commonly grouped under the umbrella of so called *Big Data*. Even if the latter has become a buzzword, these advances are real, and they are having a profound impact in domains as varied as scientific research, commerce, social media, industrial processes or e-government. Yet, looking ahead, emerging technologies related to what we now call the *Web of Data* (a.k.a the Semantic Web) have the potential to create an even larger revolution in data-driven activities, by making information accessible to machines as semistructured data [28] that eventually becomes actionable knowledge. Indeed, novel Web data models considerably ease the interlinking of semi-structured data originating from multiple independent sources. They make it possible to associate machine-processable semantics with the data. This in turn means that heterogeneous systems can exchange data, infer new data using reasoning engines, and that software agents can cross data sources, resolving ambiguities and conflicts between them [71]. Datasets are becoming very rich and very large. They are gradually being made even larger and more heterogeneous, but also much more useful, by interlinking them, as exemplified by the Linked Data initiative [47].

These advances raise research questions and technological challenges that span numerous fields of computer science research: databases, communication networks, security and trust, data mining, as well as human-computer interaction. Our research is based on the conviction that interactive systems play a central role in many data-driven activity domains. Indeed, no matter how elaborate the data acquisition, processing and storage pipelines are, data eventually get processed or consumed one way or another by users. The latter are faced with large, increasingly interlinked heterogeneous datasets (see, e.g., Figure 1) that are organized according to complex structures, resulting in overwhelming amounts of both raw data and structured information. Users thus require effective tools to make sense of their data and manipulate them.
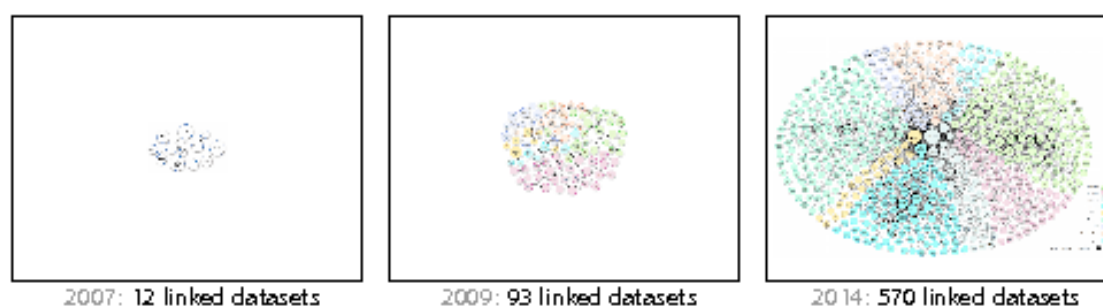
Figure 1. Linking Open Data cloud diagram from 2007 to 2014 – http://lod-cloud.net

We approach this problem from the perspective of the Human-Computer Interaction (HCI) field of research, whose goal is to study how humans interact with computers and inspire novel hardware and software designs aimed at optimizing properties such as efficiency, ease of use and learnability, in single-user or cooperative work contexts. More formally, HCI is about designing systems that lower the barrier between users' cognitive model of what they want to accomplish, and computers' understanding of this model. HCI is about the design, implementation and evaluation of computing systems that humans interact with [52], [73]. It is a highly multidisciplinary field, with experts from computer science, cognitive psychology, design, engineering, ethnography, human factors and sociology.

In this broad context, ILDA aims at designing interactive systems that display [37], [59], [80] the data and let users interact with them, aiming to help users better *navigate* and *comprehend* large webs of data represented visually, as well as *relate* and *manipulate* them.

Our research agenda consists of the three complementary axes detailed in the following subsections. Designing systems that consider interaction in close conjunction with data semantics is pivotal to all three axes. Those semantics will help drive navigation in, and manipulation of, the data, so as to optimize the communication bandwidth between users and data.

3.2. Semantics-driven Data Manipulation

Participants: Emmanuel Pietriga, Caroline Appert, Hande Ozaygen, Hugo Romat.

The Web of Data has been maturing for the last fifteen years and is starting to gain adoption across numerous application domains (Figure 1). Now that most foundational building blocks are in place, from knowledge representation, inference mechanisms and query languages [48], all the way up to the expression of data presentation knowledge [66] and to mechanisms like look-up services [79] or spreading activation [43], we need to pay significant attention to how human beings are going to interact with this new Web, if it is to *"reach its full potential"* [44].

Most efforts in terms of user interface design and development for the Web of data have essentially focused on tools for software developers or subject-matter experts who create ontologies and populate them [54], [42]. Tools more oriented towards end-users are starting to appear [34], [36], [49], [50], [53], [61], including the so-called *linked data browsers* [47]. However, those browsers are in most cases based on quite conventional point-and-click hypertext interfaces that present data to users in a very page-centric, web-of-documents manner that is ill-suited to navigating in, and manipulating, webs of data.

To be successful, interaction paradigms that let users navigate and manipulate data on the Web have to be tailored to the radically different way of browsing information enabled by it, where users directly interact with the data rather than with monolithic documents. The general research question addressed in this part of our research program is how to design novel interaction techniques that help users manipulate their data more efficiently. By data manipulation, we mean all low-level tasks related to manually creating new content, modifying and cleaning existing content, merging data from different sources, establishing connections between datasets, categorizing data, and eventually sharing the end results with other users; tasks that are currently considered quite tedious because of the sheer complexity of the concepts, data models and syntax, and the interplay between all of them.

Our approach is based on the conviction that there is a strong potential for cross-fertilization, as mentioned earlier: on the one hand, user interface design is essential to the management and understanding of webs of data; on the other hand, interlinked datasets enriched with even a small amount of semantics can help create more powerful user interfaces, that provide users with the right information at the right time.

We envision systems that focus on the data themselves, exploiting the underlying *semantics and structure* in the background rather than exposing them – which is what current user interfaces for the Web of Data often do. We envision interactive systems in which the semantics and structure are not exposed directly to users, but serve as input to the system to generate interactive representations that convey information relevant to the task at hand and best afford the possible manipulation actions.

3.3. Generalized Multi-scale Navigation

Participants: Olivier Chapuis, Emmanuel Pietriga, Caroline Appert, Anastasia Bezerianos, Olivier Gladin, Anna Gogolou, Maria Jesus Lobo Gunther, Arnaud Prouzeau.

The foundational question addressed here is what to display when, where and how, so as to provide effective support to users in their data understanding and manipulation tasks. ILDA targets contexts in which workers have to interact with complementary views on the same data, or with views on different-but-related datasets, possibly at different levels of abstraction. Being able to combine or switch between representations of the data at different levels of detail and merge data from multiple sources in a single representation is central to many scenarios. This is especially true in both of the application domains we consider: mission-critical systems (e.g., natural disaster crisis management) and the exploratory analysis of scientific data (e.g., correlate theories and heterogeneous observational data for an analysis of a given celestial body in Astrophysics).

A significant part of our research over the last ten years has focused on multi-scale interfaces. We designed and evaluated novel interaction techniques, but also worked actively on the development of open-source UI toolkits for multi-scale interfaces (see Section 6.2). These interfaces let users navigate large but relatively homogeneous datasets at different levels of detail, on both workstations [69], [31], [65], [64], [63], [32], [68], [30], [70] and wall-sized displays [5], [55], [67], [60], [33], [39], [38]. This part of the ILDA research program is about extending multi-scale navigation in two directions: 1. Enabling the representation of multiple, spatially-registered but widely varying, multi-scale data layers in Geographical Information Systems (GIS); 2. Generalizing the multi-scale navigation paradigm to interconnected, heterogeneous datasets as found on the Web of Data.

The first research problem is mainly investigated in collaboration with IGN in the context of ANR project MapMuxing (Section 9.2.1), which stands for *multi-dimensional map multiplexing*. Project MapMuxing aims at going beyond the traditional pan & zoom and overview+detail interface schemes, and at designing and evaluating novel cartographic visualizations that rely on high-quality generalization, *i.e.*, the simplification

of geographic data to make it legible at a given map scale [76], [77], and symbol specification. Beyond project MapMuxing, we are also investigating multi-scale multiplexing techniques for geo-localized data in the specific context of ultra-high-resolution wall-sized displays, where the combination of a very high pixel density and large physical surface (Figure 2) enable us to explore designs that involve collaborative interaction and physical navigation in front of the workspace. This is work done in cooperation with team Massive Data at Inria Chile.

The second research problem is about the extension of multi-scale navigation to interconnected, heterogeneous datasets. Generalization has a rather straightforward definition in the specific domain of geographical information systems, where data items are geographical entities that naturally aggregate as scale increases. But it is unclear how generalization could work for representations of the more heterogeneous webs of data that we consider in the first axis of our research program. Those data form complex networks of resources with multiple and quite varied relationships between them, that cannot rely on a single, unified type of representation (a role played by maps in GIS applications).

Addressing the limits of current generalization processes is a longer-term, more exploratory endeavor. Here again, the machine-processable semantics and structure of the data give us an opportunity to rethink how users navigate interconnected heterogeneous datasets. Using these additional data, we investigate ways to generalize the multi-scale navigation paradigm to datasets whose layout and spatial relationships can be much richer and much more diverse than what can be encoded with static linear hierarchies as typically found today in interfaces for browsing maps or large imagery. Our goal is thus to design and develop highly dynamic and versatile multi-scale information spaces for heterogeneous data whose structure and semantics are not known in advance, but discovered incrementally.

3.4. Novel Forms of Input for Groups and Individuals

Participants: Caroline Appert, Anastasia Bezerianos, Olivier Chapuis, Emmanuel Pietriga, André Spritzer, Rafael Morales Gonzalez, Bruno Fruchard.

Analyzing and manipulating large datasets can involve multiple users working together in a coordinated manner in multi-display environments: workstations, handheld devices, wall-sized displays [33]. Those users work towards a common goal, navigating and manipulating data displayed on various hardware surfaces in a coordinated manner. Group awareness [46], [27] is central in these situations, as users, who may or may not be co-located in the same room, can have an optimal individual behavior only if they have a clear picture of what their collaborators have done and are currently doing in the global context. We work on the design and implementation of interactive systems that improve group awareness in co-located situations [56], making individual users able to figure out what other users are doing without breaking the flow of their own actions.

In addition, users need a rich interaction vocabulary to handle large, structured datasets in a flexible and powerful way, regardless of the context of work. Input devices such as mice and trackpads provide a limited number of input actions, thus requiring users to switch between modes to perform different types of data manipulation and navigation actions. The action semantics of these input devices are also often too much dependent on the display output. For instance, a mouse movement and click can only be interpreted according to the graphical controller (widget) above which it is moved. We focus on designing powerful input techniques based upon technologies such as tactile surfaces (supported by UI toolkits developed in-house), 3D motion tracking systems, or custom-built controllers [58] *to complement (rather than replace) traditional input devices* such as keyboards, that remain the best method so far for text entry, and indirect input devices such as mice or trackpads for pixel-precise pointing actions.

The input vocabularies we investigate enable users to navigate and manipulate large and structured datasets in environments that involve multiple users and displays that vary in their size, position and orientation [33], [45], each having their own characteristics and affordances: wall displays [5], [81], workstations, tabletops [62], [41], tablets [6], [78], smartphones [10], [40], [74], [75], and combinations thereof [2], [9], [60], [33].

We aim at designing rich interaction vocabularies that go far beyond what current touch interfaces offer, which rarely exceeds five gestures such as simple slides and pinches. Designing larger gesture vocabularies requires identifying discriminating dimensions (e.g., the presence or absence of anchor points and the distinction between internal and external frames of reference [6]) in order to structure a space of gestures that interface designers can use as a dictionary for choosing a coherent set of controls. These dimensions should be few and simple, so as to provide users with gestures that are easy to memorize and execute. Beyond gesture complexity, the scalability of vocabularies also depends on our ability to design robust gesture recognizers that will allow users to fluidly chain simple gestures that make it possible to interlace navigation and manipulation actions.

We also plan to study how to further extend input vocabularies by combining touch [10], [6], [62] and mid-air gestures [5] with physical objects [51], [72], [58] and classical input devices such as keyboards to enable users to input commands to the system or to involve other users in their workflow (request for help, delegation, communication of personal findings, etc.) [35], [57]. Gestures and objects encode a lot of information in their shape, dynamics and direction, that can be directly interpreted in relation with the user, independently from the display output. Physical objects can also greatly improve coordination among actors for, e.g., handling priorities or assigning specific roles.

4. Application Domains

4.1. Mission-critical systems

Mission-critical contexts of use include emergency response & management, and critical infrastructure operations, such as public transportation systems, communications and power distribution networks, or the operations of large scientific instruments such as particle accelerators and astronomical observatories. Central to these contexts of work is the notion of situation awareness [27], i.e., how workers perceive and understand elements of the environment with respect to time and space, such as maps and geolocated data feeds from the field, and how they form mental models that help them predict future states of those elements. One of the main challenges is how to best assist subject-matter experts in constructing correct mental models and making informed decisions, often under time pressure. This can be achieved by providing them with, or helping them efficiently identify and correlate, relevant and timely information extracted from large amounts of raw data, taking into account the often cooperative nature of their work and the need for task coordination. With this application area, our goal is to investigate novel ways of interacting with computing systems that improve collaborative data analysis capabilities and decision support assistance in a mission-critical, often time-constrained, work context.

Relevant publications by team members this year: [22], [24].

4.2. Exploratory analysis of scientific data

Many scientific disciplines are increasingly data-driven, including astronomy, molecular biology, particle physics, or neuroanatomy. While making the right decision under time pressure is often less of critical issue when analyzing scientific data, at least not on the same temporal scale as truly time-critical systems, scientists are still faced with large-to-huge amounts of data. No matter their origin (experiments, remote observations, large-scale simulations), these data are difficult to understand and analyze in depth because of their sheer size and complexity. Challenges include how to help scientists freely-yet-efficiently explore their data, keep a trace of the multiple data processing paths they considered to verify their hypotheses and make it easy to backtrack, and how to relate observations made on different parts of the data and insights gained at different moments during the exploration process. With this application area, our goal is to investigate how data-centric interactive systems can improve collaborative scientific data exploration, where users' goals are more open-ended, and where roles, collaboration and coordination patterns [46] differ from those observed in mission-critical contexts of work.

Relevant publications by team members this year: [7].

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- ACM CHI 2016 Honorable mention for **TouchTokens: Guiding Touch Patterns with Passive Tokens** [4], awarded to the top 5% of all 2325 paper submissions.
- IEEE InfoVis 2016 Honorable mention for **The Attraction Effect in Information Visualization** [13].

6. New Software and Platforms

6.1. Smarties

FUNCTIONAL DESCRIPTION

The Smarties system provides an easy way to add mobile interactive support to collaborative applications for wall displays.

It consists of (i) a mobile interface that runs on mobile devices for input, (ii) a communication protocol between the mobiles and the wall application, and (iii) libraries that implement the protocol and handle synchronization, locking and input conflicts. The library presents the input as an event loop with callback functions and handles all communication between mobiles and wall application. Developers can customize the mobile interface from the wall application without modifying the mobile interface code.

On each mobile we find a set of cursor controllers associated with keyboards, widgets and clipboards. These controllers (pucks) can be shared by multiple collaborating users. They can control simple cursors on the wall application, or specific content (objects or groups of them). The developer can decide the types of widgets associated to pucks from the wall application side.

- Contact: Olivier Chapuis
- URL: http://smarties.lri.fr/

Smarties was used in the projects that led to the following publications this year: [7], [8], [22].

6.2. ZVTM

Zoomable Visual Transformation Machine
KEYWORDS: Information visualization - Data visualization - Visualization - Big data - Graph visualization
FUNCTIONAL DESCRIPTION

ZVTM is a toolkit enabling the implementation of multi-scale interfaces for interactively navigating in large datasets displayed as 2D graphics.

ZVTM is used for browsing large databases in multiple domains: geographical information systems, control rooms of complex facilities, astronomy, power distribution systems.

The toolkit also enables the development of applications running on ultra-high-resolution wall-sized displays.

- Participants: Caroline Appert, Maria Jesus Lobo Gunther, Arnaud Prouzeau, Hande Ozaygen, Can Liu and Olivier Chapuis
- Contact: Emmanuel Pietriga
- URL: http://zvtm.sf.net

Smarties was used in the projects that led to the following publications this year: [7], [8], [22], [19], [21].
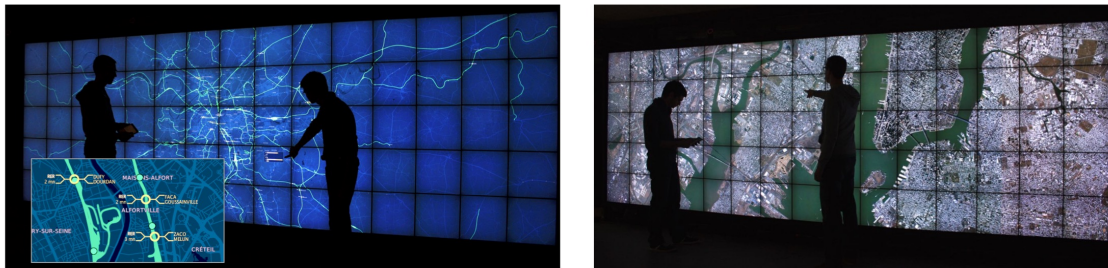
Figure 2. Geovisualization applications running on the WILDER platform. Real-time monitoring of railroad traffic in France (left), large-scale high-resolution orthoimagery visualization (right).

6.3. Platforms

6.3.1. Platform: WILDER

Ultra-high-resolution wall-sized displays [33] feature a very high pixel density over a large physical surface. Such platforms have properties that make them well-suited to the visualization of very large datasets. They can represent the data with a high level of detail while at the same time retaining context: users can transition from an overview of the data to a detailed view simply by physically moving in front of the wall display. Wall displays also offer good support for collaborative work, enabling multiple users to simultaneously visualize and interact with the displayed data. To make them interactive, wall-sized displays are increasingly coupled with input devices such as touch frames, motion-tracking systems and wireless multitouch devices, in order to enable multi-device and multi-user interaction with the displayed data. Application areas for such visualization platforms range from the monitoring of complex infrastructures and crisis management situations to tools for the exploratory visualization of scientific data.

WILDER is the latest ultra-high-resolution wall-sized display set up at Inria Saclay, and is one of the nodes of the Digiscope EquipEx. We use this platform for multiple projects, both fundamental HCI research, and research and development activities for specific application areas such as geographical informations systems (Figure 2) and astronomy (see Figure 3).

WILDER was used in the projects that led to the following publications this year: [7], [8], [22], [19], [23].

6.3.2. Platform: ANDES

ANDES is a platform similar to WILDER, set up at Inria in Santiago de Chile, that we use both as a research platform and as a showroom of our research and development activities. ANDES is the main platform used for our collaborative research project with the Millenium Institute of Astrophysics on the visualization of large FITS images (see Figure 3).

ANDES was used in the projects that led to the following publications this year: [7].

7. New Results

7.1. Wall Displays

Ultra-high-resolution wall displays feature a very high pixel density over a large physical surface, which makes them well-suited to the collaborative, exploratory visualization of large datasets (see Sections 6.3.1 and 6.3.2). We have continued working on the design, implementation and evaluation of interactive visualization
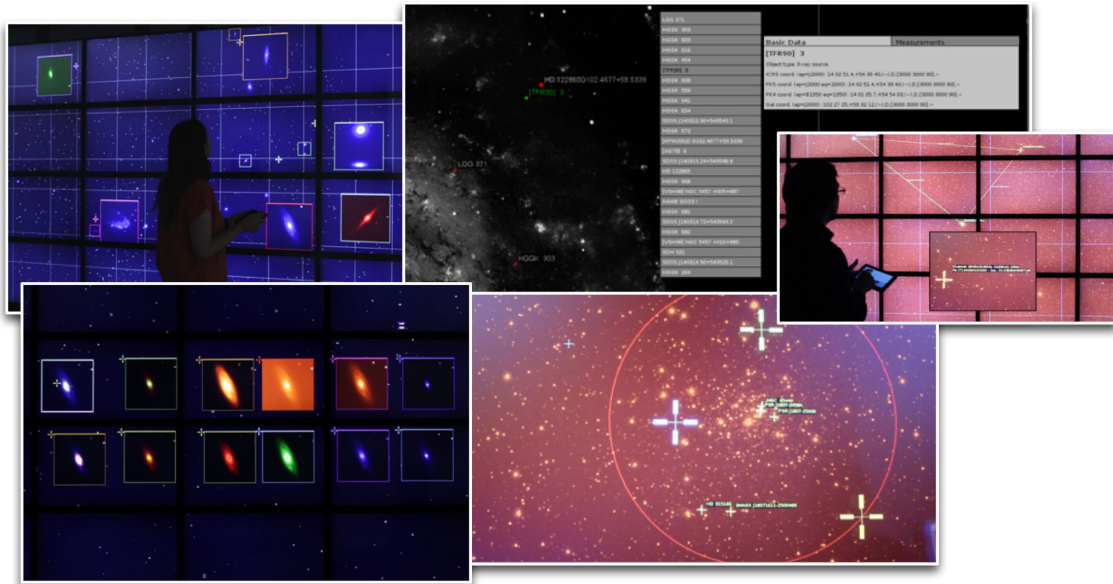
Figure 3. Visualization of high-dynamic-range FITS images and associated data catalogs in the domain of Astronomy on ANDES (collaboration with Inria Chile, Millenium Institute of Astrophysics, and Institut d'Astrophysique Spatiale).
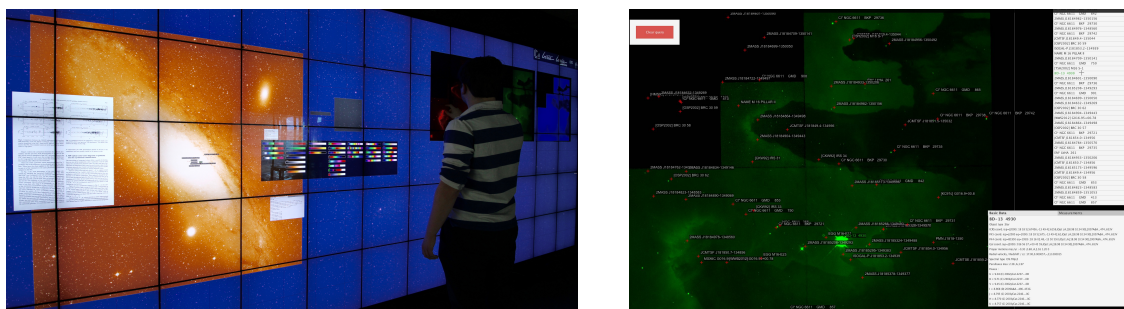
*Figure 4. **Left:** FITS-OW running on the WILDER platform, showing: multiple FITS images, (a) M31 on the left side, (b) three juxtaposed images that show observations of the Eagle nebula at different wavelengths, and (c) a much larger FITS image (86,499 × 13,474 pixels) used as a zoomable background over the entire wall; (d) the result-set of a SIMBAD query restricted to observations about galaxies; (e) basic measurements for galaxy M31; (e) a page of a research paper (PDF) discussing that particular galaxy; (f) the color map selector. **Right:** Results of a SIMBAD query superimposed on the corresponding FITS image, along with a sorted list of all items in the result-set. Selecting an element in this list updates the detailed info in the lower right window and highlights the source in the image. All windows can be freely repositioned on the wall.*

techniques for such ultra-high-resolution wall-sized displays, focusing, in some of these projects, on the collaboration between users who perform different data manipulation and analysis tasks.

- We continued working on FITS-OW, an application designed for such wall displays, that enables astronomers to navigate in large collections of FITS images, query astronomical databases, and display detailed, complementary data and documents about multiple sources simultaneously. We published a paper about FITS-OW [7], in which we describe the system, reporting on the technical challenges we addressed in terms of distributed graphics rendering and data sharing over the computer clusters that drive wall displays. The article also describes how astronomers interact with their data using both the wall's touch-sensitive surface and handheld devices. This work was also featured as a short article in the SPIE Newsroom (see Section 10.3).

- Wall-sized displays support small groups of users working together on large amounts of data. Observational studies of such settings have shown that users adopt a range of collaboration styles, from loosely to closely coupled. Shared interaction techniques, in which multiple users perform a command collaboratively, have also been introduced to support co-located collaborative work. In [19], we operationalized five collaborative situations with increasing levels of coupling, and tested the effects of providing shared interaction support for a data manipulation task in each situation. The results show the benefits of shared interaction for close collaboration: it encourages collaborative manipulation, it is more efficient and preferred by users, and it reduces physical navigation and fatigue. We also identify the time costs caused by disruption and communication in loose collaboration and analyze the trade-offs between parallelization and close collaboration. These findings inform the design of shared interaction techniques to support collaboration on wall-sized displays.

- We also studied how pairs explore graphs on a touch enabled wall-display [16], using two selection techniques adapted for collaboration: a basic localized selection, and a propagation selection technique that uses the idea of diffusion/transmission from an origin node. We assessed in a controlled experiment the impact of selection technique on a shortest path identification task. Pairs consistently divided space even though the task is not spatially divisible. The basic selection technique, that has a localized visual effect, led to parallel work that negatively impacted accuracy. The large visual footprint of the propagation technique led to close coordination, improving speed and accuracy for complex graphs only. We then observed the use of propagation on additional graph topology tasks, confirming pair strategies on spatial division and coordination.

- In [22], we focused on road traffic control center. Road traffic control centers are of vital importance to modern cities. Interviews with controllers in two such centers identified the need to incorporate the visualization of results from predictive traffic models with real traffic, to help operators choose among different interventions on the network. We explore this idea in a prototype that runs on a wall display, and supports direct touch and input from workstations and mobile devices. Apart from basic functionality to manage the current traffic such as changing traffic light duration or speed limits, the prototype incorporates traffic simulations for forecasting results of possible actions, highlighting their differences to current traffic. Based on needs identified in our interviews, we offered two techniques that visually combine simulated and real situations, taking advantage of the large display space: multiple independent views and DragMagic, a variation of magic lenses. A preliminary laboratory experiment suggests that both techniques are viable design options, even for monitoring several simulations and areas of interest, contrary to expectations from previous work. However DragMagics are easier to master.

- Immersion is the subjective impression of being deeply involved in a specific situation, and can be sensory or cognitive. In a position paper [23], we used a basic model of visual perception to study how ultra-high resolution wall displays can provide visual immersion. With their large size, depending on the position of viewers in front of them, wall displays can provide a surrounding and vivid environment. Users close to the wall can have their visual field filled by the wall and they are able to see clearly a large amount information with a fine resolution. However, when close to the wall, visual distortion due to large possible viewing angles, can affect the viewing of data. On the

contrary, from far away, distortion is no longer an issue, but the viewers' visual field is not fully contained inside the wall, and the information details seen are less fine.
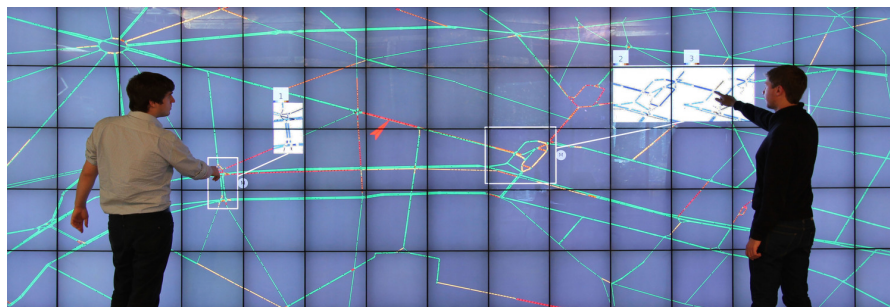
Figure 5. **Top:** *A pair using the propagation technique described in [16] to explore a graph. They discuss two communities, in orange and purple, selected using the propagation technique. The communities are linked by a specific node shown by the right user. The remaining 3 orange-purple nodes show how by propagating the purple community, it flows into the orange one through this node.* **Bottom:** *Visualization from [22] of traffic in a city with two "DragMagics" (white rectangles) showing one (left) and two (right) simulations associated with different possible interventions on the traffic. The simulation visualizations use difference color maps to highlight differences with the real traffic.*

7.2. Gestures, Tangibles and Sound

- We designed a new way of implementing tangible interfaces with TouchTokens [4]. The approach requires only passive tokens and a regular multi-touch surface. The tokens constrain users' grasp, and thus, the relative spatial configuration of fingers on the surface, theoretically making it possible to design algorithms that can recognize the resulting touch patterns. We performed a formative user study to collect and analyze touch patterns with tokens of varying shape and size. The analysis of this pattern collection showed that individual users have a consistent grasp for each token, but that this grasp is user-dependent and that different grasp strategies can lead to confounding patterns. We thus designed a second set of tokens featuring notches that constrain users' grasp. Our recognition algorithm can classify the resulting patterns with a high level of accuracy ($>95\%$) without any training, enabling application designers to associate rich touch input vocabularies with command triggers and parameter controls.
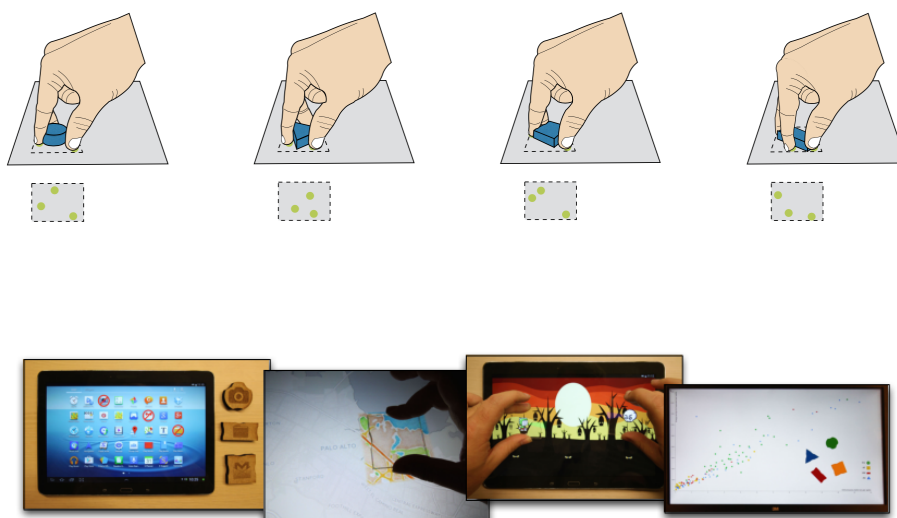
*Figure 6. **Top:** TouchTokens are passive tokens that guide users' fingers to specific spatial configurations, resulting in distinguishable touch patterns. **Bottom:** Proof-of-concept applications: access control, tangible magic lenses, character controllers in a game, data visualization.*

- In collaboration with IRCAM, we introduced SoundGuides [17], a user adaptable tool for auditory feedback on movement. The system is based on a interactive machine learning approach, where both gestures and sounds are first conjointly designed and conjointly learned by the system. The system can then automatically adapt the auditory feedback to any new user, taking into account the particular way each user performs a given gesture. SoundGuides is suitable for the design of continuous auditory feedback aimed at guiding users' movements and helping them to perform a specific movement consistently over time. Applications span from movement-based interaction techniques to auditory-guided rehabilitation. We first describe our system and report a study that demonstrates a "stabilizing effect" of our adaptive auditory feedback method.

7.3. Interacting with Linked Data

As part of the team's novel research theme on Semantics-Driven Data Manipulation 3.2, and in collaboration with Aba-Sah Dadzie from the Open University, Emmanuel Pietriga coordinated a special issue of the Semantic Web Journal and wrote a follow-up [12] to the 2011 survey about Approaches to Visualizing Linked Data [42]. Linked Data promises to serve as a disruptor of traditional approaches to data management and use, promoting the push from the traditional Web of documents to a Web of data. The ability for data consumers to adopt a follow your nose approach, traversing links defined within a dataset or across independently-curated datasets, is an essential feature of this new Web of Data, enabling richer knowledge retrieval thanks to synthesis across multiple sources of, and views on, interrelated datasets. But for the Web of Data to be successful, we must design novel ways of interacting with the corresponding very large amounts of complex, interlinked, multi-dimensional data throughout its management cycle. The design of user interfaces for Linked Data, and more specifically interfaces that represent the data visually, play a central role in this respect. Contributions to this special issue on Linked Data visualization investigate different approaches to harnessing visualization as a tool for exploratory discovery and basic-to-advanced analysis. The papers in this volume illustrate the design and construction of intuitive means for end-users to obtain new insight and gather more knowledge, as they follow links defined across datasets over the Web of Data.

7.4. Visualization

- The attraction effect is a well-studied cognitive bias in decision making research, where one's choice between two alternatives is influenced by the presence of an irrelevant (dominated) third alternative. In collaboration with EPI Aviz, we examined in [13] whether this cognitive bias, so far only tested with three alternatives and simple presentation formats such as numerical tables, text and pictures, also appears in visualizations. In a series of crowdsource experiments, we observed this cognitive bias in visualizations (namely scatterplots), even in larger sets of alternatives, never considered before, where the number of alternatives is too large for numerical tables to be practical. We discussed implications for future research on how to further study and possibly alleviate the attraction effect.

- With colleagues from the University of Konstanz [14] we concluded previous work on data glyphs, i.e., visual marks that encode multiple dimensions to one or more visual variables. We provided a systematic review of experimental studies on data glyphs from the past 60 years, describing the types of glyphs and design variations tested, the tasks under which they were analyzed, and study results. Based on our meta analysis of all results, we further contributed a set of design implications and a discussion on open research directions.

- In [11], with colleagues from INRA, we provided an overview of a framework for Evolutionary Visual Exploration (EVE) that guides users in exploring large search spaces. EVE uses an interactive evolutionary algorithm to steer the exploration of multidimensional datasets towards two-dimensional projections that are interesting to the analyst. Our method smoothly combines automatically calculated metrics and user input in order to propose pertinent views to the user. While previously we showed that using EVE, domain experts were able to formulate interesting hypothesis and reach new insights when exploring freely, our new findings indicate that users, guided by the interactive evolutionary algorithm, are able to converge quickly to an interesting view of their data when a clear task is specified. Our work aims at building a bridge between the domains of visual analytics and interactive evolution.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Tecknowmetrix (TKM): ANRT/CIFRE PhD (Hugo Romat), 3 years, starting June 2016.

9. Partnerships and Cooperations

9.1. Regional Initiatives

ScaleGest. Surface Gestures for Advanced Graphical Interfaces: Which Gesture for What. (2014-2017) Funded by Digiteo. In collaboration with Telecom ParisTech: **109Keuros**. Participants: Caroline Appert (PI), Rafael Morales Gonzalez, Emmanuel Pietriga.

The project aims at designing gesture-based interaction for expert users who navigate and manipulate large datasets. In the context of advanced graphical applications, the number of gestures should be large-enough to cover the set of controls (*i.e.*, commands and parameter settings) but remain simple-enough to avoid exceeding human abilities. Making gesture-based interaction scale with graphical applications' growing complexity can be achieved only by understanding the foundational aspects of this input modality. This project is about characterizing and structuring both the space of application controls and the space of surface gestures in order to establish guidelines for appropriate control-gesture mappings. It is also about the definition of a sound and systematic evaluation methodology that will serve as a reference benchmark for evaluating these mappings. The resulting control-gesture mappings are demonstrated in the specific application domains of cartography and astronomy.

9.2. National Initiatives

9.2.1. ANR

MapMuxing - Multi-dimensional Map Multiplexing. (2014-2018) Funded by the French National Research Agency (ANR). In collaboration with IGN (Institut National de l'Information Géographique et Forestière): **208Keuros**/499Keuros. Participants: Emmanuel Pietriga (PI), Caroline Appert, Olivier Chapuis, María-Jesús Lobo. http://mapmuxing.ign.fr

The project explores novel ways of combining different maps and data layers into a single cartographic representation, and investigates novel interaction techniques for navigating in it. The project aims at going beyond the traditional pan & zoom and overview+detail interface schemes, and at designing and evaluating novel cartographic visualizations that rely on high-quality generalization, *i.e.*, the simplification of geographic data to make it legible at a given map scale, and symbol specification.

9.2.2. *Collaborations with other French Research Organizations*

CorTextViz. (2015-2016) Funded by INRA (Institut National de la Recherche Agronomique). In collaboration with project-team Aviz at Inria Saclay (Jean-Daniel Fekete) and INRA (Jean-Philippe Cointet, Guy Riba). Interactive visualization of medium-scale multi-level networks, supporting data storytelling on wall displays. Participants: André Spritzer, Emmanuel Pietriga (PI), Anastasia Bezerianos.

9.3. European Initiatives

9.3.1. *Collaborations with Major European Organizations*

- European Southern Obseervatory (ESO)
- ALMA Operations Monitoring and Control - design and implementation of state-of-the-art inter-active visualization components for the operations monitoring and control software that runs the ALMA radio-observatory in Chile.
- Deutsches Elektronen-Synchrotron (DESY)
- Scientific collaboration on the design and implementation of user interfaces for array operations monitoring and control for the Cherenkov Telescope Array (CTA) project [24], to be built in the Canary Islands (Spain) and in the Atacama desert (Chile).

9.4. International Initiatives

9.4.1. *Inria International Labs*

Inria Chile / CIRIC. From 2012 to 2015, Emmanuel Pietriga was the scientific leader of the Massive Data team at Inria Chile, working on projects in collaboration with the ALMA radio-telescope and the Millenium Institute of Astrophysics [15]. He is now scientific advisor to Inria Chile's visualization lab.

9.4.2. *Inria International Partners*

9.4.2.1. *Informal International Partners*

- University of Konstanz: Daniel Keim and Johannes Fuchs on mapping out the design space for visualization glyphs [14]. Participants: Anastasia Bezerianos.

9.5. International Research Visitors

9.5.1. *Visits of International Scientists*

- Shumin Zhai, Google, June 2016
- Iftach Sadeh, DESY/CTA Observatory, April 2016

9.5.1.1. *Internships*

- María Grazia Prato, Inria Chile, April 2016

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific events organisation

10.1.1.1. General chair, scientific chair

- CHI 2017, 35th ACM SIGCHI Conference on Human Factors in Computing Systems: Caroline Appert (papers and notes co-chair)

10.1.2. Scientific events selection

10.1.2.1. Chair of conference program committees

- CHI 2017, 35th ACM SIGCHI Conference on Human Factors in Computing Systems: Emmanuel Pietriga (SC - subcommittee chair)

10.1.2.2. Member of the conference program committees

- CHI 2017, 35th ACM SIGCHI Conference on Human Factors in Computing Systems: Anastasia Bezerianos (AC - associate chairs)
- CHI 2016, 34th ACM SIGCHI Conference on Human Factors in Computing Systems: Anastasia Bezerianos, Emmanuel Pietriga (AC - associate chairs)
- UIST 2016, 29th ACM ACM User Interface Software and Technologies Symposium: Caroline Appert, Anastasia Bezerianos (AC - associate chairs)
- InfoVis 2016, IEEE Information Visualization Conference: Anastasia Bezerianos (AC - associate chair)
- EICS 2016, 8th ACM SIGCHI Symposium on Engineering Interactive Computing Systems: Emmanuel Pietriga
- VOILA @ ISWC 2016, Visualizations and User Interfaces for Ontologies and Linked Data, workshop co-located with ISWC 2016: Emmanuel Pietriga
- SSDBM 2016: 28th International Conference on Scientific and Statistical Database Management: Emmanuel Pietriga
- Immersive Analytics @ ISS 2016: Exploring Future Interaction and Visualization Technologies for Data Analytics, workshop co-located with ACM ISS 2016: Emmanuel Pietriga
- IHM 2016 – Conference of the Association Francophone d'Interaction Homme-Machine: Olivier Chapuis (AC - associate chair)

10.1.2.3. Reviewer

- ACM CHI 2016, Conference on Human Factors in Computing Systems: Olivier Chapuis, María-Jesús Lobo, Arnaud Prouzeau
- ACM UIST 2016, Symposium on User Interface Software and Technology: Olivier Chapuis, Emmanuel Pietriga
- EuroVis EG/VGTC 2016, Conference on Data Visualization: Caroline Appert
- ACM ISS 2016, International Conference on Interactive Surfaces and Spaces: Caroline Appert, Olivier Chapuis
- ACM MobileHCI 2016, International Conference on Human-Computer Interaction with Mobile Devices and Services: Caroline Appert
- ACM IUI 2017, International Conference on Intelligent User Interfaces: Caroline Appert

- IEEE VR 2016, Virtual Reality: Olivier Chapuis
- ACM EICS 2016, Symposium on Engineering Interactive Computing Systems: Olivier Chapuis

10.1.3. Journal

10.1.3.1. Member of the editorial boards

- Semantic Web Journal: Emmanuel Pietriga (guest editor, special issue on Visual Exploration and Analysis of Linked Data)

10.1.3.2. Reviewer - Reviewing activities

- ACM ToCHI, Transactions on Computer-Human Interaction: Olivier Chapuis
- IJHCI, International Journal of Human-Computer Interaction: Olivier Chapuis, Emmanuel Pietriga
- IEEE TVCG, Transactions on Visualization and Computer Graphics: Anastasia Bezerianos

10.1.4. Invited talks

- Olivier Chapuis, L'équipement d'excellence Digiscope: un réseau de plateformes pour la visualisation interactive, CNRS Innovatives Big Data Forum, October 2016, Paris France.
- Anna Gogolou, Visualization and human-computer interaction methods: Challenges in Interactive Visual Exploration of Large Data Series, 2nd Interdisciplinary Time Series Workshop, December 2016, Paris France.
- Anastasia Bezerianos, Hands-on session on visualization and human-computer interaction methods: Designing interactive visualization techniques for large data series, 2nd Interdisciplinary Time Series Workshop, December 2016, Paris France.

10.1.5. Scientific expertise

- Evaluator for the Sponsored Program for Industrial Data Science (EDF & Thales) - Fondation Mathématique Jacques Hadamard/IRSDI: Emmanuel Pietriga

10.1.6. Research administration

- Executive committee of the Laboratoire de Recherche en Informatique (LRI): Olivier Chapuis
- President of Inria Saclay - Île de France's Commission for Technological Development (CDT): Emmanuel Pietriga

10.1.7. Learned societies

- Association Francophone d'Interaction Homme-Machine (AFIHM): Olivier Chapuis (board member, vice-president).

10.1.8. Hiring committees

- Univ. Paris-Sud hiring committee, Commission Consultative des Spécialistes de l'Université 27ème section (computer science), members: Caroline Appert.
- Univ. Paris-Sud hiring committee, Comités de Sélection PRCE/PRAG 1195 (english language), members: Anastasia Bezerianos.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master: Anastasia Bezerianos, Co-Head of EIT masters M1 & M2 HCID, Univ. Paris-Sud

Master: Anastasia Bezerianos, Co-Head of M2 Interaction, Univ. Paris-Sud, Université Paris-Saclay

Master: Anastasia Bezerianos, Programming of Interactive Systems – 21h CM, M1/M2 HCID + M2R Interaction, Univ. Paris-Sud

Master: Anastasia Bezerianos, Mixed Reality and Tangible Interaction – 11h CM, M1/M2 HCID + M2R Interaction, Univ. Paris-Sud

Master: Anastasia Bezerianos, Career Seminar, – 12h CM, M2 HCID + M2R Interaction, Univ. Paris-Sud

Master: Anastasia Bezerianos, Design Project in HCI, – 21h CM, M1 HCID, Univ. Paris-Sud

Master: Anastasia Bezerianos, HCI Project, – 21h CM, M2 HCID + M2 Interaction, Univ. Paris-Sud

Master: Anastasia Bezerianos, Introduction to Java – 18h CM, M1/M2 HCID + M2R Interaction, Univ. Paris-Sud

Master: Caroline Appert, Evaluation of Interactive Systems – Intro, 21h CM, M1/M2 HCID + M2R Interaction, Univ. Paris-Sud

Master: Caroline Appert, Evaluation of Interactive Systems – Advanced, 21h CM, M1/M2 HCID + M2R Interaction, Univ. Paris-Sud

Master: Emmanuel Pietriga, Data Visualization, 24h CM, M2 Informatique Décisionelle, Univ. Paris-Dauphine.

Licence: María-Jesús Lobo, Informatique Graphique, 28h, L3 Univ. Paris-Sud

Licence: María-Jesús Lobo, Programmation d'Interfaces Interactives avancées, 29h, L3 Univ. Paris-Sud

Master: Arnaud Prouzeau, Programming of Interactive Systems: 27h, Univ. Paris-Sud.

Licence: Arnaud Prouzeau, Programmation Impérative en C++, 22.4h, Univ. Paris-Sud.

Licence: Arnaud Prouzeau, Programmation Impérative Avancée en C++, 16.5h, Univ. Paris-Sud.

Licence: Arnaud Prouzeau, Interaction Homme-Machine, 12h, olytech Paris-Sud.

Licence: Bruno Fruchard, Développement d'Applications Mobiles, 13h, Télécom Paristech.

Licence: Bruno Fruchard, Paradigmes de programmation, 24.5h, Télécom Paristech.

10.2.2. Supervision

PhD in progress : Anna Gogolou, Iterative and expressive querying for big data series, October 2016, Advisors: Anastasia Bezerianos, Themis Palpanas

PhD in progress : María-Jesús Lobo, Interaction Techniques for Map Multiplexing, since October 2014, Advisors: Caroline Appert, Emmanuel Pietriga

PhD in progress : Rafael Morales Gonzalez, Surface Gestures for Advanced Graphical Interfaces: Which Gesture for What, since November 2014, Advisors: Caroline Appert, Gilles Bailly, Emmanuel Pietriga

PhD in progress : Arnaud Prouzeau, Collaboration around Wall-Displays in Time Critical and Command and Control Contexts, since October 2014, Advisors: Anastasia Bezerianos, Olivier Chapuis

PhD in progress : Hugo Romat, Nouvelles techniques de visualisation et de manipulation interactive de collections de documents et données hétérogènes, since June 2016, Advisors: Caroline Appert, Emmanuel Pietriga

PhD in progress : Evanthia Dimara, Merging Interactive Visualization and Automated Analysis for Group Decision-Making Involving Large Datasets, since October 2014, Advisors: Pierre Dragicevic, Anastasia Bezerianos

PhD in progress : Bruno Fruchard, Techniques d'interaction exploitant la mémoire spatiale pour faciliter l'accès rapide aux commandes et aux données, since October 2015, Advisors: Eric Lecolinet, Olivier Chapuis

Master: Dylan Lebout, Indirect vs Direct surface gestures for wall-sized displays, May - August 2016, Advisors: Caroline Appert, Olivier Chapuis

10.2.3. Juries

PhD: Rémy Dautriche, Université Grenoble-Alpes: Emmanuel Pietriga (rapporteur)

PhD: Pascal Goffin, Inria / Université Paris-Saclay: Emmanuel Pietriga (président du jury)

10.3. Popularization

- SPIE Newsroom: E. Pietriga, F. del Campo, A. Ibsen, R. Primet, C. Appert, O. Chapuis, M. Hempel, R. Muñoz, S. Eyheramendy Duerr, A. Jordan, H. Dole, Ultra-high-resolution walls for visualizing very large datasets, May 2016. http://spie.org/newsroom/6505-ultra-high-resolution-walls-for-visualizing-very-large-datasets

11. Bibliography

Major publications by the team in recent years

[1] C. APPERT, O. CHAPUIS, E. PIETRIGA, M.-J. LOBO.*Reciprocal Drag-and-Drop*, in "ACM Transactions on Computer-Human Interaction", September 2015, vol. 22, n° 6, p. 29:1–29:36 [*DOI : 10.1145/2785670*], https://hal.archives-ouvertes.fr/hal-01185805.

[2] O. CHAPUIS, A. BEZERIANOS, S. FRANTZESKAKIS.*Smarties: An Input System for Wall Display Development*, in "CHI '14", Toronto, Canada, ACM, April 2014, p. 2763-2772 [*DOI : 10.1145/2556288.2556956*], https://hal.archives-ouvertes.fr/hal-00979034.

[3] M.-J. LOBO, E. PIETRIGA, C. APPERT.*An Evaluation of Interactive Map Comparison Techniques*, in "CHI '15 Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems", Seoul, South Korea, ACM, April 2015, p. 3573-3582 [*DOI : 10.1145/2702123.2702130*], https://hal.inria.fr/hal-01144163.

[4] R. MORALES GONZALEZ, C. APPERT, G. BAILLY, E. PIETRIGA.*TouchTokens: Guiding Touch Patterns with Passive Tokens*, in "2016 CHI Conference on Human Factors in Computing Systems", San Jose, CA, United States, May 2016 [*DOI : 10.1145/2858036.2858041*], https://hal.archives-ouvertes.fr/hal-01315130.

[5] M. NANCEL, E. PIETRIGA, O. CHAPUIS, M. BEAUDOUIN-LAFON.*Mid-air Pointing on Ultra-Walls*, in "ACM Transactions on Computer-Human Interaction", August 2015, vol. 22, n° 5, p. 21:1–21:62 [*DOI : 10.1145/2766448*], https://hal.inria.fr/hal-01184544.

[6] H. OLAFSDOTTIR, C. APPERT.*Multi-Touch Gestures for Discrete and Continuous Control*, in "International Working Conference on Advanced Visual Interfaces (AVI)", Como, Italy, May 2014, 8 [*DOI : 10.1145/2598153.2598169*], https://hal.archives-ouvertes.fr/hal-00998971.

[7] E. PIETRIGA, F. DEL CAMPO, A. IBSEN, R. PRIMET, C. APPERT, O. CHAPUIS, M. HEMPEL, R. MUÑOZ, S. EYHERAMENDY, A. JORDAN, H. DOLE.*Exploratory Visualization of Astronomical Data on Ultra-high-resolution Wall Displays*, in "Proceedings SPIE", July 2016, vol. 9913, 15 [*DOI : 10.1117/12.2231191*], https://hal.inria.fr/hal-01350722.

[8] A. PROUZEAU, A. BEZERIANOS, O. CHAPUIS.*Evaluating Multi-User Selection for Exploring Graph Topology on Wall-Displays*, in "IEEE Transactions on Visualization and Computer Graphics", 2016, 14 [*DOI : 10.1109/TVCG.2016.2592906*], https://hal.archives-ouvertes.fr/hal-01348578.

[9] T. TSANDILAS, A. BEZERIANOS, T. JACOB.*SketchSliders: Sketching Widgets for Visual Exploration on Wall Displays*, in "Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems", Seoul, South Korea, ACM, April 2015, p. 3255-3264 [*DOI : 10.1145/2702123.2702129*], https://hal.archives-ouvertes.fr/hal-01144312.

[10] S. ZHAI, P. O. KRISTENSSON, C. APPERT, T. H. ANDERSEN, X. CAO.*Foundational Issues in Touch-Screen Stroke Gesture Design - An Integrative Review*, in "Foundations and Trends in Human-Computer Interaction", December 2012, vol. 5, n° 2, p. 97-205 [*DOI : 10.1561/1100000012*], https://hal.inria.fr/hal-00765046.

Publications of the year

Articles in International Peer-Reviewed Journal

[11] N. BOUKHELIFA, A. BEZERIANOS, W. CANCINO, E. LUTTON.*Evolutionary Visual Exploration: Evaluation of an IEC Framework for Guided Visual Search*, in "Evolutionary Computation", 2016, forthcoming [*DOI : 10.1162/EVCO_A_00161*], https://hal.inria.fr/hal-01218959.

[12] A.-S. DADZIE, E. PIETRIGA.*Visualisation of Linked Data - Reprise*, in "Open Journal Of Semantic Web", November 2016, vol. 8, n° 1, p. 1 - 21 [*DOI : 10.3233/SW-160249*], https://hal.inria.fr/hal-01406437.

[13] E. DIMARA, A. BEZERIANOS, P. DRAGICEVIC.*The Attraction Effect in Information Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", 2017, vol. 23, n° 1 [*DOI : 10.1109/TVCG.2016.2598594*], https://hal.inria.fr/hal-01355750.

[14] J. FUCHS, P. ISENBERG, A. BEZERIANOS, D. KEIM.*A Systematic Review of Experimental Studies on Data Glyphs*, in "IEEE Transactions on Visualization and Computer Graphics", 2017, forthcoming [*DOI : 10.1109/TVCG.2016.2549018*], https://hal.inria.fr/hal-01378429.

[15] E. PIETRIGA, F. DEL CAMPO, A. IBSEN, R. PRIMET, C. APPERT, O. CHAPUIS, M. HEMPEL, R. MUÑOZ, S. EYHERAMENDY, A. JORDAN, H. DOLE.*Exploratory Visualization of Astronomical Data on Ultra-high-resolution Wall Displays*, in "Proceedings SPIE", July 2016, vol. 9913, 15 [*DOI : 10.1117/12.2231191*], https://hal.inria.fr/hal-01350722.

[16] A. PROUZEAU, A. BEZERIANOS, O. CHAPUIS.*Evaluating Multi-User Selection for Exploring Graph Topology on Wall-Displays*, in "IEEE Transactions on Visualization and Computer Graphics", 2016, 14 [*DOI : 10.1109/TVCG.2016.2592906*], https://hal.archives-ouvertes.fr/hal-01348578.

International Conferences with Proceedings

[17] J. FRANÇOISE, O. CHAPUIS, S. HANNETON, F. BEVILACQUA.*SoundGuides: Adapting Continuous Auditory Feedback to Users*, in "CHI EA '16", San Jose, United States, ACM (editor), Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, May 2016, p. 2829–2836 [*DOI : 10.1145/2851581.2892420*], https://hal.archives-ouvertes.fr/hal-01317069.

[18] B. FRUCHARD, E. LECOLINET, O. CHAPUIS.*MarkPad: Augmenting Touchpads for Command Selection*, in "Proceedings of the 35th international conference on Human factors in computing systems", Denver, United States, CHI '17, May 2017 [*DOI : 10.1145/3025453.3025486*], https://hal.archives-ouvertes.fr/hal-01437093.

[19] C. LIU, O. CHAPUIS, M. BEAUDOUIN-LAFON, E. LECOLINET.*Shared Interaction on a Wall-Sized Display in a Data Manipulation Task*, in "Proceedings of the 34th international conference on Human factors in computing systems", San Jose, United States, ACM (editor), CHI '16, SigCHI and ACM, May 2016, p. 1-12 [*DOI : 10.1145/2858036.2858039*], https://hal.inria.fr/hal-01275535.

[20] C. LIU, O. CHAPUIS, M. BEAUDOUIN-LAFON, E. LECOLINET.*CoReach: Cooperative Gestures for Data Manipulation on Wall-sized Displays*, in "Proceedings of the 35th international conference on Human factors in computing systems", Denver, United States, ACM (editor), CHI '17, May 2017 [*DOI : 10.1145/3025453.3025594*], https://hal.archives-ouvertes.fr/hal-01437091.

[21] R. MORALES GONZALEZ, C. APPERT, G. BAILLY, E. PIETRIGA.*TouchTokens: Guiding Touch Patterns with Passive Tokens*, in "2016 CHI Conference on Human Factors in Computing Systems", San Jose, CA, United States, May 2016 [*DOI : 10.1145/2858036.2858041*], https://hal.archives-ouvertes.fr/hal-01315130.

[22] A. PROUZEAU, A. BEZERIANOS, O. CHAPUIS.*Towards Road Traffic Management with Forecasting on Wall Displays*, in "Proceedings of the 2016 International Conference on Interactive Surfaces and Spaces", Niagara Falls, Canada, November 2016, p. 119–128 [*DOI : 10.1145/2992154.2992158*], https://hal.archives-ouvertes.fr/hal-01370231.

[23] A. PROUZEAU, A. BEZERIANOS, O. CHAPUIS.*Visual Immersion in the Context of Wall Displays*, in "Interactive Surfaces and Spaces Surfaces Companion Proceedings", Niagara Falls, Canada, November 2016 [*DOI : 10.1145/3009939.3009945*], https://hal.archives-ouvertes.fr/hal-01383587.

[24] I. SADEH, I. OYA, J. SCHWARZ, E. PIETRIGA.*Prototyping the graphical user interface for the operator of the Cherenkov Telescope Array*, in "SPIE Astronomical Telescopes and Instrumentation", Edinburgh, United Kingdom, Proc. SPIE, Software and Cyberinfrastructure for Astronomy III, SPIE, June 2016, vol. 9913, 9 [*DOI : 10.1117/12.2231606*], https://hal.inria.fr/hal-01350713.

Research Reports

[25] E. DIMARA, A. BEZERIANOS, P. DRAGICEVIC.*Testing the Attraction Effect on Two Information Visualization Datasets*, Inria Saclay Ile de France, March 2016, n⁰ RR-8895, 13, https://hal.inria.fr/hal-01295624.

[26] A. PROUZEAU.*Collaboration around Wall-Displays in Command and Control contexts*, LRI, Univ. Paris-Sud, CNRS, Inria, Université Paris-Saclay, December 2016, Consortium Doctoral de IHM 2016, https://hal.archives-ouvertes.fr/hal-01420098.

References in notes

[27] M. R. ENDSLEY, D. G. JONES (editors). *Designing for Situation Awareness: an Approach to User-Centered Design*, CRC Press, Taylor & Francis, 2012.

[28] S. ABITEBOUL, P. BUNEMAN, D. SUCIU.*Data on the Web: From Relations to Semistructured Data and XML*, Morgan Kaufmann, 1999.

[29] S. ABITEBOUL, P. SENELLART, V. VIANU.*The ERC Webdam on Foundations of Web Data Management*, in "Proc. of the 21st International Conference Companion on World Wide Web", WWW '12 Companion, ACM, 2012, p. 211–214, http://doi.acm.org/10.1145/2187980.2188011.

[30] J. ALVINA, C. APPERT, O. CHAPUIS, E. PIETRIGA.*RouteLens: Easy Route Following for Map Applications*, in "Proceedings of the International Working Conference on Advanced Visual Interfaces", AVI '14, ACM, 2014, p. 125–128 [*DOI : 10.1145/2598153.2598200*].

[31] C. APPERT, O. CHAPUIS, E. PIETRIGA.*High-precision magnification lenses*, in "Proceedings of the 28th international conference on Human factors in computing systems", CHI '10, ACM, 2010, p. 273–282 [*DOI : 10.1145/1753326.1753366*].

[32] C. APPERT, J.-D. FEKETE.*OrthoZoom Scroller: 1D Multi-scale Navigation*, in "Proceedings of the SIGCHI Conference on Human Factors in Computing Systems", New York, NY, USA, CHI '06, ACM, 2006, p. 21–30, http://doi.acm.org/10.1145/1124772.1124776.

[33] M. BEAUDOUIN-LAFON, S. HUOT, M. NANCEL, W. MACKAY, E. PIETRIGA, R. PRIMET, J. WAGNER, O. CHAPUIS, C. PILLIAS, J. R. EAGAN, T. GJERLUFSEN, C. KLOKMOSE.*Multisurface Interaction in the WILD Room*, in "IEEE Computer", 2012, vol. 45, n° 4, p. 48-56 [*DOI : 10.1109/MC.2012.110*].

[34] E. BENSON, D. KARGER.*End-users Publishing Structured Information on the Web: An Observational Study of What, Why, and How*, in "Proc. of the SIGCHI Conference on Human Factors in Computing Systems", CHI '14, ACM, 2014, p. 1265–1274, http://doi.acm.org/10.1145/2556288.2557036.

[35] J. BERNDTSSON, M. NORMARK.*The coordinative functions of flight strips: air traffic control work revisited*, in "Proceedings of the international ACM SIGGROUP conference on Supporting group work", GROUP '99, ACM, 1999, p. 101–110, http://doi.acm.org/10.1145/320297.320308.

[36] T. BERNERS-LEE, Y. CHEN, L. CHILTON, D. CONNOLLY, R. DHANARAJ, J. HOLLENBACH, A. LERER, D. SHEETS.*Tabulator: Exploring and Analyzing linked data on the Semantic Web*, in "Proc. of the International Semantic Web User Interaction Workshop", 2006.

[37] J. BERTIN.*Semiology of Graphics*, University of Wisconsin Press, 1983.

[38] A. BEZERIANOS, R. BALAKRISHNAN.*View and Space Management on Large Displays*, in "IEEE Computer Graphics and Applications", 2005, vol. 25, n° 4, p. 34-43.

[39] A. BEZERIANOS.*Using alternative views for layout, comparison and context switching tasks in wall displays*, in "OZCHI", 2007, p. 303-310.

[40] D. BONNET, C. APPERT, M. BEAUDOUIN-LAFON.*Extending the Vocabulary of Touch Events with ThumbRock*, in "Proceedings of Graphics Interface", GI '13, CIPS, 2013, p. 221-228, http://dl.acm.org/citation.cfm?id=2532129.2532166.

[41] A. COLLINS, A. BEZERIANOS, G. MCEWAN, M. RITTENBRUCH, R. WASINGER, J. KAY.*Understanding File Access Mechanisms for Embedded Ubicomp Collaboration Interfaces*, in "Proceedings of the 11th International Conference on Ubiquitous Computing", New York, NY, USA, UbiComp '09, ACM, 2009, p. 135–144, http://doi.acm.org/10.1145/1620545.1620567.

[42] A.-S. DADZIE, M. ROWE.*Approaches to Visualising Linked Data: A Survey*, in "Semantic Web", 2011, vol. 2, n° 2, http://dx.doi.org/10.3233/SW-2011-0037.

[43] A. Dix, G. Lepouras, A. Katifori.*From the web of data to a world of action*, in "Web Semantics", 2010, vol. 8, p. 394–408, http://dx.doi.org/10.1016/j.websem.2010.04.007.

[44] D. Fensel, J. Hendler, H. Lieberman, W. Wahlster.*Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*, MIT Press, 2005.

[45] C. Forlines, A. Esenther, C. Shen, D. Wigdor, K. Ryall.*Multi-user, Multi-display Interaction with a Single-user, Single-display Geospatial Application*, in "Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology", UIST '06, ACM, 2006, p. 273–276, http://doi.acm.org/10.1145/1166253.1166296.

[46] J. Grudin.*Computer-Supported Cooperative Work: History and Focus*, in "IEEE Computer", May 1994, vol. 27, n° 5, p. 19–26, http://dx.doi.org/10.1109/2.291294.

[47] T. Heath, C. Bizer.*Linked Data: Evolving the Web into a Global Data Space*, Morgan & Claypool, 2011.

[48] J. Hebeler, M. Fisher, R. Blace, A. Perez-Lopez.*Semantic web programming*, Wiley & Sons, 2011.

[49] M. Hildebrand, J. van Ossenbruggen, L. Hardman./facet: *A Browser for Heterogeneous Semantic Web Repositories*, in "ISWC'06: Proc. of the International Semantic Web Conference", Springer, 2006, p. 272–285.

[50] D. F. Huynh, R. C. Miller, D. Karger.*Potluck: Data Mash-up Tool for Casual Users*, in "Web Semantics", 2008, vol. 6, n° 4, p. 274–282, http://dx.doi.org/10.1016/j.websem.2008.09.005.

[51] H. Ishii, B. Ullmer.*Tangible Bits: Towards Seamless Interfaces Between People, Bits and Atoms*, in "Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems", CHI '97, ACM, 1997, p. 234–241, http://doi.acm.org/10.1145/258549.258715.

[52] J. A. Jacko.*Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, 3rd Edition*, 3rd, CRC Press, 2012.

[53] D. Karger, S. Ostler, R. Lee.*The Web Page As a WYSIWYG End-user Customizable Database-backed Information Management Application*, in "Proc. of the Symposium on User Interface Software and Technology", UIST '09, ACM, 2009, p. 257–260, http://doi.acm.org/10.1145/1622176.1622223.

[54] A. Katifori, C. Halatsis, G. Lepouras, C. Vassilakis, E. Giannopoulou.*Ontology visualization methods—a survey*, in "ACM Computing Surveys", 2007, vol. 39, n° 4, p. 10:1-10:42.

[55] C. Liu, O. Chapuis, M. Beaudouin-Lafon, E. Lecolinet, W. Mackay.*Effects of Display Size and Navigation Type on a Classification Task*, in "Proceedings of the 32nd international conference on Human factors in computing systems", CHI '14, ACM, 2014, p. 4147-4156, http://doi.acm.org/10.1145/2556288.2557020.

[56] C. Liu, O. Chapuis, E. Lecolinet, M. Beaudouin-Lafon.*Shared Interaction on a Wall-Sized Display in a Data Manipulation Task*, in "Proceedings of the SIGCHI Conference on Human Factors in Computing Systems", New York, NY, USA, CHI '16, ACM, 2016, To appear.

[57] W. MACKAY.*Is paper safer? The role of paper flight strips in air traffic control*, in "ACM Trans. Comput.-Hum. Interact.", December 1999, vol. 6, p. 311–340, http://doi.acm.org/10.1145/331490.331491.

[58] R. MORALES, C. APPERT, G. BAILLY, E. PIETRIGA.*TouchTokens: Guiding Touch Patterns with Passive Tokens*, in "Proceedings of the SIGCHI Conference on Human Factors in Computing Systems", New York, NY, USA, CHI '16, ACM, 2016, To appear.

[59] T. MUNZNER.*Visualization Analysis and Design*, CRC Press, 2014.

[60] M. NANCEL, O. CHAPUIS, E. PIETRIGA, X.-D. YANG, P. P. IRANI, M. BEAUDOUIN-LAFON.*High-precision pointing on large wall displays using small handheld devices*, in "Proceedings of the SIGCHI Conference on Human Factors in Computing Systems", CHI '13, ACM, 2013, p. 831–840, http://doi.acm.org/10.1145/2470654.2470773.

[61] A.-C. NGONGA NGOMO, L. BÜHMANN, C. UNGER, J. LEHMANN, D. GERBER.*Sorry, I Don'T Speak SPARQL: Translating SPARQL Queries into Natural Language*, in "Proc. of the International Conference on World Wide Web", WWW '13, ACM, 2013, p. 977–988, http://dl.acm.org/citation.cfm?id=2488388.2488473.

[62] H. OLAFSDOTTIR, T. TSANDILAS, C. APPERT.*Prospective Motor Control on Tabletops: Planning Grasp for Multitouch Interaction*, in "Proceedings of the 32nd international conference on Human factors in computing systems", CHI '14, ACM, 2014, p. 2893-2902, http://doi.acm.org/10.1145/2556288.2557029.

[63] E. PIETRIGA, C. APPERT, M. BEAUDOUIN-LAFON.*Pointing and beyond: an operationalization and preliminary evaluation of multi-scale searching*, in "CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems", ACM Press, 2007, p. 1215–1224 [*DOI :* 10.1145/1240624.1240808].

[64] E. PIETRIGA, C. APPERT.*Sigma lenses: focus-context transitions combining space, time and translucence*, in "CHI '08: Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems", ACM, 2008, p. 1343–1352 [*DOI :* 10.1145/1357054.1357264].

[65] E. PIETRIGA, O. BAU, C. APPERT.*Representation-Independent In-Place Magnification with Sigma Lenses*, in "IEEE Transactions on Visualization and Computer Graphics (TVCG)", 2010, vol. 16, n° 03, p. 455-467 [*DOI :* 10.1109/TVCG.2009.98].

[66] E. PIETRIGA, C. BIZER, D. KARGER, R. LEE.*Fresnel - A Browser-Independent Presentation Vocabulary for RDF*, in "Proceedings of the 5th International Semantic Web Conference (ISWC 2006)", Springer, November 2006, p. 158-171.

[67] E. PIETRIGA, S. HUOT, M. NANCEL, R. PRIMET.*Rapid development of user interfaces on cluster-driven wall displays with jBricks*, in "Proceedings of the 3rd ACM SIGCHI symposium on Engineering interactive computing systems", EICS '11, ACM, 2011, p. 185–190, http://doi.acm.org/10.1145/1996461.1996518.

[68] E. PIETRIGA.*A Toolkit for Addressing HCI Issues in Visual Language Environments*, in "IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)", IEEE Computer Society, 2005, p. 145-152 [*DOI :* 10.1109/VLHCC.2005.11].

[69] C. PINDAT, E. PIETRIGA, O. CHAPUIS, C. PUECH.*JellyLens: Content-Aware Adaptive Lenses*, in "UIST - 25th Symposium on User Interface Software and Technology - 2012", Cambridge, MA, United States,

Proceedings of the 25th Symposium on User Interface Software and Technology, ACM, October 2012, p. 261-270 [*DOI :* 10.1145/2380116.2380150], https://hal.inria.fr/hal-00721574.

[70] C. PINDAT, E. PIETRIGA, O. CHAPUIS, C. PUECH.*Drilling into complex 3D models with gimlenses*, in "Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology", VRST '13, ACM, 2013, p. 223–230, http://doi.acm.org/10.1145/2503713.2503714.

[71] N. SHADBOLT, T. BERNERS-LEE, W. HALL.*The Semantic Web Revisited*, in "IEEE Intelligent Systems", 2006, vol. 21, n° 3, p. 96-101, http://doi.ieeecomputersociety.org/10.1109/MIS.2006.62.

[72] O. SHAER, E. HORNECKER.*Tangible User Interfaces: Past, Present, and Future Directions*, in "Found. Trends Hum.-Comput. Interact.", January 2010, vol. 3, n° 1–2, p. 1–137, http://dx.doi.org/10.1561/1100000026.

[73] B. SHNEIDERMAN, C. PLAISANT.*Designing the User Interface, 4th Edition*, Pearson Addison Wesley, 2005.

[74] D. SPELMEZAN, C. APPERT, O. CHAPUIS, E. PIETRIGA.*Controlling widgets with one power-up button*, in "Proceedings of the 26th annual ACM symposium on User interface software and technology", UIST '13, ACM, 2013, p. 71–74, http://doi.acm.org/10.1145/2501988.2502025.

[75] D. SPELMEZAN, C. APPERT, O. CHAPUIS, E. PIETRIGA.*Side pressure for bidirectional navigation on small devices*, in "Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services", MobileHCI '13, ACM, 2013, p. 11–20, http://doi.acm.org/10.1145/2493190.2493199.

[76] G. TOUYA, J.-F. GIRRES.*ScaleMaster 2.0: a ScaleMaster extension to monitor automatic multi-scales generalizations*, in "Cartography and Geographic Information Science", 2013, vol. 40, n° 3, p. 192–200, http://dx.doi.org/10.1080/15230406.2013.809233.

[77] G. TOUYA.*Social Welfare to Assess the Global Legibility of a Generalized Map*, in "International Conference on Geographic Information", Springer, 2012, p. 198–211, http://dx.doi.org/10.1007/978-3-642-33024-7_15.

[78] T. TSANDILAS, C. APPERT, A. BEZERIANOS, D. BONNET.*Coordination of Tilt and Touch in One- and Two-Handed Use*, in "Proceedings of the 32nd international conference on Human factors in computing systems", CHI '14, ACM, 2014, p. 2001-2004, http://doi.acm.org/10.1145/2556288.2557088.

[79] G. TUMMARELLO, R. DELBRU, E. OREN.*Sindice.com: Weaving the Open Linked Data*, in "Proc. of the 6th Int. Semantic Web Conf. (ISWC2007)", Springer-Verlag, 2007, p. 552–565.

[80] C. WARE.*Information visualization: perception for design*, Elsevier, 2012.

[81] R. A. DE ALMEIDA, C. PILLIAS, E. PIETRIGA, P. CUBAUD.*Looking behind bezels: french windows for wall displays*, in "Proceedings of the International Working Conference on Advanced Visual Interfaces", AVI '12, ACM, 2012, p. 124–131, http://doi.acm.org/10.1145/2254556.2254581.

Project-Team INFINE

INFormation NEtworks

Table of contents

Project-Team INFINE

Creation of the Team: 2014 April 01, updated into Project-Team: 2016 December 01

Keywords:

Computer Science and Digital Science:

   1.2.3. - Routing
   1.2.5. - Internet of things
   1.2.6. - Sensor networks
   1.2.9. - Social Networks
   1.3. - Distributed Systems
   2.6.1. - Operating systems
   3.3.2. - Data mining
   3.4.2. - Unsupervised learning
   3.5. - Social networks
   3.5.1. - Analysis of large graphs
   3.5.2. - Recommendation systems
   4.8. - Privacy-enhancing technologies
   5.9.2. - Estimation, modeling

Other Research Topics and Application Domains:

   4.4. - Energy delivery
   4.4.1. - Smart grids
   6.4. - Internet of things
   8.1.2. - Sensor networks for smart buildings
   8.2. - Connected city

1. Members

Research Scientists
   Laurent Massoulié [Team leader, Inria, Researcher, HDR]
   Cedric Adjih [Inria, Researcher]
   Emmanuel Baccelli [Inria, Researcher, HDR]
   Aline Carneiro Viana [Inria, Researcher, HDR]

Technical Staff
   Francisco Javier Acosta Padilla [Inria]
   Oliver Hahm [Inria]
   Virag Shah [Inria]

PhD Students
   Guangshuo Chen [Inria]
   Loic Dauphin [Inria]
   Lennart Gulikers [Microsoft Research-Inria Joint Centre]
   Rémi Varloot [Microsoft Research-Inria Joint Centre]

Post-Doctoral Fellows
   Ehsan Ebrahimi Khaleghi [Inria, from May 2016]
   Sahar Hoteit [Inria]

Administrative Assistant
 Helene Bessin Rousseau [Inria]
Others
 Panagiota Katsikouli [University of Edinburgh, PhD Student, Intern at Inria for 5 months]
 Roni Shigueta [PUC-BR, Brazil, PhD Student]
 Felipe Cunha [Federal University of Minas Gerais, Brazil, PhD Student]
 Jose Ignacio Alamos Aste [Inria, Intern, from Mar 2016 until Aug 2016]
 Ichrak Amdouni [TelecomSudParis, external collaborator]

2. Overall Objectives

2.1. Overall Objectives

The INFINE proposal aims to design and analyse novel communication paradigms, protocols and architectures based on concepts of ultra distributed, information- and user-centric networking. The project is motivated by the recent and forthcoming evolution of Internet uses. Based on an information- and user-centric perspective, not only does it address issues pertaining to physical communication networks such as traffic routing, regulation and caching, but also issues about online social networks such as content recommendation and privacy protection.

INFINE team is engaged in research along three main themes: *Online social networking*, *Traffic and Resource Management*, and *Spontaneous Wireless Networks*. All these research activities encompass both theoretical research (on elaboration of models, algorithms, protocols and formal characterization of their performance), and applied research (to validate and/or experiment the proposed concepts against real networking scenarios). INFINE fits in the theme "Networks and Telecommunications" of the research field "Networks, Systems and Services, Distributed Computing" at Inria.

2.1.1. *New challenging demands*

Nowadays, we use networks not only to transport information from where it resides to ourselves but also, with online social networks, to determine what information might be of interest to us. Such a social recommendation functionality holds the promise of allowing us to access more relevant information. At the same time there is ample scope for improving its efficiency. Moreover it creates threats to user privacy.

At the same time, the physical context in which we access communication networks has drastically changed. While in the past, Internet was mostly accessed through fixed desktop computers, users are now mobile about 50% of their time online. In addition, while communicating machines used to be sparse and wired, with the advent of the Internet of Things we now evolve in a dense, interconnected environment of heterogeneous devices communicating via wireless and/or via wires.

This new context of Internet uses challenges several aspects of currently deployed networks. Some aspects pertain to the physical architecture of the Internet. In particular, at the core of the Internet, a drastic increase in volume of data traffic is anticipated due to the emergence of new applications, generalization of cloud services, or the advent of the Internet of Things (IoT) and Machine-to-Machine (M2M) communications. On the other hand, at the edge of the Internet, user mobility and today's pervasiveness of computing devices with increasingly higher capabilities (i.e., processing, storage, sensing) have a fundamental impact on the adequacy of algorithms and communication mechanisms.

Other aspects concern the logical architecture of the network. For instance, currently deployed protocols at layers above IP must now carry massive publish-subscribe traffic, preserve user privacy, be social-aware, and support delay tolerant communications and paradigms for which they were not initially designed. Concerning actual content distribution, the avalanche of data and privacy concerns puts more and more pressure on filter/push mechanisms to provide users with relevant information.

While considering physical and logical aspects of networks, the INFINE team will pursue research activities combining theoretical and experimental approaches.

2.1.2. Research agenda

Our general goal is to develop distributed mechanisms for optimizing the operation of networks both at the mentioned logical and physical levels of the architecture. Taking an information- and user-centric perspective, we envision networks as means to convey relevant information to users, while adapting to customary practices (in terms of context, interests, or content demands) of such users. At the logical level, online social networks (OSNs) allow users to choose what information to access. At the physical level, communication, computation, and memory resources allow users to retrieve some content eventually selected on the basis of the online social network.

The two setups feature scarce resources: for instance, in OSNs, these are the users' budget of attention, which must be used sparingly by recommending only relatively few potential content items. At the physical level this is typically the channels' capacity or networking resources, which cannot be oversubscribed.

Beyond a formal resemblance between the optimizations that one must carry at these two levels, there is a strong commonality in the methods adequate for conducting optimizations in the two setups. To illustrate this point, consider *contact recommendation*, that is a key objective in our agenda on online social networks. This entails automatically proposing to users potential contacts for optimizing the subsequent efficiency of social content filtering. We envision addressing contact recommendation by first performing some community detection, i.e. identification of similarly behaving users. Similarly, at the physical level, user-centric approaches, sometimes also related to community detection, have guided routing decisions in challenged network environments, where delay-tolerant networking is used. Still, associated with dynamic centrality metrics, community detection can guide the replication of a specific content in well-selected users, while exploiting the advantages of distributed decentralized storage and opportunistic communications.

As an additional example at the logical level, we consider *content recommendation*, whereby a list of potential contents is filtered before being presented to a user, with the aim of maximizing the chance this user finds an item of interest therein. This has an exact analogue at the physical level, where by taking an information- and user-centric approach, we intend to off-load communication resources via pre-loaded content replicas at various storage points in the network. The problem of determining which content to cache so as to maximize the chance of it being accessed in the vicinity of the corresponding cache memory corresponds precisely to the aforementioned content recommendation problem.

We now detail further our agenda along three main specific axes, namely Online Social Networks, Traffic and Resource Management, and Spontaneous Wireless Networks/Internet of Things, bearing in mind that we will develop generic solutions relevant to several of these axes wherever possible.

3. Research Program

3.1. Online Social Networks (OSN)

Large-scale online social networks such as Twitter or FaceBook provide a powerful means of selecting information. They rely on "social filtering", whereby pieces of information are collectively evaluated and sorted by users. This gives rise to information cascades when one item reaches a large population after spreading much like an epidemics from user to user in a viral manner. Nevertheless, such OSNs expose their users to a large amount of content of no interest to them, a sign of poor "precision" according to the terminology of information retrieval. At the same time, many more relevant content items never reach those users most interested in them. In other words, OSNs also suffer from poor "recall" performance.

This leads to a first challenge: *what determines the optimal trade-off between precision and recall in OSNs? And what mechanisms should be deployed in order to approach such an optimal trade-off?* We intend to study this question at a theoretical level, by elaborating models and analyses of social filtering, and to validate the resulting hypotheses and designs through experimentation and processing of data traces. More specifically, we envision to reach this general objective by solving the following problems.

3.1.1. *Community Detection*

Identification of implicit communities of like-minded users and contact recommendation for helping users "rewire" the information network for better performance. Potential schemes may include variants of spectral clustering and belief propagation-style message passing. Limitations / relative merits of candidate schemes, their robustness to noise in the input data, will be investigated.

3.1.2. *Incentivization*

Design of incentive mechanisms to limit the impact of users' selfishness on system behavior: efficiency should be maintained even when users are gaming the system to try and increase their estimated expertise. By offering rewards to users on the basis of their involvement in filtering and propagation of content, one might encourage them to adjust their action and contribute to increase the overall efficiency of the OSN as a content access platform.

One promising direction will be to leverage the general class of Vickrey-Clarke-Groves incentive-compatible mechanisms of economic theory to design so-called marginal utility reward mechanisms for OSN users.

3.1.3. *Social Recommendation and Privacy*

So far we have only alluded to the potential benefits of OSNs in terms of better information access. We now turn to the risks they create. Privacy breaches constitute the greatest of these risks: OSN users disclose a wealth of personal information and thereby expose themselves to discrimination by potential employers, insurers, lenders, government agencies...Such privacy concerns are not specific to OSNs: internauts' online activity is discretely tracked by companies such as Bluekai, and subsequently monetized to advertisers seeking better ad targeting. While disclosure of personal data creates a privacy risk, on the other hand it fuels personalized services and thereby potentially benefits everyone.

One line of research will be to focus on the specific application scenario of content categorization, and to characterize analytically the trade-off between user privacy protection (captured by differential privacy), accuracy of content categorization, and sample complexity (measured in number of probed users).

3.2. Traffic and Resource Management

Despite the massive increases in transmission capacity of the last few years, one has every reason to believe that networks will remain durably congested, driven among other factors by the steadily increasing demand for video content, the proliferation of smart devices (i.e., smartphones or laptops with mobile data cards), and the forecasted additional traffic due to machine-to-machine (M2M) communications. Despite this rapid traffic growth, there is still a rather limited understanding of the features protocols have to support, the characteristics of the traffic being carried and the context where it is generated. There is thus a strong need for smart protocols that transport requested information at the cheapest possible cost on the network as well as provide good quality of service to network subscribers. One particularly new aspect of up-and-coming networks is that networks are now used to not only (i) access information, but also (ii) distributively process information, en-route.

We intend to study these issues at the theoretical and protocol design levels, by elaborating models and analysis of content demands and/or mobility of network subscribers. The resulting hypothesis and designs will be validated through experimentation, simulation, or data trace processing. It is also worth mentioning the provided solutions may bring benefits to different entities in the network: to content owners (if applied at the core of Internet) or to subscribers or network operators (if applied at the edge of the Internet).

3.2.1. *At the Internet Core*

One important optimization variable consists in content replication: users can access the closest replica of the content they are interested in. Thus the memory resource can be used to create more replicas and reduce the usage of the bandwidth resource. Another interesting arbitrage between resources arises because content is no longer static but rather dynamic. Here are two simple examples: i) a video could be encoded at several resolutions. There is then a choice between pre-recording all possible resolutions, or alternatively synthesizing

a lower-resolution version on the fly from a higher resolution version when a request arises. ii) A user requests the result of a calculation, say the average temperature in a building; this can either be kept in memory, or recomputed each time such a query arises. Optimizing the joint use of all three resources, namely bandwidth, memory, computation, is a complex task. Content Delivery Network companies such as Akamai or Limelight have worked on the memory/bandwidth trade-off for some years, but as we will explain more can be done on this. On the other hand optimizing the memory/computation trade-off has received far less attention. We aim to characterize the best possible content replication strategies by leveraging fine-grained prediction of i) users' future requests, and ii) wireless channels' future bandwidth fluctuations. In the past these two determining inputs have only been considered at a coarse-grained, aggregate level. It is important to assess how much bandwidth saving can be had by conducting finer-grained prediction. We are developing light-weight protocols for conducting these predictions and automatically instantiating the corresponding optimal replication policies. We are also investigating generic protocols for automatically trading replication for computation, focusing initially on the above video transcoding scenario.

3.2.2. *At the Internet Edge*

Cellular and wireless data networks are increasingly relied upon to provide users with Internet access on devices such as smartphones, laptops or tablets. In particular, the proliferation of handheld devices equipped with multiple advanced capabilities (e.g., significant CPU and memory capacities, cameras, voice to text, text to voice, GPS, sensors, wireless communication) has catalyzed a fundamental change in the way people are connected, communicate, generate and exchange data. In this evolving network environment, users' social relations, opportunistic resource availability, and proximity between users' devices are significantly shaping the use and design of future networking protocols.

One consequence of these changes is that mobile data traffic has recently experienced a staggering growth in volume: Cisco has recently foreseen that the mobile data traffic will increase 18-fold within 2016, in front of a mere 9-fold increase in connection speeds. Hence, one can observe today that the inherently centralized and terminal-centric communication paradigm of currently deployed cellular networks cannot cope with the increased traffic demand generated by smartphone users. This mismatch is likely to last because (1) forecasted mobile data traffic demand outgrows the capabilities of planned cellular technological advances such as 4G or LTE, and (2) there is strong skepticism about possible further improvements brought by 5G technology.

Congestion at the Internet's edge is thus here to stay. Solutions to this problem relates to: densify the infrastructure, opportunistically forward data among neighbors wireless devices, to offload data to alternate networks, or to bring content from the Internet closer to the subscribers. Our recent work on leveraging user mobility patterns, contact and inter-contact patterns, or content demand patterns constitute a starting point to these challenges. The projected increase of mobile data traffic demand pushes towards additional complementary offloading methods. Novel mechanisms are thus needed, which must fit both the new context that Internet users experience now, and their forecasted demands. In this realm, we will focus on new approaches leveraging ultra-distributed, user-centric approaches over IP.

3.3. Spontaneous Wireless Networks (SWN) and Internet of Things (IoT)

The unavailability of end-to-end connectivity in emergent wireless mobile networks is extremely disruptive for IP protocols. In fact, even in simpler cases of spontaneous wireless networks where end-to-end connectivity exists, such networks are still disruptive for the standard IP protocol stack, as many protocols rely on atomic link-local services (such as link-local multicast/broadcast), while these services are inherently unavailable in such networks due to their opportunistic, wireless multi hop nature. In this domain, we will aim to characterize the achievable performance in such IP-disruptive networks and to actively contribute to the design of new, deployable IP protocols that can tolerate these disruptions, while performing well enough compared to what is achievable and remaining interoperable with the rest of the Internet.

Spontaneous wireless networking is also a key aspect of the Internet of Things (IoT). The IoT is indeed expected to massively use this networking paradigm to gradually connect billions of new devices to the Internet, and drastically increase communication without human source or destination – to the point where the

amount of such communications will dwarf communications involving humans. Large scale user environment automation require communication protocols optimized to efficiently leverage the heterogeneous and unreliable wireless vicinity (the scope of which may vary according to the application). In fact, extreme constraints in terms of cost, CPU, battery and memory capacities are typically experienced on a substantial fraction of IoT devices. We expect that such constraints will not vanish any time soon for two reasons. On one hand the progress made over the last decade concerning the cost/performance ratio for such small devices is quite disappointing. On the other hand, the ultimate goal of the IoT is ubiquitous Internet connectivity between devices as tiny as dust particles. These constraints actually require to redesign not only the network protocol stack running on these devices, but also the software platform powering these machines. In this context, we will aim at contributing to the design of novel network protocols and software platforms optimized to fit these constraints while remaining compatible with legacy Internet.

3.3.1. *Design & Development of Open Experimental IoT Platforms*

Manufacturers announce on a regular basis the availability of novel tiny devices, most of them featuring network interfaces: the Internet of Things (IoT) is already here, from the hardware perspective, and it is expected in the near future that we will see a massive increase of the number of muti-purpose smart objects (from tiny sensors in industrial automation to devices like smart watches and tablets). Thus, one of the challenges is to be able to test architectures, protocols and applications, in realistic conditions and at large scale.

One necessity for research in this domain is to establish and improve IoT hardware platforms and testbeds, that integrate representative scenarios (such as Smart Energy, Home Automation etc.) and follow the evolution of technology, including radio technologies, and associated experimentation tools. For that, we plan to build upon the IoT-LAB federated testbeds, that we have participated in designing and deploying recently. We plan to further develop IoT-LAB with more heterogeneous, up-to-date IoT hardware and radios that will provide a usable and realistic experimentation environment. The goal is to provide a tool that enables testing a validation of upcoming software platforms and network stacks targeting concrete IoT deployments.

In parallel, on the software side, IoT hardware available so far made it uneasy for developers to build apps that run across heterogeneous hardware platforms. For instance Linux does not scale down to small, energy-constrained devices, while microcontroller-based OS alternatives were so far rudimentary and yield a steep learning curve and lengthy development life-cycles because they do not support standard programming and debugging tools. As a result, another necessity for research in this domain is to allow the emergence of it more powerful, unifying IOT software platforms, to bridge this gap. For that, we plan to build upon RIOT, a new open source software platform which provides a portable, Linux-like API for heterogeneous IoT hardware. We plan to continue to develop the systems and network stacks aspects of RIOT, within the open source developer community currently emerging around RIOT, which we co-founded together with Freie Universitaet Berlin. The key challenge is to improve usability and add functionalities, while maintaining architectural consistency and a small enough memory footprint. The goal is to provide an IoT software platform that can be used like Linux is used for less constrained machines, both (i) in the context of research and/or teaching, as well as (ii) in industrial contexts. Of course, we plan to use it ourselves for our own experimental research activities in the domain of IoT e.g., as an API to implement novel network protocols running on IoT hardware, to be tested and validated on IoT-LAB testbeds.

3.3.2. *Design & Standardization of Architectures and Efficient Protocols for Internet of Things*

As described before, and by definition, the Internet of Things will integrate not only a massive number of homogeneous devices (e.g., networks of wireless sensors), but also heterogeneous devices using various communication technologies. Most devices will be very constrained resources (memory resources, computational resources, energy). Communicating with (and amongst) such devices is a key challenge that we will focus on. The ability to communicate efficiently, to communicate reliably, or even just to be able to communicate at all, is non-trivial in many IoT scenarios: in this respect, we intend to develop innovative protocols, while following and contributing to standardization in this area. We will focus and base most of our work on standards developed in the context of the IETF, in working groups such as 6lo, CORE, LWIG etc., as well as IRTF

research groups such as NWCRG on network coding and ICNRG on Information Centric Networking. We note however that this task goes far beyond protocol design: recently, radical rearchitecturing of the networks with new paradigms such as Information Centric Networking, ICN, (or even in wired networks, software-defined networks), have opened exciting new avenues. One of our direction of research will be to explore these content-centric approaches, and other novel architectures, in the context of IoT.

4. Highlights of the Year

4.1. Highlights of the Year

4.1.1. *Conferences and Presentations*

We organized a high-profile conference in May 2016 at the Institut Henri Poincaré on "Networks: learning, information and complexity" (see: http://www.msr-inria.fr/conferences-workshops/workshop-on-networks-learning-information-and-complexity/) which gathered leading scientists in computer science, maths and statistical physics.

We organized in January 2016 a workshop at the Turing building involving top executives of LVMH and Inria researchers to exchange on innovation opportunities for LVMH notably around advertising with online social networks, data visualization, and computer vision.

We gave several invited talks at: Stochastic Networks Conference, UCSD; CIRM workshop on random matrices; Institut Henri Poincaré's "Nexus" of Information and Computation Theories; EPFL workshop for birthdays of Shannon, Urbanke and Telatar (see: http://www.etru50.com/invited-speakers/) ; LINCS scientific advisory board.

4.1.2. *RIOT Summit*

We successfully organized in July 2016 the first RIOT Summit in Berlin. The RIOT Summit 2016 gathered 100+ enthusiastic industrial participants, makers and academics involved in RIOT. Relevant partners such as Cisco, Samsung, Siemens, Nordic Semiconductors, as well as a number of SMEs and startups from various places in Europe gave talks on aspects of IoT communication, use cases IoT hardware, IoT open source community aspects and concepts for future IoT software and networks, as well as hands-on sessions and tutorials. See: http://summit.riot-os.org/#speakers.

4.1.3. *Opening of the IoT-LAB experimental platform at the site Saclay*

The project Equipex FIT deploys experimental facilities on several sites. In 2016, at the site of Saclay, the opening of the FIT IoT-LAB site followed the move from its previous location at Rocquencourt.

The platform of Saclay is an Internet-of-Things testbed and includes more than 300 nodes (175 A8-M3, 12 M3 and 120 WSN430), deployed in large experimentation rooms and space. All A8 nodes are equipped with GPS.

More information about the topology and the resources of this new site is available here: https://www.iot-lab.info/deployment/saclay/.

4.1.4. *Awards*

The team members have received a number of awards:

M1 intern Davi Castro de Silva received best internship prize of LIX for his work on modifying spectral methods for community detection to increase their robustness.

Best Poster Award   [26] O. Hahm, C. Adjih, E. Baccelli, T. C. Schmidt, M. Waehlisch.
   ICN over TSCH: Potentials for Link-Layer Adaptation in the IoT, September 2016, pp. 195-196, ACM-ICN '16 Proceedings of the 3rd ACM Conference on Information-Centric Networking , Poster. [DOI:10.1145/2984356.2985226] https://hal.inria.fr/hal-01369704

Best Demo Award   [18] H. Petersen, C. Adjih, O. Hahm, E. Baccelli.

Demo: IoT Meets Robotics - First Steps, RIOT Car, and Perspectives, in: ACM International Conference on Embedded Wireless Systems and Networks (EWSN), Graz, Austria, February 2016. https://hal.inria.fr/hal-01262638

BEST PAPERS AWARDS :

[15] **5th IFIP International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks (PEMWN 2016)**. F. SOMAA, C. ADJIH, I. E. KORBI, L. A. SAIDANE.

5. New Software and Platforms

5.1. RIOT

KEYWORDS: Internet of things - Operating system - Sensors - Iot - Wireless Sensor Networks - Internet protocols

SCIENTIFIC DESCRIPTION

While requiring as low as 1,5kB of RAM and 5kB or ROM, RIOT offers real time and energy efficiency capabilities, as well as a powerful API (partially POSIX compliant) that is consistent across heterogeneous low-end IoT hardware (8-bit, 16-bit and 32-bit architectures). This API is developer-friendly in that it enables multi-threading, standard C and C++ application programming (as well as scripting) and the use of standard debugging tools – all of which was not possible so far for such embedded programming. On top of this, RIOT supports a large number of software libraries (e.g. crypto, maths, drivers...) and aggregates in a simple manner a wide variety of third-party open source software packages. In particular, various network stacks are available with RIOT, such as a standard IPv6/6LoWPAN stack and a information-centric network stack (based on CCN).

FUNCTIONAL DESCRIPTION

RIOT is an Open Source operating system that provides standard protocols for embedded systems. RIOT allows, for example, the development of applications that collect sensor data and transmit it to a central node (e.g. a server). This data can then be used for smart energy management for instance. Other use-cases include also IoT-enabled low-cost robots.

RIOT is specially designed for embedded systems, which are strongly constrained in memory and energy. Further, RIOT can easily be ported to different hardware devices and follows the latest evolution of IP standards.

RIOT applications can readily be tested in the FIT IoT-Lab, which provides a large-scale infrastructure facility with 3000 nodes for testing remotely small wireless devices.

- Participants: Emmanuel Baccelli, Oliver Hahm, Cedric Adjih, Francisco Acosta
- Partner: Freie Universität Berlin
- Contact: Emmanuel Baccelli
- URL: https://github.com/RIOT-OS/RIOT

5.2. MACACO

Mobile context-Adaptive CAching for COntent-centric networking

FUNCTIONAL DESCRIPTION

MACACOapp is developed in the context of the EU CHIST-ERA MACACO project. It consists in a mobile phone application that periodically samples phone's information on the mobility (through, e.g., GPS sensor, accelerometer and WiFi/Bluetooth/Cellular environment, connectivity type) and on the data traffic it generates (through, e.g., Internet browser history and applications data consumption). The information collected will be time-stamped and will be periodically sent to the central servers for analysis and visualization. We expect that (1) the collected information will allow us studying the correlation between mobility and content demand patterns and that (2) the results of this analysis will allow us inferring the best times and places to transfer content from/to users' phones location and/or from/to the wireless infrastructure closest to the users' phones location. Users will be also invited to fill a non-mandatory questionnaire relevant to this study. Our questionnaire collects information about the personality traits and application preferences of people. We expect that the information collected from questionnaire will allow us to analyse the correlation between users personality traits and their application preferences and interests. User's application preferences and interests will be inferred from the Internet browsing history and running app information obtained from the MACACO App.

- Participants: Aline Carneiro Viana, Katia Jaffres and Marco Fiore
- Contact: Aline Carneiro Viana
- URL: https://macaco.inria.fr/macacoapp/

5.3. GardiNet (previously DragonNet)

FUNCTIONAL DESCRIPTION

GardiNet (previously known as DragonNet) is a generic framework for network coding in wireless networks. It is initially a result of the GETRF project of the Hipercom2 team.

It is based on intra-flow coding where the source divides the flow in a sequence of payloads of equal size (padding may be used). The design keys of DragonNet are simplicity and universality, GardiNet does not use explicit or implicit knowledge about the topology (such as the direction or distance to the source, the loss rate of the links, ...). Hence, it is perfectly suited to the most dynamic wireless networks. The protocol is distributed and requires minimal coordination. DragonNet architecture is modular, it is based on 5 building blocks (LIB, SIG, Protocol, SEW and DRAGON). Each block is almost independent. This makes DragonNet generic and hence adaptable to many application scenarios. DragonNet derives from a prior protocol called DRAGONCAST. Indeed, DragonNet shares the same principles and theoretical overview of DRAGONCAST. It enriches DRAGONCAST by the information base and signaling required to perform broadcast in wireless networks and in wireless sensor networks in particular.

- Participants: Cedric Adjih, Ichrak Amdouni, Hana Baccouch and Antonia Masucci
- Contact: Cedric Adjih
- URL: http://getrf.gforge.inria.fr/https://github.com/GardiNet/gardinet

6. New Results

6.1. Online Social Networks (OSN)

Community detection; bandit algorithms; privacy preservation; reward mechanisms

6.1.1. Capacity of Information Processing Systems

Participants: Laurent Massoulié, Kuang Xu.

We propose and analyze a family of information processing systems, where a finite set of experts or servers are employed to extract information about a stream of incoming jobs. Each job is associated with a hidden label drawn from some prior distribution. An inspection by an expert produces a noisy outcome that depends both on the job's hidden label and the type of the expert, and occupies the expert for a finite time duration. A decision maker's task is to dynamically assign inspections so that the resulting outcomes can be used to accurately recover the labels of all jobs, while keeping the system stable. Among our chief motivations are applications in crowd-sourcing, diagnostics, and experiment designs, where one wishes to efficiently learn the nature of a large number of items, using a finite pool of computational resources or human agents. We focus on the capacity of such an information processing system. Given a level of accuracy guarantee, we ask how many experts are needed in order to stabilize the system, and through what inspection architecture. Our main result provides an adaptive inspection policy that is asymptotically optimal in the following sense: the ratio between the required number of experts under our policy and the theoretical optimal converges to one, as the probability of error in label recovery tends to zero.

This work was accepted and presented under the title "On the capacity of information processing systems" at the COLT 2016 conference.

6.2. Spontaneous Wireless Networks and Internet of Things

internet of things; wireless sensor networks; dissemination; resource management

6.2.1. Platform Design for the Internet of Things

Participants: Emmanuel Baccelli, Cedric Adjih, Oliver Hahm, Francisco Acosta, Hauke Petersen.

Within this activity, we have further developed the platforms we champion for the Internet of Things: the open source operating system RIOT on one hand, and open-access IoT-lab testbeds on the other hand. RIOT now aggregates open source contributions from 130+ people (and counting) from all over the world, coming both from academia and from industry, and received financial backing from top companies including Cisco and Google. We further developed RIOT for low-cost mobile robots and received the Best Demo Award at the ACM EWSN'16 conference for our work on this topic. As steering RIOT community members, we also participated in the prestigious Internet Architecture Board (IAB) workshop on IoT Software Updates, a hot and essential topic for the future of Internet of Things. The year culminated in this domain with the successful organization of the first RIOT Summit in Berlin, where 100+ participants from all over the world, from industry, academia as well as hackers/makers involved in RIOT gathered to discuss various aspects of the future of RIOT and open source IoT software. In addition, 2016, at the site of Saclay, one of the testbeds from FIT IoT-LAB was opened: the platform of Saclay includes more than 300 IoT nodes (175 A8-M3, 12 M3, 120 WSN430, some Arduinos and some SAMR21-xpro). In parallel, the platform from Freie Universitat Berlin also joined the OneLab/FIT IoT-LAB testbed federation.

6.2.2. Energy-Efficient Communication Protocols for the Internet of Things

Participants: Oliver Hahm, Emmanuel Baccelli, Cedric Adjih, Matthias Waehlisch, Thomas Schmidt.

Within this activity, we have designed distributed algorithms providing improved trade-off between content availability and energy efficiency (which plays a crucial role). The approach we developed leverages distributed caching for IoT content, based on an information-centric networking paradigm. We extended the NDN protocol with a variety of caching and replacement strategies, and we analyzed alternative approaches for extending NDN to accommodate such IoT use cases. Based on extensive experiments on real IoT hardware in a network gathering hundreds of nodes, we demonstrate these caching strategies can bring 90% reduction in energy consumption while maintaining IoT content availability above 90%. This work was published in IEEE Globecom'16 workshop on Named Data Networks for Challenged Communication Environments.

We also have designed new mechanisms to jointly exploit ICN communication patterns and dynamically optimize the use of TSCH (Time Slotted Channel Hopping), a wireless link layer technology increasingly popular in the IoT. Through a series of experiments on FIT IoT-LAB interconnecting typical IoT hardware, we find that our proposal is fully robust against wireless interference, and almost halves the energy consumed for transmission when compared to CSMA. Most importantly, our adaptive scheduling prevents the time-slotted MAC layer from sacrificing throughput and delay. Our work on ICN and on TSCH was published at NTMS'16, at ACM ICN'16, and in Proceedings of the IEEE.

6.2.3. Standards for Spontaneous Wireless Networks

Participant: Emmanuel Baccelli.

Within this activity, we have contributed to new network protocol standards for spontaneous wireless networking, applied to ad hoc networks and the Internet of Things. In particular, collaborating with Fraunhofer, we have published RFC 7779, standardizing Directional Airtime Metric (DAT), a new wireless metric standard targeting wireless mesh networks. Furthermore, collaborating with ARM and Sigma Designs, we published RFC 7733, which provides guidance in the configuration and use of protocols from the RPL protocol suite to implement the features required for control in building and home environments. In collaboration with various industrial partners, with have also published a number of other Internet drafts, including an analysis of the characteristics of multi-hop ad hoc wireless communication between interfaces in the context of IP networks, and an analysis of the challenges of information-centric networking in the Internet of Things.

6.2.4. Spatio-Temporal Predictability of Cellular Data Traffic

Participants: Guangshuo Chen, Aline Carneiro Viana, Marco Fiore, Sahar Hoteit, Carlos Sarraute.

The ability to foresee the data traffic activity of subscribers opens new opportunities to reshape mobile network management and services. In this work, we leverage two large-scale real-world datasets collected by a major mobile carrier in Mexico to study how predictable are the cellular data traffic demands generated by individual users. We focus on the predictability of mobile traffic consumption patterns in isolation. Our results show that it is possible to anticipate the individual demand with a typical accuracy of 85%, and reveal that this percentage is consistent across all user types. Despite the heterogeneity in usage patterns of users, we also find a lack of significant variability in predictability when considering demographic factors or different mobility or mobile service usage. We also analyze the joint predictability of the traffic demands and mobility patterns. We find that the two dimensions are correlated, which improves the predictability upper bound to 90% on average. This first work is in submission in an international conference.

6.2.5. Completion of Sparse Call Detail Records for Mobility Analysis

Participants: Guangshuo Chen, Aline Carneiro Viana, Marco Fiore, Sahar Hoteit.

Call Detail Records (CDRs) have been widely used in the last decades for studying different aspects of human mobility. The accuracy of CDRs strongly depends on the user-network interaction frequency: hence, the temporal and spatial sparsity that typically characterize CDR can introduce a bias in the mobility analysis. In this work, we evaluate the bias induced by the use of CDRs for inferring important locations of mobile subscribers, as well as their complete trajectories. Besides, we propose a novel technique for estimating real human trajectories from sparse CDRs. Compared to previous solutions in the literature, our proposed technique reduces the error between real and estimated human trajectories and at the same time shortens the temporal period where users' locations remain undefined. This work has been published as an invited paper at the ACM CHANTS 2016 workshop in conjunction with ACM MobiCom 2016. Related to CDRs, we have also investigated whether the information of user's instantaneous whereabouts provided by CDRs enables us to estimate positions over longer time spans. Our results confirm that CDRs ensure a good estimation of radii of gyration and important locations, yet they lose some location information. Most importantly, we show that temporal completion of CDRs is straightforward and efficient: thanks to the fact that they remain fairly static before and after mobile communication activities, the majority of users' locations over time can be accurately inferred from CDRs. Finally, we observe the importance of user's context, i.e.,of the size of the current network cell, on the quality of the CDR temporal completion. This work is in submission in an international conference.

Finally, driven by real-world data across a large population, we propose two approaches as the refinement of the legacy solution, which complete CDR data adaptively according to the information of users and activities. Our proposed methods outperform the legacy solution in terms of the combination of accuracy and temporal coverage. Besides, our work reveals the important factors to the data completion. This paper has been accepted for publication at the IEEE DAWM workshop in conjunction with IEEE Percom 2017.

6.2.6. *Completion of Sparse Call Detail Records for Mobility Analysis*
Participants: Panagiota Katsikouli, Aline Carneiro Viana, Marco Fiore, Alessandro Nordio, Alberto Tarable.

The increasing usage of smart devices and location-tracking systems has made it possible to study and understand the behaviour of users as well as human mobility at an unprecedented scale. The insights of such studies can help improve many aspects of our everyday lives, from road network infrastructure to mobile network quality of service. Human mobility is repetitive and regular.In addition to our tendency to revisit the same locations, those visits happen with relevant temporal regularity, where each visited location has been assigned with an ID. The daily interaction with our smart devices, such as smartphones, results in collecting fine grained information on our activities and whereabouts. This information can be used to detect and analyze the routinary behaviour of humans but also to discover interests, preferences and hidden patterns of mobility. However, frequent recording of data tends to quickly drain the battery of the smartphone. A natural alternative is to sample the collected data. Maintaining a summary or sample as close to the original collected data as possible is the key challenge. Deciding what constitutes a representative sample depends on the type of information we wish to maintain from the data collected. In this work, we wish to sparsely sample mobility traces of GPS data with the goal to reconstruct the movement of the users both in space and time at the desired granularity. An ideal sample would allow us to reconstruct the traces in such a way that we preserve the frequency of visits and the time spent to the various locations. Therefore, the problem we tackle here is to *sparsely sample the mobility trace of a user with the goal to reconstruct her complete trace in space and time*. This is an on-going work and will be submitted to an international conference in the next months.

6.3. Resource and Traffic Management

Traffic offloading; infrastructure deployment; opportunistic routing; traffic modeling; intermittently connected networks.

6.3.1. *Utility Optimization Approach to Network Cache Design*
Participants: Mostafa Dehghan, Laurent Massoulié, Don Towsley, Daniel Menasche, Y.c. Tay.

In any caching system, the admission and eviction policies determine which contents are added and removed from a cache when a miss occurs. Usually, these policies are devised so as to mitigate staleness and increase the hit probability. Nonetheless, the utility of having a high hit probability can vary across contents. This occurs, for instance, when service level agreements must be met, or if certain contents are more difficult to obtain than others. In this paper, we propose utility-driven caching, where we associate with each content a utility, which is a function of the corresponding content hit probability. We formulate optimization problems where the objectives are to maximize the sum of utilities over all contents. These problems differ according to the stringency of the cache capacity constraint. Our framework enables us to reverse engineer classical replacement policies such as LRU and FIFO, by computing the utility functions that they maximize. We also develop online algorithms that can be used by service providers to implement various caching policies based on arbitrary utility functions.

This work was published and presented at the IEEE Infocom 2016 conference as "A Utility Optimization Approach to Network Cache Design".

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

- Participation to Microsoft Research – Inria Joint Centre, which funds two PhD students (Lennart Gulikers and Remi Varloot).
- During 2016, Cisco and Nordic Semiconductors have funded further development of RIOT and sponsored the RIOT Summit.

7.2. GranData

Participants: Aline Carneiro Viana, Eduardo Mucelli.

Since June 2014, we have a collaboration with GranData (http://grandata.com/), Buenos Aires, Argentina on traffic vs mobility modeling of smartphone users. GranData is a small company that integrates first-party and telco partner data to understand key market trends, to predict customer behavior, and to deliver business results. Its products integrates and analyzes diverse data traces (e.g., telco, social media, or mobile data) to generate behavioral insights and deliver targeted mobile marketing. Part of the thesis of Eduardo Mucelli analysis data traffic using telco traces provided by GranDatas. While this collaboration allow us collaborating with machine learning experts, GranData has the opportunity to get our expertise in mobility analysis.

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. LiveGrid

Participants: Cedric Adjih, Emmanuel Baccelli.

Infine is one of the teams from Inria participating to LiveGrid: LiveGrid is a consortium of the main actors of industry, research organisations, local authorities and competitive cluster from the Paris-Saclay campus. The goal of LiveGrid is make the Paris-Saclay campus one of the leader regions of smart grids. Infine expertise is in infrastructure: testbeds, communication protocols, embedded open source OS.

8.2. National Initiatives

8.2.1. Equipex FIT

Participants: Cedric Adjih, Emmanuel Baccelli, Alexandre Abadie, Philippe Lubrano, Ichrak Amdouni, Alaeddine Weslati, Vincent Ladeveze.

Partners: Inria (Lille, Sophia-Antipolis, Grenoble), INSA, UPMC, Institut Telecom Paris, Institut Télécom Evry, LSIIT Strasbourg.

FIT (Future Internet of Things) aims to develop an experimental facility, a federated and competitive infrastructure with international visibility and a broad panel of customers. It provides this facility with a set of complementary components that enable experimentation on innovative services for academic and industrial users. The project gives french internet stakeholders a means to experiment on mobile wireless communications at the network and application layers thereby accelerating the design of advanced networking technologies for the future internet. FIT was one of 52 winning projects from the first wave of the French Ministry of Higher Education and Research's "Équipements d'Excellence" (Equipex) research grant program, in 2011.

One component of the FIT platform is the sets of IoT-LAB testbeds (see the IoT-LAB web site). These were motivated by the observation that the world is moving towards an "Internet of Things", in which most communication over networks will be between objects rather than people.

The Infine team is more specifically managing the FIT IoT-LAB site formerly at Rocquencourt, which recently moved to Saclay (on-going re-deployment), and is participating in the deployment of an additional IoT-lab testbed in Berlin (at Freie Universitaet Berlin).

The Infine team is actively collaborating with UPEC on wireless sensor network testbeds (and protocols): in 2015 and 2016, the testbed from UPEC with 45 Arduino nodes has been integrated with the FIT IoT-LAB testbed.

8.3. European Initiatives

8.3.1. FP7 & H2020 Projects

8.3.1.1. AGILE (H2020 project)
Participants: Emmanuel Baccelli, Cedric Adjih.

Program: H2020 ICT-30-2015 Topic: Internet of Things and Platforms for Connected Smart Objects

Project acronym: AGILE

Project title: Adoptive Gateways for dIverse muLtiple Environments

Duration: 2015-2017

Coordinator: Emmanuel Baccelli

Other partners: Canonical (UK), Eclipse IoT Foundation (IE), Mobistar (BE), Libelium (ES), Startupbootcamp IoT (SP), CREATE-NET (IT), iMinds (BE), Atos (SP), Rulemotion (UK), Jolocom (DE), Passau University (DE), Sky-Watch (DN), BioAssist (GR), Graz Technical University (AT), Eurotech (IT), IoTango (US).

Abstract:

The AGILE project is a 3-year H2020 project started in January 2016, which will deliver an integrated framework of open source tools and platforms that interoperate for enabling the delivery of adaptive, self-configurable and secure IoT elements (both software and hardware) that can be utilized in a variety of scenarios. Such tools target actors with heterogeneous skills, including entrepreneurs, researchers, and individuals, aiming to enable the realization of IoT applications respecting user privacy and data ownership.

8.3.1.2. ARMOUR (H2020 project)
Participants: Cedric Adjih, Emmanuel Baccelli, Oliver Hahm.

Program: H2020 ICT-12-2015 Topic: Integrating experiments and facilities in FIRE+

Project acronym: ARMOUR

Project title: Large-Scale Experiments of IoT Security Trust

Duration: 2016-2018

Coordinator: Serge Fdida (UPMC)

Other partners: UPMC (France), Synelixis (Greece), SMA (France), UI (Portugal), JRC (Belgium), EGM (France), OdinS (Spain).

Abstract:

The ARMOUR project is a 2-year H2020 project started in February 2016. The ARMOUR project is aimed at providing duly tested, benchmarked and certified Security & Trust technological solutions for large-scale IoT using upgraded FIRE large scale IoT/Cloud testbeds properly-equipped for Security & Trust experimentations. To this, ARMOUR will: (1) Enhance two outstanding FIRE testbeds (> 2700 nodes; 500 users) with the ARMOUR experimentation toolbox for enabling large-scale IoT Security & Trust experiments; (2) Deliver six properly experimented, suitably validated and duly benchmarked methods and technologies for enabling Security & Trust in the large-scale IoT; and (3) Define a framework to support the design of Secure & Trusted IoT applications as well as establishing a certification scheme for setting confidence on Security & Trust IoT solutions.

8.3.2. Collaborations in European Programs, Except FP7 & H2020

8.3.2.1. EU CHIST-ERA MACACO

Participants: Aline Carneiro Viana, Emmanuel Baccelli, Eduardo Mucelli.

Program: EU CHIST-ERA, topic Context- and Content-Adaptive Communication Networks

Project acronym: MACACO

Project title: Mobile context-Adaptive CAching for COntent-centric networking

Duration: 2013-2016

Coordinator: Aline Carneiro Viana

Other partners: INPT-ENSEEIHT at University of Toulouse, University of Birmingham (UK), SUPSI (Switzerland), CNR (Italy) and Federal University of Minas Gerais (Brazil)

Abstract:

MACACO (Mobile context-Adaptive CAching for COntent-centric networking) is a 3-year CHIST-ERA European Project addressing the topic Context- and Content-Adaptive Communication Networks. It is funded by ANR in France, SNSF in Switzerland, and ESPRC in UK. It focus on data offloading mechanisms that take advantage of context and content information. Our intuition is that if it is possible to extract and forecast the behaviour of mobile network users in the three dimensional space of time, location and interest (i.e. what, when and where users are pulling data from the network), it is possible to derive efficient data offloading protocols. Such protocols would prefetch the identified data and cache it at the network edge at an earlier time, preferably when the mobile network is less charged, or offers better quality of service. This project has officially started in November 2013.

8.4. International Initiatives

8.4.1. Inria International Partners

8.4.1.1. Declared Inria International Partners

- Renewed IOTPUSH collaboration with Freie Universitaet Berlin around the long-term stay of Emmanuel Baccelli in Berlin, on research topics about the Internet of Things, RIOT and Information-Centric Networking.

- The Inria teams Infine and Eva are part of the "D2D Communication for LTE Advanced Cellular Network" , a project funded by the Indo-French Centre for the Promotion of Advanced Research (CEFIPRA). With industrial partners, and also with Indian partners, this project is focusing on the evolution of cellular networks towards 5G: this includes exploration of device-to-device (D2D) communication, and more generally IoT communication in a cellular context. Research directions include efficient access for IoT devices (massive numbers of devices with low volume communication); combination of random access protocols/error coding/physical layer ; efficient neighbor discovery, ...

8.4.1.2. Informal International Partners

- On-going collaboration with Hamburg University of Applied Science around RIOT.
- Informal collaborations with UIUC and UMass.
- Informal collaborations with ENSI Tunis and Sesame Tunis.

8.4.2. Participation in Other International Programs

8.4.2.1. PHC PESSOA 2015

Participant: Aline Carneiro Viana.

PHC PESSOA 2015 with University of Coimbra (2015-2016).

Program: -FCT - Programa PESSOA

Project title: Routine-based Enhanced Connectivity under User Mobility

Duration: 2015-2016

Coordinator: Aline Carneiro Viana and João Paulo da Silva Machado Garcia Vilela (University of Coimbra)

Abstract: The main goal of this project is to improve WiFi connectivity of users under mobility. The steady growth of smart-phones usage has put cellular networks under great strain, justifying the need for WiFi offloading as a solution that transfers part of the demand on cellular networks to WiFi hotspots that are in many cases already available. However, this must be performed in a way that provides benefits to the cellular operator while ensuring users a similar level of connectivity that they would achieve with cellular networks, even under user mobility (e.g. walking, taking a bus/train, etc). In this work we aim at (1) developing prediction mechanisms for selection of best hotspots by users under mobility, and (2) develop lightweight security schemes to reduce the burden of the association/authentication process of WiFi networks, therefore making WiFi offloading an effective and secure alternative to the growing demand on cellular networks.

8.5. International Research Visitors

8.5.1. Visits to International Teams

8.5.1.1. Research Stays Abroad

Emmanuel Baccelli was Visiting Professor at Freie Universitaet (FU) Berlin in 2016, within the context of the formal collaboration IOTPUSH with this university on research topics about the Internet of Things, RIOT and Information-Centric Networking.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

- In 2016, Emmanuel Baccelli was general chair of the RIOT Summit.

9.1.1.2. Member of the Organizing Committees

- In 2016, Aline Viana was on the conference program committee for IEEE International Conference on Advanced Information Networking and Applications (IEEE AINA-2017).
- In 2016, Emmanuel Baccelli was in the conference programm committee of the ACM Mobile-Health'16.

9.1.2. Scientific Events Selection

9.1.2.1. Reviewer

- In 2016, Cédric Adjih was reviewer for WINCOM'16, and Globecom 2016.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- In 2016, Aline Viana was Associate Editor of ACM Computer Communication Review (ACM CCR) and Editorial Board member of "Wireless Communications and Mobile Computing Open Access Journal" of John Wiley&Sons and Hindawi.

9.1.3.2. Reviewer - Reviewing Activities

- In 2016, Aline Viana was reviewer for IEEE Transaction on Mobile Computing (TMC), Elsevier Pervasive and Mobile Computing, Elsevier AdHoc Networks, Elsevier Computer Networks.
- In 2016, Emmanuel Baccelli was member of the IETF Routing Directorate, and as such was reviewing RFCs in the making.

9.1.4. Invited Talks

- In 2016, Emmanuel Baccelli gave a number of invited talks on the topic of RIOT, including at the IoT Meetup (April 2016 in Athens, Greece), at the RIOT Summit (July 2016 in Berlin, Germany), at TelecomParisTech for ICSSEA'16 (June 2016, Paris, France), at EclipseCon'16 (October 2016 in Ludwigsburg, Germany), at Fraunhofer FOKUS for ASQF Day (December 2016 in Berlin, Germany)
- In 2016, Cédric Adjih gave a talk on RIOT at IFSSTAR (November 2016 in paris, France).
- In 2016, we also gave several invited talks at: Stochastic Networks Conference, UCSD; CIRM workshop on random matrices; Institut Henri Poincaré's "Nexus" of Information and Computation Theories; EPFL workshop for birthdays of Shannon, Urbanke and Telatar (see: http://www.etru50.com/invited-speakers/) ; LINCS scientific advisory board.

9.1.5. Scientific Expertise

- In 2016, Cédric Adjih served as a expert reviewer for one ANR call for proposals.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

- **Masters**
- Engineering school, third year students at Ecole Polytechnique: Laurent Massoulié taught a course on "Networks: distributed control and emerging phenomena".
- Course for Corps des Mines: Emmanuel Baccelli gave a half day course on "Communication Protocols and Operating Systems for the Internet of Things" in the context of the formation PESTO.
- Engineering school, third year students at Telecom ParisTech: Aline Carneiro Viana taught an 1 hour lesson on "Wireless Network" course.
- Engineering school, third year students at Telecom SudParis: Aline Carneiro Viana taught an 6 hour lessons on "Wireless Network" course.

9.2.2. Supervision

- Laurent Massoulie currently advises 2 PhD students: Lennart Gulikers and Remi Varloot.
- Cedric Adjih currently advises Fatma Somaa, on "handling mobility in routing in wireless sensor networks", thesis started in 2013 and co-advised with Inès El Korbi (Institut Supérieur d'Informatique et de Gestion de Kairouan, Tunisie).
- Emmanuel Baccelli and Cédric Adjih are advising: Loïc Dauphin on "Robotics meets the Internet of Things", (June 2016-).
- Aline Viana currently advises 2 PhD students: Roni Shigueta, on "Resource allocation in highly mobile wireless networks" (January 2012-2017). Guangshuo Chen, on "Understanding and predicting human demanded content and mobility" (September 2014-).
- PhD completed: Felipe Domingos on "identifying social attributes in VANETs" (January 2013-June 2016). Advisor: Aline Carneiro Viana.
- PhD completed: Oliver Hahm, "Enabling Energy Efficient Smart Object Networking at Internet-Scale", 2012-2016. Advisor: Emmanuel Baccelli.

9.2.3. Juries

- Laurent Massoulié was on the PhD jury as reviewer for the PhD theses of Anna Benhamou and Alaa Saade, and he presided the PhD thesis committee of Kevin Scaman.
- Aline Viana was on the PhD jury as reviewer of Wenjing Shuai, ”Management of electric vehicle systems with self-interested actors” Telecom Bretagne, September 2016.
- Aline Viana was on the PhD jury as examiner of Mouna Rekik, ”Protocols for Smart Girds”, Universite de Lille, France July 2016.
- Aline Viana was on the Master jury of Fausto Silva Moraes ”Explorando Interações em Redes Sociais Online, Comunicação D2D e Estratégias de Cache para Uso Eficiente de Recursos em Redes Celulares”, Brazil November 2016.
- Emmanuel Baccelli was on the PhD jury of Francisco Acosta, "Self-adaptation for Internet of Things applications", Universite de Rennes 1, Dec. 2016.
- Emmanuel Baccelli was on the PhD jury of Kevin Roussel, "Évaluation et amélioration des plates-formes logicielles pour réseaux de capteurs sans-fil, pour optimiser la qualite de service et l'énergie," Université de Lorraine, June 2016.

9.3. Popularization

- Publication of "RIOT, le Systeme d'Exploitation pour l'Internet des Objets" in the magazine Open Silicium, Mars 2016. Co-authors: Alexandre Abadie, Cedric Adjih, Emmanuel Baccelli, Francisco Acosta, Hauke Petersen.
- Francisco Padilla and Loïc Dauphin participated to the "Fête de la Science 2016".

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] F. DOMINGOS DA CUNHA, L. VILLAS, A. BOUKERCHE, G. MAIA, A. CARNEIRO VIANA, R. A. F. MINI, A. A. F. LOUREIRO.*Data Communication in VANETs: Survey, Applications and Challenges*, in "Ad Hoc Networks", December 2016, vol. 44, n⁰ C, p. 90-103 [*DOI : 10.1016/J.ADHOC.2016.02.017*], https://hal.archives-ouvertes.fr/hal-01369972.

[2] K. FRAHM, K. JAFFRES-RUNSER, D. L. SHEPELYANSKY.*Wikipedia mining of hidden links between political leaders*, in "The European Physical Journal B", December 2016, vol. 89, 269 [*DOI : 10.1140/EPJB/E2016-70526-3*], https://hal.archives-ouvertes.fr/hal-01362129.

[3] O. HAHM, E. BACCELLI, H. PETERSEN, N. TSIFTES.*Operating Systems for Low-End Devices in the Internet of Things: a Survey*, in "IEEE Internet of Things Journal", October 2016, vol. 3, n⁰ 5, p. 720-734 [*DOI : 10.1109/JIOT.2015.2505901*], https://hal.inria.fr/hal-01245551.

[4] A. C. KOCHEM VENDRAMIN, A. MUNARETTO, M. R. DELGADO, A. CARNEIRO VIANA, M. FONSECA.*A Social-aware Routing Protocol for Opportunistic Networks*, in "Expert Systems with Applications", July 2016, vol. 34, p. 351–363 [*DOI : 10.1016/J.ESWA.2016.01.039*], https://hal.archives-ouvertes.fr/hal-01369964.

[5] K. THILAKARATHNA, A. CARNEIRO VIANA, A. SENEVIRATNE, H. PETANDER.*Design and Analysis of an Efficient Friend-to-Friend Content Dissemination System*, in "IEEE Transactions on Mobile Computing", December 2016 [*DOI : 10.1109/TMC.2016.2570747*], https://hal.archives-ouvertes.fr/hal-01369954.

[6] T. WATTEYNE, V. HANDZISKI, X. VILAJOSANA, S. DUQUENNOY, O. HAHM, E. BACCELLI, A. WOLISZ.*Industrial Wireless IP-based Cyber Physical Systems*, in "Proceedings of the IEEE", March 2016, vol. 104, nº 5, p. 1025-1038 [*DOI : 10.1109/JPROC.2015.2509186*], https://hal.inria.fr/hal-01282597.

Invited Conferences

[7] G. CHEN, A. CARNEIRO VIANA, C. SARRAUTE.*Towards an Adaptive Completion of Sparse Call Detail Records for Mobility Analysis*, in "Workshop on Data Analytics for Mobile Networking", Kona, United States, March 2017, https://hal.inria.fr/hal-01448822.

International Conferences with Proceedings

[8] I. AMDOUNI, C. ADJIH, N. AITSAADI, P. MUHLETHALER.*Experiments with ODYSSE: Opportunistic Duty cYcle based routing for wirelesS Sensor nEtworks*, in "IEEE LCN 2016: The 41st IEEE Conference on Local Computer Networks (LCN), November 7-10, 2016, Dubai, UAE", Dubai, United Arab Emirates, November 2016, https://hal.inria.fr/hal-01407525.

[9] I. AMDOUNI, C. ADJIH, P. MINET, T. PLESSE.*Delay analysis of STDMA in grid wireless sensor networks*, in "ICMCIS 2016: International Conference on Military Communications and Information Systems ICMCIS (former MCC) Brussels, Belgium, 23rd - 24th May 2016", Brussels, Belgium, May 2016, p. 1 - 8 [*DOI : 10.1109/ICMCIS.2016.7496580*], https://hal.inria.fr/hal-01405540.

[10] M. DEHGHAN, L. MASSOULIÉ, D. TOWSLEY, D. S. MENASCHE, Y. TAY.*A utility optimization approach to network cache design*, in "IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications", San Francisco, United States, April 2016 [*DOI : 10.1109/INFOCOM.2016.7524445*], https://hal.archives-ouvertes.fr/hal-01377841.

[11] F. DOMINGOS DA CUNHA, D. ABREU ALVARENGA, G. MAIA, A. CARNEIRO VIANA, R. A. F. MINI, A. A. F. LOUREIRO.*Exploring Interactions in Vehicular Networks*, in "The 14th ACM International Symposium on Mobility Management and Wireless Access", Malte, Malta, November 2016, https://hal.inria.fr/hal-01369341.

[12] C. GÜNDOGAN, C. ADJIH, O. HAHM, E. BACCELLI.*Let healthy links bloom: Scalable Link Checks in Low-Power Wireless Networks for Smart Health*, in "MobileHealth 2016 : ACM 6th International Workshop on Pervasive Wireless Healthcare", Paderborn, Germany, ACM, July 2016, https://hal.inria.fr/hal-01369693.

[13] O. HAHM, C. ADJIH, E. BACCELLI, T. C. SCHMIDT, M. WÄHLISCH.*Designing Time Slotted Channel Hopping and Information-Centric Networking for IoT*, in "IEEE / IFIP International Conference on New Technologies, Mobility and Security (NTMS'2016)", Larnaca, Cyprus, November 2016, https://hal.inria.fr/hal-01404318.

[14] L. MASSOULIÉ, K. XU.*On the capacity of information processing systems*, in "29th Annual Conference on Learning Theory", New York, United States, June 2016, https://hal.archives-ouvertes.fr/hal-01377837.

[15] *Best Paper*
F. SOMAA, C. ADJIH, I. E. KORBI, L. A. SAIDANE.*A Bayesian Model for Mobility Prediction in Wireless Sensor Networks*, in "5th IFIP International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks (PEMWN 2016)", Paris, France, November 2016, https://hal.inria.fr/hal-01405277.

Conferences without Proceedings

[16] F. J. ACOSTA PADILLA, E. BACCELLI, T. EICHINGER, K. SCHLEISER.*The Future of IoT Software Must be Updated*, in "IAB Workshop on Internet of Things Software Update (IoTSU)", Dublin, Ireland, Internet Architecture Board (IAB), June 2016, https://hal.inria.fr/hal-01369681.

[17] O. HAHM, E. BACCELLI, T. C. SCHMIDT, M. WAEHLISCH, C. ADJIH.*A Named Data Network Approach to Energy Efficiency in IoT*, in "IEEE GLOBECOM Workshop on Named Data Networking for Challenged Communication Environments", Washington DC, United States, IEEE, December 2016, https://hal.inria.fr/hal-01369698.

[18] H. PETERSEN, C. ADJIH, O. HAHM, E. BACCELLI.*Demo: IoT Meets Robotics - First Steps, RIOT Car, and Perspectives*, in "ACM International Conference on Embedded Wireless Systems and Networks (EWSN)", Graz, Austria, February 2016, https://hal.inria.fr/hal-01262638.

Research Reports

[19] I. AMDOUNI, C. ADJIH, N. AITSAADI, P. MUHLETHALER.*ODYSSE: A routing Protocol for Wireless Sensor Networks*, UPEC ; Inria Saclay ; Inria – Centre Paris-Rocquencourt, February 2016, n^o RR-8873, https://hal.inria.fr/hal-01292479.

[20] E. BACCELLI, K. SCHLEISER.*Powering the Internet of Things with RIOT: Why? How? What is RIOT?*, ArXiv, March 2016, https://hal.inria.fr/hal-01369669.

[21] G. CHEN.*Understanding Mobile Data Demand regarding Mobility: The report for mid-term thesis evaluation*, Inria Saclay, June 2016, https://hal.inria.fr/hal-01323916.

[22] G. CHEN, S. HOTEIT, A. C. VIANA, M. FIORE, C. SARRAUTE.*Relevance of Context for the Temporal Completion of Call Detail Record*, Inria Saclay, November 2016, n^o RT-0482, 14, https://hal.inria.fr/hal-01393364.

[23] G. CHEN, S. HOTEIT, A. C. VIANA, M. FIORE, C. SARRAUTE.*Spatio-Temporal Predictability of Cellular Data Traffic*, Inria Saclay - Ile-de-France, November 2016, n^o RT-0483, 17, https://hal.inria.fr/hal-01393361.

[24] O. HAHM, C. ADJIH, E. BACCELLI, T. C. SCHMIDT, M. WÄHLISCH. *A Case for Time Slotted Channel Hopping for ICN in the IoT*, ArXiv, February 2016, https://hal.inria.fr/hal-01369709.

Other Publications

[25] G. CHEN, S. HOTEIT, A. CARNEIRO VIANA, M. FIORE.*On the Quest for Representative Behavioral Datasets: Mobility and Content Demand*, May 2016, CHIST-ERA Projects Seminar 2016, Poster, https://hal.inria.fr/hal-01323917.

[26] O. HAHM, C. ADJIH, E. BACCELLI, T. C. SCHMIDT, M. WAEHLISCH.*ICN over TSCH: Potentials for Link-Layer Adaptation in the IoT*, September 2016, p. 195-196, ACM-ICN '16 Proceedings of the 3rd ACM Conference on Information-Centric Networking , Poster [*DOI : 10.1145/2984356.2985226*], https://hal.inria.fr/hal-01369704.

Project-Team LIFEWARE

Computational systems biology and optimization

Table of contents

Project-Team LIFEWARE

Creation of the Team: 2014 January 01, updated into Project-Team: 2015 April 01

Keywords:

Computer Science and Digital Science:
- 2.1.1. - Semantics of programming languages
- 2.1.5. - Constraint programming
- 2.1.10. - Domain-specific languages
- 2.2.1. - Static analysis
- 2.3.2. - Cyber-physical systems
- 2.4. - Verification, reliability, certification
- 2.4.1. - Analysis
- 2.4.2. - Model-checking
- 2.4.3. - Proofs
- 6. - Modeling, simulation and control
- 6.1. - Mathematical Modeling
- 6.1.1. - Continuous Modeling (PDE, ODE)
- 6.1.2. - Stochastic Modeling (SPDE, SDE)
- 6.1.3. - Discrete Modeling (multi-agent, people centered)
- 6.1.4. - Multiscale modeling
- 6.1.5. - Multiphysics modeling
- 6.2.4. - Statistical methods
- 6.2.6. - Optimization
- 6.3.1. - Inverse problems
- 6.3.4. - Model reduction
- 7.2. - Discrete mathematics, combinatorics
- 7.3. - Optimization
- 7.4. - Logic in Computer Science
- 7.9. - Graph theory
- 8.7. - AI algorithmics

Other Research Topics and Application Domains:
- 1. - Life sciences
- 1.1.2. - Molecular biology
- 1.1.3. - Cellular biology
- 1.1.9. - Bioinformatics
- 1.1.10. - Mathematical biology
- 1.1.11. - Systems biology
- 1.1.12. - Synthetic biology
- 1.4. - Pathologies
- 2. - Health
- 2.2.3. - Cancer
- 2.4.1. - Pharmaco kinetics and dynamics

1. Members

Research Scientists
François Fages [Team leader, Senior researcher, Inria, HDR]
Grégory Batt [Team co-leader, Researcher, Inria, HDR]
Jakob Ruess [Researcher, Inria, from Oct 2016]
Sylvain Soliman [Researcher, Inria, HDR]

PhD Students
Virgile Andreani [Ecole Polytechnique, with Duke Univ., from Sept 2016]
Jean Baptiste Caron [Inria, from Dec 2016]
Jean-Baptiste Lugagne [Inria, with MSC lab (CNRS/Paris7)]
Jonas Senizergues [Inria, until Aug 2016]
Pauline Traynard [Inria, until Feb 2016]

Post-Doctoral Fellows
Chiara Fracassi [Inria, with MSC lab (CNRS/Paris7)]
Artémis Llamosi [CNRS, with MSC lab (CNRS/Paris7), until Aug 2016]
Sucheendra Palaniappan [Inria, with EPI SUMO, from April to June 2016]

Visiting Scientists
David Rosenblueth [Univ. Mexico, from Mar 2016 to Sep 2016]
Denis Thieffry [ENS Paris, HDR]
Pascal Hersen [CNRS, MSC lab (CNRS/Paris7), HDR]
Thierry Martinez [Inria Paris, SED]
Laura Guyot [Dassault Systèmes]

Administrative Assistants
Régine Bricquet [Inria, until June 2016]
Corinne Petitot [Inria, from July 2016]

Others
Virgile Andreani [Inria, Internship, until Aug 2016]
Jean Baptiste Caron [MSC lab (CNRS/Paris7), Internship, from Apr 2016 to June 2016]
Ewen Corre [Inria, Internship, until Mar 2016]
Guillaume Le Guludec [Inria, Internship MPRI, Sup Telecom Paris, from Apr 2016 to Sep 2016]
Sebastian Ramon Sosa Carrillo [Inria, Internship, from Dec 2016]
Nicolas Vasselin [Inria Internship, Ecole Centrale de Paris, from Aug 2016]

2. Overall Objectives

2.1. Overall Objectives

This project aims at developing formal methods and experimental settings for **understanding the cell machinery** and establishing formal paradigms in cell biology. It is based on the vision of **cells as machines**, **biochemical reaction systems as programs**, and on the use of concepts and tools from computer science to master the complexity of cell processes. While for the biologist, as well as for the mathematician, the size of the biological networks and the number of elementary interactions constitute a complexity barrier, for the computer scientist the difficulty is not that much in the size of the networks than in the unconventional

nature of biochemical computation. Unlike most programs, biochemical reaction systems involve transitions that are stochastic rather than deterministic, continuous-time rather than discrete-time, poorly localized in compartments instead of well-structured in modules, and created by evolution instead of by rational design. It is our belief however that some form of modularity is required by an evolutionary system to survive, and that the elucidation of these modules in biochemical computation is now a key to apply engineering methods in cell biology on a large scale.

Concretely, we keep developing a theory of biochemical computation with a prototype implementation in the Biochemical Abstract Machine BIOCHAM, a modeling and analysis platform for Systems Biology. The reaction rule-based language used in this system allows us to reason about biochemical reaction networks at different levels of abstraction, in either the **stochastic, differential, discrete, logical or hybrid semantics** of the reactions. This makes it possible to apply a variety of **static analysis** methods, before going to simulations and **dynamical analyses**, for which we use **temporal logics** as a specification language to formalize biological behaviours with imprecise data, validate models w.r.t. observations, constrain model building, and calibrate models in high dimension by optimization methods

A **tight integration between dry lab and wet lab** efforts is also essential for the success of the project. In collaboration with biologists, we investigate concrete biological questions and develop computational models fitted to quantitative data which allow us to make quantitative predictions. In collaboration with Pascal Hersen, MSC lab, we contribute to the development of an experimental platform for the closed-loop control of intracellular processes. This platform combines hardware (microfluidic device and microscope), software (cell tracking and model-based predictive control algorithms) and genetically modified living cells. It is used to investigate the possibilities to externalize the control of intracellular processes for systems and synthetic biology applications, and perform accurate observations, modifications and real-time control at both single cell and cell population levels. We are affiliated with the Doctorate Schools "Frontières du vivant (FdV)" of University Sorbonne Paris Cité and "Sciences et technologies de l'information et de la communication (STIC)" of University Paris-Saclay.

This project addresses fundamental research issues in computer science on the interplay between **structure and dynamics** in large interaction networks, and on the mixing of **continuous and discrete computation**. Many static analysis problems of biological networks are NP-hard. The recourse to constraint logic programming (CLP) to model and solve them, is our secret weapon, which probably explains our capability to experiment ideas in computational systems biology in very short time, by implementing them in CLP, integrating them as new components in our modeling platform BIOCHAM, and evaluating them directly on a large scale in systems biology model repositories such as BIOMODELS.NET.

3. Research Program

3.1. Computational Systems Biology

Bridging the gap between the complexity of biological systems and our capacity to model and **quantitatively predict system behaviors** is a central challenge in systems biology. We believe that a deeper understanding of the concept and theory of biochemical computation is necessary to tackle that challenge. Progress in the theory is necessary for scaling, and enabling the application of static analysis, module identification and decomposition, model reductions, parameter search, and model inference methods to large biochemical reaction systems. A measure of success on this route will be the production of better computational modeling tools for elucidating the complex dynamics of natural biological processes, designing synthetic biological circuits and biosensors, developing novel therapy strategies, and optimizing patient-tailored therapeutics.

Progress on the **coupling of models to data** is also necessary. Our approach based on quantitative temporal logics provides a powerful framework for formalizing experimental observations and using them as formal specification in model building. Key to success is a tight integration between *in vivo* and *in silico* work, and on the mixing of dry and wet experiments, enabled by novel biotechnologies. In particular, the use of microfluidic devices makes it possible to measure behaviors at both single-cell and cell population levels *in vivo*, provided innovative modeling, analysis and control methods are deployed *in silico*.

In synthetic biology, while the construction of simple intracellular circuits has shown feasible, the design of larger, **multicellular systems** is a major open issue. In engineered tissues for example, the behavior results from the subtle interplay between intracellular processes (signal transduction, gene expression) and intercellular processes (contact inhibition, gradient of diffusible molecule), and the question is how should cells be genetically modified such that the desired behavior robustly emerges from cell interactions.

3.2. Modeling of Phenotypic Heterogeneity in Cellular Processes

Since nearly two decades, a significant interest has grown for getting a quantitative understanding of the functioning of biological systems at the cellular level. Given their complexity, proposing a model accounting for the observed cell responses, or better, predicting novel behaviors, is now regarded as an essential step to validate a proposed mechanism in systems biology. Moreover, the constant improvement of stimulation and observation tools creates a strong push for the development of methods that provide predictions that are increasingly precise (single cell precision) and robust (complex stimulation profiles).

It is now fully apparent that cells do not respond identically to a same stimulation, even when they are all genetically-identical. This phenotypic heterogeneity plays a significant role in a number of problems ranging from cell resistance to anticancer drug treatments to stress adaptation and bet hedging.

Dedicated modeling frameworks, notably **stochastic** modeling frameworks, such as chemical master equations, and **statistic** modeling frameworks, such as ensemble models, are then needed to capture biological variability.

Appropriate mathematical and computational should then be employed for the analysis of these models and their calibration to experimental data. One can notably mention **global optimization** tools to search for appropriate parameters within large spaces, **moment closure** approaches to efficiently approximate stochastic models [0], and (stochastic approximations of) the **expectation maximization** algorithm for the identification of mixed-effects models [0].

3.3. Logical Paradigm for Systems Biology

Our group was among the first ones in 2002 to apply **model-checking** methods to systems biology in order to reason on large molecular interaction networks, such as Kohn's map of the mammalian cell cycle (800 reactions over 500 molecules) [0]. The logical paradigm for systems biology that we have subsequently developed for quantitative models can be summarized by the following identifications :

$$\text{biological model} = \text{transition system } K$$

$$\text{dynamical behavior specification} = \text{temporal logic formula } \phi$$

$$\text{model validation} = \text{model-checking} \quad K, \ s \models? \ \phi$$

$$\text{model reduction} = \text{sub-model-checking} \quad K'? \subset K, \ K', \ s \models \phi$$

$$\text{model prediction} = \text{formula enumeration} \quad K, \ s \models \phi?$$

$$\text{static experiment design} = \text{symbolic model-checking} \quad K, \ s? \models \phi$$

$$\text{model inference} = \text{constraint solving} \quad K?, \ s \models \phi$$

$$\text{dynamic experiment design} = \text{constraint solving} \quad K?, \ s? \models \phi$$

---

[0]Moment-based inference predicts bimodality in transient gene expression, C. Zechner C, J. Ruess, P. Krenn, S. Pelet, M. Peter, J. Lygeros, and H. Koeppl, Proceedings of the National Academy of Sciences USA, 9(5):109(21):8340-5, 2012

[0]What population reveals about individual cell identity: estimation of single-cell models of gene expression in yeast, A. Llamosi, A.M. Gonzalez-Vargas, C. Versari, E. Cinquemani, G. Ferrari-Trecate, P. Hersen, and G. Batt, PLoS Computational Biology, 9(5): e1003056, 2015

[0]N. Chabrier-Rivier, M. Chiaverini, V. Danos, F. Fages, V. Schächter. Modeling and querying biochemical interaction networks. Theoretical Computer Science, 325(1):25–44, 2004.

In particular, the definition of a continuous satisfaction degree for **first-order temporal logic** formulae with constraints over the reals, was the key to generalize this approach to quantitative models, opening up the field of model-checking to model optimization [0] This line of research continues with the development of temporal logic patterns with efficient constraint solvers and their generalization to handle stochastic effects.

3.4. External Control of Cell Processes

External control has been employed since many years to regulate culture growth and other physiological properties. Recently, taking inspiration from developments in synthetic biology, closed loop control has been applied to the regulation of intracellular processes. Such approaches offer unprecedented opportunities to investigate how a cell process dynamical information by maintaining it around specific operating points or driving it out of its standard operating conditions. They can also be used to complement and help the development of synthetic biology through the creation of hybrid systems resulting from the interconnection of in vivo and in silico computing devices.

In collaboration with Pascal Hersen (CNRS MSC lab), we developed a platform for gene expression control that enables to control protein concentrations in yeast cells. This platform integrates microfluidic devices enabling long-term observation and rapid change of the cells environment, microscopy for single cell measurements, and software for real-time signal quantification and model based control. We demonstrated recently that this platform enables controlling the level of a fluorescent protein in cells with unprecedented accuracy and for many cell generations [0].

More recently, motivated by an analogy with a benchmark control problem, the stabilization of an inverted pendulum, we investigated the possibility to balance a genetic toggle switch in the vicinity of its unstable equilibrium configuration. We searched for solutions to balance an individual cell and even an entire population of heterogeneous cells, each harboring a toggle switch.

3.5. Chemical Reaction Network Theory

Feinberg's chemical reaction network theory and Thomas's influence network analyses provide sufficient and/or necessary structural conditions for the existence of multiple steady states and oscillations in regulatory networks, which can be predicted by static analyzers without making any simulation. In this domain, most of our work consists in analyzing the interplay between the **structure** (Petri net properties, influence graph, subgraph epimorphisms) and the **dynamics** (Boolean, CTMC, ODE, time scale separations) of biochemical reaction systems. In particular, our study of influence graphs of reaction systems, our generalization of Thomas' conditions of multi-stationarity and Soulé's proof to reaction systems [0], the inference of reaction systems from ODEs [0], the computation of structural invariants by constraint programming techniques, and the analysis of model reductions by subgraph epimorphisms now provide solid ground for developing static analyzers, using them on a large scale in systems biology, and elucidating modules.

3.6. Mixed Analog-Digital Computation with Biochemical Reactions

The continuous nature of many protein interactions leads us to consider models of analog computation, and in particular, the recent results in the theory of analog computability and complexity obtained by Amaury

---

[0] On a continuous degree of satisfaction of temporal logic formulae with applications to systems biology A. Rizk, G. Batt, F. Fages, S. Soliman International Conference on Computational Methods in Systems Biology, 251-268

[0] Jannis Uhlendorf, Agnés Miermont, Thierry Delaveau, Gilles Charvin, François Fages, Samuel Bottani, Grégory Batt, Pascal Hersen. Long-term model predictive control of gene expression at the population and single-cell levels. Proceedings of the National Academy of Sciences USA, 109(35):14271–14276, 2012.

[0] Sylvain Soliman. A stronger necessary condition for the multistationarity of chemical reaction networks. Bulletin of Mathematical Biology, 75(11):2289–2303, 2013.

[0] François Fages, Steven Gay, Sylvain Soliman. Inferring reaction systems from ordinary differential equations. Journal of Theoretical Computer Science (TCS), Elsevier, 2015, 599, pp.64–78.

Pouly [0] and Olivier Bournez, establish fundamental links with digital computation. In [18], we derive from these results a Turing completeness result for elementary reaction systems (without polymerization) under the differential semantics. The proof of this result shows how mathematical functions described by Ordinary Differential Equations, namely by Polynomial Initial Value Problems (PIVP), can be compiled into elementary biochemical reactions, furthermore with a notion of analog computation complexity defined as the length of the trajectory to reach a given precision on the result. This opens a whole research avenue to analyze natural circuits in Systems Biology, transform behavioural specifications into biochemical reactions for Synthetic Biology, and compare artificial circuits with natural circuits acquired through evolution, from the novel point of view of analog computation complexity.

3.7. Constraint Solving and Optimization

Constraint solving and optimization methods are important in our research [17]. On the one hand, static analysis of biochemical reaction networks involves solving hard combinatorial optimization problems, for which **constraint programming** techniques have shown particularly successful, often beating dedicated algorithms and allowing to solve large instances from model repositories. On the other hand, parameter search and model calibration problems involve similarly solving hard continuous optimization problems, for which **evolutionary algorithms** such as the covariance matrix evolution strategy (CMA-ES) [0] has shown to provide best results in our context, for up to 100 parameters, for building challenging quantitative models, gaining model-based insights, revisiting admitted assumptions, and contributing to biological knowledge [0].

4. Application Domains

4.1. Preamble

Our collaborative work on biological applications is expected to serve as a basis for groundbreaking advances in cell functioning understanding, cell monitoring and control, and novel therapy design and optimization. We work mainly on eukaryotic cells. Our collaborations with biologists are focused on **concrete biological questions**, and on the building of predictive models of biological systems to answer them. However, one important application of our research is the development of a **modeling platform** for systems biology.

4.2. Modeling platform for systems biology

Since 2002, we develop an open-source software environment for modeling and analyzing biochemical reaction systems. This software, called the Biochemical Abstract Machine (BIOCHAM), is compatible with SBML for importing and exporting models from repositories such as BioModels. It can perform a variety of static analyses, specify behaviors in Boolean or quantitative temporal logics, search parameter values satisfying temporal constraints, and make various simulations. While the primary reason of this development effort is to be able to **implement our ideas and experiment them quickly on a large scale**, BIOCHAM is used by other groups either for building models, for comparing techniques, or for teaching (see statistics in software section). BIOCHAM-WEB is a web application which makes it possible to use BIOCHAM without any installation. We plan to continue developing BIOCHAM for these different purposes and improve the software quality.

---

[0]Amaury Pouly, "Continuous models of computation: from computability to complexity", PhD Thesis, Ecole Polytechnique, Nov. 2015.
[0]N. Hansen, A. Ostermeier (2001). Completely derandomized self-adaptation in evolution strategies. Evolutionary Computation, 9(2) pp. 159–195.
[0]Domitille Heitzler, Guillaume Durand, Nathalie Gallay, Aurélien Rizk, Seungkirl Ahn, Jihee Kim, Jonathan D. Violin, Laurence Dupuy, Christophe Gauthier, Vincent Piketty, Pascale Crépieux, Anne Poupon, Frédérique Clément, François Fages, Robert J. Lefkowitz, Eric Reiter. Competing G protein-coupled receptor kinases balance G protein and β-arrestin signaling. Molecular Systems Biology, 8(590), 2012.

4.3. Couplings between the cell cycle and the circadian clock

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. These questions are at the heart of our collaboration with Franck Delaunay (CNRS Nice) and Francis Lévi (Univ. Warwick, GB, formerly INSERM Hopital Paul Brousse, Villejuif) and of our participation in the ANR Hyclock project and in the submitted EU H2020 C2SyM proposal, following the former EU EraNet Sysbio C5Sys and FP6 TEMPO projects. In the past, we developed a coupled model of the Cell Cycle, Circadian Clock, DNA Repair System, Irinotecan Metabolism and Exposure Control under Temporal Logic Constraints [0]. We now focus on the bidirectional coupling between the cell cycle and the circadian clock and expect to gain fundamental insights on this complex coupling from computational modeling and single-cell experiments.

4.4. Biosensor design and implementation in non-living protocells

In collaboration with Franck Molina (CNRS, Sys2Diag, Montpellier) and Jie-Hong Jiang (NTU, Taiwan) we ambition to apply our techniques to the design and implementation of biosensors in non-living vesicles for medical applications. Our approach is based on purely protein computation and on our ability to compile controllers and programs in biochemical reactions. The realization will be prototyped using a microfluidic device at CNRS Sys2Diag which will allow us to precisely control the size of the vesicles and the concentrations of the injected proteins. It is worth noting that the choice of non-living chassis, in contrast to living cells in synthetic biology, is particularly appealing for security considerations and compliance to forthcoming EU regulation.

5. Highlights of the Year

5.1. Highlights of the Year

Creation of the Exploratory Action *InBio* with Pasteur Institute in Paris

The InBio project has been selected in the context of a call for new research groups organized by the Center for Bioinformatics, Biostatistics and Integrative Biology (C3BI) of Institut Pasteur.

The main scientific question investigated in InBio is how one can exploit cell-to-cell differences to better learn and control the functioning of biological systems. That is, instead of seeing phenotypic variability as undesired noise that beclouds the processes of interest, we will try to harness cellular heterogeneity. In particular for control problems, because one interacts with the system, it is important to be able to predict the dynamical evolution of phenotypic heterogeneity.

A second important scientific objective of InBio is to develop more rational and systematic interactions between experimental and computational work. The virtuous loop in which experiments nurture models, that in turn, orient further experiments is universally acclaimed and installing such a loop is a central objective of many research projects. In interdisciplinary research, it is expected that this exchange of information will emerge from the interactions between the two disciplinary groups. For both practical and theoretical reasons, this is actually often not the case. In InBio, we will adopt a multidisciplinary research approach and develop an integrated environment around the design-and-test loop. This will notably involve the rational design of cell stains and of experimental plans, so that experiments are maximally-informative, and of efficient model calibration and discrimination methods. This specific focus explains the full name given to the InBio group: "Experimental and Computational Methods for Modeling Cellular Processes" (InBio simply abridges integrative biology).

---

[0] Elisabetta De Maria, François Fages, Aurélien Rizk, Sylvain Soliman. Design, Optimization, and Predictions of a Coupled Model of the Cell Cycle, Circadian Clock, DNA Repair System, Irinotecan Metabolism and Exposure Control under Temporal Logic Constraints. Theoretical Computer Science, 412(21):2108 2127, 2011.

InBio will be hosted at Institut Pasteur and will host experimental and theoretical research. It is a mixed structure between Inria (action exploratoire attached to Lifeware) and Institut Pasteur (research unit attached to the C3BI), and is headed by Grégory Batt.

The Dogma of the Control of the Cell Cycle by the Circadian Clock Revisited

Our long-standing and tight collaboration with Franck Delaunay's lab in Nice culminated this year with a revisiting of the dogma of the control of the cell cycle by the circadian clock. In [9] we showed, using a coupled reaction model of the cell cycle and the circadian clock and BIOCHAM analysers [4], that a selective upregulation of Reverb-α (or an inhibition of Bmal1) during mitosis is necessary to explain the period and phase data observed in NIH3T3 fibroblasts in different serum concentrations. This mechanism constitutes a reverse control of the circadian clock by the cell divisions which was previously outlooked but is overriding in some spontaneously dividing cell types such as non-confluent NIH3T3 fibroblasts.

Figure 1.

Céline Feillet and Franck Delaunay, CNRS Nice,

with the large-scale time-lapse video microscope which produced the unicellular 72h data studied in [9].

6. New Software and Platforms

6.1. BIOCHAM

The Biochemical Abstract Machine
KEYWORDS: Systems Biology - Bioinformatics
FUNCTIONAL DESCRIPTION

The Biochemical Abstract Machine (BIOCHAM) is a software environment for modeling and analyzing biochemical reaction systems, performing static analyses, making simulations, specifying behaviors in temporal logic and searching parameter values in high dimension.

This year BIOCHAM has been completely rewritten with a modular architecture. The new version v4.0 will be released soon with new features for synthesizing biochemical reaction systems from input/output function specifications.

- Participants: François Fages, Guillaume Le Guludec, Thierry Martinez Sylvain Soliman
- Contact: François Fages
- URL: http://lifeware.inria.fr/biocham/

6.2. CellStar

KEYWORDS: Systems biology - Bioinformatics
FUNCTIONAL DESCRIPTION

In close collaboration with Kirill Batmanov, Cédric Lhoussaine and Cristian Versari (LIFL, CNRS/Lille Univ), with Szymon Stoma (Inria; now ETHZ), and with Pascal Hersen (MSC, CNRS/Paris7), we developed CellStar, a tool-chain for image processing and analysis dedicated to segmentation and tracking of yeast cells in brightfield time-lapse microscopy movies. To estimate algorithm quality we developed a benchmark made of manually-verified images illustrating various situations. On this benchmark, CellStar outperformed 5 other state-of-the-art tools. The tool-chain is implemented in Matlab and is provided together with the Python Yeast Image Toolkit benchmark tool.

- Participants: Pascal Hersen, Grégory Batt, Artémis Llamosi and Szymon Stoma
- Contact: Grégory Batt
- URL: http://cellstar-algorithm.org/

6.3. CLP2Zinc

KEYWORDS: Modeling language - Constraint programming - Search
FUNCTIONAL DESCRIPTION

CLP2Zinc is a rule-based modeling language for constraint programming. It extends the MiniZinc modeling language with Horn clauses which can be used to express search strategies as constraints in the model. This system was developed in the framework of the ANR Net-WMS-2 project and is a follow-up of the Rules2CP modeling language.

- Participants: Thierry Martinez, François Fages and Sylvain Soliman
- Contact: Thierry Martinez
- URL: http://lifeware.inria.fr/~tmartine/clp2zinc/

7. New Results

7.1. Analog computation in the cell: specifications, compilation into biochemical reactions and computational complexity

Participants: François Fages, Guillaume Le Guludec.

The continuous nature of many protein interactions leads us to consider mixed analog-digital computation methods, for which the recent results in the theory of analog computability and complexity obtained by Amaury Pouly [0] and Olivier Bournez, establish fundamental links with digital computation. In [18], we derive from these results a Turing completeness result for elementary reaction systems (without polymerization) under the differential semantics, and present a compiler of behavioural specifications into biochemical reactions which can be compared to natural circuits acquired through evolution. We illustrate this approach through the example of the MAPK signaling module which has a function of analog-digital converter in the cell, and through the cell cycle control.

The biochemical compiler is implemented in BIOCHAM v4.0 which will be soon released. We plan to use it in the ANR-MOST project BIOPSY on "Biochemical Programming" for the design of artificial biosensors and the programming of (non-living) protocells in collaboration with Franck Molina, CNRS Sys2diag lab, Montpellier.

---

[0] Amaury Pouly, "Continuous models of computation: from computability to complexity", PhD Thesis, Ecole Polytechnique, Nov. 2015.

7.2. Influence systems vs reaction systems

Participants: François Fages, Thierry Martinez, David Rosenblueth, Sylvain Soliman, Denis Thieffry.

In Systems Biology, modelers develop more and more reaction-based models to describe the mechanistic biochemical reactions underlying cell processes. They may also work, however, with a simpler formalism of influence graphs, to merely describe the positive and negative influences between molecular species. The first approach is promoted by reaction model exchange formats such as SBML, and tools like CellDesigner, while the second is supported by other tools that have been historically developed to reason about boolean gene regulatory networks. In practice, modelers often reason with both kinds of formalisms, and may find an influence model useful in the process of building a reaction model. In [11], we introduce a formalism of influence systems with forces, and put it in parallel with reaction systems with kinetics, in order to develop a similar hierarchy of boolean, discrete, stochastic and differential semantics. We show that the expressive power of influence systems is the same as that of reaction systems under the differential semantics, but weaker under the other interpretations, in the sense that some discrete behaviours of reaction systems cannot be expressed by influence systems. This approach leads us to consider a positive boolean semantics which we compare to the asynchronous semantics of gene regulatory networks à la Thomas. We study the monotonicity properties of the positive boolean semantics and derive from them an efficient algorithm to compute attractors.

7.3. What population reveals about individual cell identity: Single-cell parameter estimation of models of gene expression in yeast

Participants: Grégory Batt, Pascal Hersen, Artémis Llamosi.

Significant cell-to-cell heterogeneity is ubiquitously observed in isogenic cell populations. Consequently, parameters of models of intracellular processes, usually fitted to population-averaged data, should rather be fitted to individual cells to obtain a population of models of similar but non-identical individuals. In [6], we propose a quantitative modeling framework that attributes specific parameter values to single cells for a standard model of gene expression. We combine high quality single-cell measurements of the response of yeast cells to repeated hyperosmotic shocks and state-of-the-art statistical inference approaches for mixed-effects models to infer multidimensional parameter distributions describing the population, and then derive specific parameters for individual cells. The analysis of single-cell parameters shows that single-cell identity (e.g. gene expression dynamics, cell size, growth rate, mother-daughter relationships) is, at least partially, captured by the parameter values of gene expression models (e.g. rates of transcription, translation and degradation). Our approach shows how to use the rich information contained into longitudinal single-cell data to infer parameters that can faithfully represent single-cell identity. This is the first demonstration that biologically-meaningful values for parameter of intracellular processes can be attributed to individual cells.

7.4. The cost of cellular adaptation to stress: tradeoff between short-term and long-term adaptation to osmotic stress in yeast

Participants: Grégory Batt, Ewen Corre, Pascal Hersen, Artémis Llamosi.

Upon stress, cells have evolved complex adaptation strategies to environmental variations which include sensing, information processing and modification of metabolic and transcriptional activity. The reaction of yeast cells to hyperosmotic stress spans several timescales and includes massive gene-expression changes, bio-compatible osmolyte production, and direct action on the cell cycle. Despite a detailed knowledge of molecular events, the impact of stress-response on cellular resources is poorly known. In particular, strong and fast adaptation which alter proliferation in the short term while conferring advantage on the long term are important drivers of stress-response evolution.

In this study, we use microfluidics to vary dynamically both the source of cost (osmotic stress) and the available metabolic resources (glucose) while monitoring cellular proliferation. We show that, under hyper-osmotic stress, metabolic resources are rerouted towards the production of glycerol through activation of an essential enzyme in glycerol production. This reveals the nature of the burden imposed by osmotic stress and, more generally, allows us to better understand the evolutionary tradeoffs between stress response and proliferation.

7.5. Balancing a genetic inverted pendulum

Participants: Grégory Batt, Pascal Hersen, Jean-Baptiste Lugagne, Jean-Baptiste Caron.

The ability to routinely control complex genetic circuits in vivo and in real-time promises quantitative understanding of cellular processes of unprecedented precision and quality. With combined efforts in microfluidic design, microscope automation, image analysis, modeling and control theory, we propose a platform for real-time, single-cell, *in silico* control of genetic networks in *E. coli*. The circuit we are trying to control is a genetic toggle switch, a foundational circuit in synthetic biology, which consists of two genes that repress each other. This genetic system features two stable equilibrium points where one of the genes has taken over. Our objective is to dynamically balance the circuit in single cells around a third, unstable equilibrium point at which no gene dominates and their mutual repression strengths are balanced. This is similar to the landmark problem in control theory of stabilizing an inverted pendulum in its upright position. Although our work indicates that this real-time control approach can drive convoluted genetic networks towards states that are inaccessible to traditional genetic perturbations such as knock-outs and promoter induction, the a priori quantitative knowledge of the system required for achieving this control is minimal. We show that even a simple Proportional-Integral controller can maintain in a balanced state the toggle switch in single cells. Finally, we demonstrate that similar results can be obtained by applying periodic inductions, identical to all cells in the population. Given the fact that all cells behave differently, this result was highly unexpected. It can however be understood as an example of dynamic stabilization, analogous to the solution proposed by Kapitza for the inverted pendulum.

These results are presented in the PhD thesis of Jean-Baptiste Lugagne [2].

7.6. A look-ahead simulation algorithm for DBN models of biochemical pathways

Participants: Grégory Batt, Sucheendra Palaniappan.

Dynamic Bayesian Networks (DBNs) have been proposed as an efficient abstraction formalism of biochemical models. They have been shown to approximate well the dynamics of biochemical models, while offering improved efficiency for their analysis. In [14], we compare different representations and simulation schemes on these DBNs, testing their efficiency and accuracy as abstractions of biological pathways. When generating these DBNs, many configurations are never explored by the underlying dynamics of the biological systems. This can be used to obtain sparse representations to store and analyze DBNs in a compact way. On the other hand, when simulating these DBNs, singular configurations may be encountered, that is configurations from where no transition probability is defined. This makes simulation more complex. We initially evaluate two simple strategies for dealing with singularities: first, re-sampling simulations visiting singular configurations; second filling up uniformly these singular transition probabilities. We show that both these approaches are error prone. Next, we propose a new algorithm which samples only those configurations that avoid singularities by using a look-ahead strategy. Experiments show that this approach is the most accurate while having a reasonable run time.

7.7. Logical model specification aided by model-checking techniques: application to the mammalian cell cycle regulation

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

Understanding the temporal behaviour of biological regulatory networks requires the integration of molecular information into a dynamical model. However, the analysis of model dynamics faces a combinatorial explosion as the number of regulatory components and interactions increases. In [8], we use model-checking techniques to verify sophisticated dynamical properties resulting from the model influence structure in the absence of kinetic assumption. We demonstrate the power of this approach by analysing a Boolean influence model of the molecular network controlling mammalian cell cycle. This approach enables a systematic analysis of model properties, the delineation of model limitations, and the assessment of various refinements and extensions based on recent experimental observations. The resulting logical model accounts for the main irreversible transitions between cell cycle phases, the sequential activation of cyclins, and the inhibitory role of Skp2, and further emphasizes the multifunctional role for the cell cycle inhibitor Rb.

7.8. Model-based investigation of the circadian clock and cell cycle coupling in mouse embryonic fibroblasts: prediction of RevErb–α up-regulation during mitosis

Participants: François Fages, Sylvain Soliman, Pauline Traynard.

Experimental observations have put in evidence autonomous self-sustained circadian oscillators in most mammalian cells, and proved the existence of molecular links between the circadian clock and the cell cycle. Some mathematical models have also been built to assess conditions of control of the cell cycle by the circadian clock. However, recent studies in individual NIH3T3 fibroblasts have shown an unexpected acceleration of the circadian clock together with the cell cycle when the culture medium is enriched with growth factors, and the absence of such acceleration in confluent cells. In [9], in order to explain these observations, we study a possible entrainment of the circadian clock by the cell cycle through a regulation of clock genes around the mitosis phase. We develop a computational model and a formal specification of the observed behavior to investigate the conditions of entrainment in period and phase. We show that either the selective activation of RevErb-α or the selective inhibition of Bmal1 transcription during the mitosis phase, allow us to fit the experimental data on both period and phase, while a uniform inhibition of transcription during mitosis seems incompatible with the phase data. We conclude on the arguments favouring the RevErb-α up-regulation hypothesis and on some further predictions of the model.

7.9. Stochastic continuous optimization backend for the constraint modelling language MiniZinc with applications to geometrical placement problems

Participants: François Fages, Thierry Martinez, Sylvain Soliman.

MiniZinc is a solver-independent constraint modeling language which is increasingly used in the constraint programming community. It can be used to compare different solvers which are currently based on either Constraint Programming, Boolean satisfiability, Mixed Integer Linear Programming, and recently Local Search. In [12], [13] we present a stochastic continuous optimization backend for MiniZinc models over real numbers. More specifically, we describe the translation of FlatZinc models into objective functions over the reals, and their use as fitness functions for the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) solver. We illustrate this approach with the declarative modeling and solving of hard geometrical placement problems [16], motivated by packing applications in logistics [10] involving mixed square-curved shapes and complex shapes defined by Bézier curves.

Beyond these applications to packing problem, our real motivation for these developments is the solving of parameter search problems in computational systems biology and its implementation in BIOCHAM.

7.10. Mixture model CMA-ES

Participants: François Fages, Nicolas Vasselin.

In [19], we report on our attempt to improve the CMA-ES global optimization algorithm based on two ideas: first the use of Sobol's quasi-random low discrepancy numbers instead of pseudo-random numbers, second the design of a mixture model extension of CMA-ES (MM-CMA-ES) which, instead of doing restarts with an important loss of information at each restart, evolves a dynamic set of multivariate normal distributions in parallel, using an EM clustering algorithm at each step to decide of population splittings and mergings. On the standard Coco benchmark for evaluating global stochastic optimization methods, the use of Sobol numbers shows a quite uniform improvement by 30% as was already shown by Teytaud last year [0]. On the other hand, MM-CMA-ES does not show speed-up w.r.t. CMA-ES with IBOP restart strategy, even on objective functions with many local minima such as the Rastragin function. The reason is the overhead in the number of evaluation of the objective functions, introduced by the MM strategy, and the very subtle effect of the adaptive step size strategy of CMA-ES to escape from the covering of several local minima by one (large) normal distribution.

---

[0]O. Teytaud. Quasi-random numbers improve the CMA-ES on the BBOB testbed. Artificial Evolution (EA2015), 2015, Lyon, France. Springer Verlag, pp.13

7.11. Metro energy optimization through rescheduling

Participants: François Fages, Thierry Martinez.

The use of regenerative braking is a key factor to reduce the energy consumption of a metro line. In the case where no device can store the energy produced during braking, only the metros that are accelerating at the same time can benefit from it. Maximizing the power transfers between accelerating and braking metros thus provides a simple strategy to benefit from regenerative energy without any other hardware device. In [15], we use a mathematical timetable model to classify various metro energy optimization problems studied in the literature and prove their NP-hardness by polynomial reductions of SAT. We then focus on the problem of minimizing the global energy consumption of a metro timetable by modifying the dwell times in stations. We present a greedy heuristic algorithm which aims at locally synchronizing braking trains along the timetable with accelerating trains in their time neighbourhood, using a non-linear approximation of energy transfers. On a benchmark of the litterature composed of six small size timetables, we show that our greedy heuristics performs better than CPLEX using a MILP formulation of the problem with a linear approximation of the objective function. We also show that it runs ten times faster than a state-of-the-art evolutionary algorithm, called the covariance matrix adaptation evolution strategy (CMA-ES), using the same non-linear objective function on these small size instances. On real data leading to 10000 decision variables on which both MILP and CMA-ES do not provide solutions, our dedicated algorithm computes solutions with a reduction of energy consumption ranging from 5% to 9%.

This work done in 2014 in the Cifre PhD Thesis of David Fournier with General Electric Transportation has received this year the Gold Medal of the *Annual Alstom Contest* "I Nove You" in the "Green Innovation" category.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR Projects

- ANR-MOST BIOPSY (2016-2020) on "Biochemical Programming System", coordinated by F. Molina (CNRS, Sys2diag, Montpellier) and J.H. Jiang (National Taiwan University), with F. Fages.

- ANR MEMIP (2016-2020) on "Mixed-Effects Models of Intracellular Processes", coordinated by G. Batt, with P. Hersen, (CNRS/Paris7), E. Cinquemani (Inria EPI IBIS) and M. Lavielle (Inria/CNRS/Polytechnique, EPI XPOP).

- ANR COGEX (2016-2019) on "Computer Aided Control of Gene Expression" coordinated by P. Hersen (MSC lab, CNRS/Paris7), with G. Batt and G. Truan (LISBP, CNRS/INSA).

- ANR Blanc HYCLOCK (2014-2018) on "Hybrid modeling of time for Circadian Clock Biology and Chronopharmacology", coordinated by F. Delaunay (CNRS, Nice), with F. Lévi (INSERM Paris-Sud), G. Bernot (CNRS I3S, Nice), O. Roux (Ecole Centrale Nantes), F. Fages and S. Soliman.

- ANR Blanc STOCH-MC (2014-2018) on "Stochastic Models: Scalable Model Checking", coordinated by Blaise Genest (Inria Rennes), with Grégory Batt, Wieslaw Zielonka (LIAFA), and Hugo Gimbert (LaBRI).

- ANR Investissement Avenir ICEBERG project (2011-2016) "From population models to model populations", coordinated by Grégory Batt, with Pascal Hersen (MSC lab, Paris Diderot Univ./CNRS), Reiner Veitia (Institut Jacques Monod, Paris Diderot Univ./CNRS), Olivier Gandrillon (BM2A lab, Lyon Univ./CNRS), Cédric Lhoussaine (LIFL/CNRS), and Jean Krivine (PPS lab, Paris Diderot Univ./CNRS).

8.1.2. GENCI Contract

- GENCI (2009-2016) attribution of 300000 computation hours per year on the Jade cluster of 10000 cores of GENCI at CINES, Montpellier. Used for our hardest parameter search problems in BIOCHAM-parallel.

8.2. International Initiatives

8.2.1. *Inria International Partners*

8.2.1.1. Informal International Partners

In the context of the PhD thesis of Virgile Andréani, we initiated a collaboration with the lab of Lingchong You in the Biomedical Engineering department of Duke University (NC, USA).

8.2.2. *Participation in Other International Programs*

- French-German PROCOPE (2015-2017) grant on "Réduction de modèle et analyse de grands réseaux biochimiques par des méthodes stoechiométriques et tropicales", coord. Prof. Andreas Weber, University of Bonn, Germany, and Prof. Ovidiu Radulescu, Univ. Montpellier, France.

8.3. International Research Visitors

8.3.1. *Visits of International Scientists*

Our group received for a sabbatical stay of six months

- Prof. David Rosenblueth, University of Mexico, Mexico.

We also received for short visits:

- Prof. Mark Chaplain, University of St-Andrews, UK,
- Prof. Attila Attila Csikász-Nagy, King's College London,
- Dr. Jakob Ruess, IST Austria,
- Dr. Amaury Pouly, Univ. Oxford, UK,
- Dr. Christoph Zechner, ETH Zurich,
- Prof. Natalio Krasnogor, Newcastle University, UK,

8.3.1.1. Research Stays Abroad

Virgile Andréani visited Lingchong You's lab (Duke U.) for two weeks in March 2016.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. *Scientific Events Selection*

9.1.1.1. Member of the Conference Program Committees

- Grégory Batt was member of the program committees of:
  - FOSBE'16 6th IFAC Conference on Foundations of Systems Biology in Engineering, Laboratory for Systems Theory and Automatic Control, Otto von Guericke University, Magdeburg, Germany, 9th - 12th October 2016.
  - CMSB'16 14th International Conference on Computational Methods in Systems Biology, Computer Laboratory, University of Cambridge, UK, 21st - 23rd September 2016.
  - HSB'16 5th international workshop on Hybrid Systems Biology, Grenoble, October 20-21, 2016.
- François Fages was member of the program committees of:

- CIE'16 Computability in Europe, Paris, 2016.
- CMSB'16 14th International Conference on Computational Methods in Systems Biology, Computer Laboratory, University of Cambridge, UK, 21st - 23rd September 2016.
- CP'16 22nd International Conference on Principles and Practice of Constraint Programming, Toulouse, France, September 5-9, 2016.
- HSB'16 5th international workshop on Hybrid Systems Biology, Grenoble, October 20-21, 2016.
- IJCAI'16 25th International Joint Conference on Artificial Intelligence, New York, 2016.
- SASB'16 The Seventh International Workshop on Static Analysis and Systems Biology, September 8–10, 2016, Edinburgh, UK
- WCB'16 12th International Workshop on Constraint-Based Methods for Bioinformatics, Toulouse, 2016.

- Sylvain Soliman was member of the program committees of:
  - CP'16 22nd International Conference on Principles and Practice of Constraint Programming, Toulouse, France, September 5-9, 2016.
  - WCB'16 12th International Workshop on Constraint-Based Methods for Bioinformatics, Toulouse, 2016.
- Sucheendra K. Palaniappan was member of the program committee of Formal Methods for Biological and Biomedical Systems (FMBBS 2016), Shenzhen, China, December 2016.

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

François Fages is member of

- the Editorial Board of the Computer Science area of the Royal Society Open Science journal since 2014,
- the Editorial Board of the journal RAIRO OR Operations Research since 2004.

9.1.2.2. Reviewer - Reviewing Activities

Beyond their Editorial Board and Program Committee duties,

- Grégory Batt reviewed an article for *ACS Synthetic Biology*
- François Fages reviewed articles for *BMC Systems Biology*, *Fundamenta Informaticae*, *BioSystems*, *PLOS-One* and *Computers and Industrial Engineering*.
- Sylvain Soliman reviewed articles for *BMC Systems Biology*, *BioSystems* and *AMS Math Reviews*.

9.1.3. Invited Talks

- François Fages gave invited talks at
  - Workshop on Symbolic Computation for Biological Systems, "Biochemical Programs and Mixed Analog-Digital Algorithms in the Cell", Univ. Bonn, 22 November 2016
  - Bilille scientific day on Systems Biology, "Biochemical Programs and Mixed Analog-Digital Algorithms in the Cell", Univ. Lille, 15 November 2016,
  - Colloque de Cérisy, Sciences de la vie, science de l'information, "Biochemical Programs and Mixed Analog-Digital Algorithms in the Cell", 20 September 2016,
  - Workshop on Formal Verification of Real-Time Systems, "Continuous Valuations of Temporal Logic Specifications with applications to Parameter Optimization and Robustness Measures", ENSTA Brest, 28 June 2016,
  - Workshop on Verification of Biological Systems, "Hybrid Analog/Digital Computation with Biochemical Reaction Systems", ENS Cachan, 17 May 2016,

- International Worskhop on Entropy and Information, "Digital/Analog Computation in the Cell", Univ Paris-Diderot, 9 May 2016.
- GT TheorBio, "Computational Systems Biology and Optimization", Orsay, 9 Février 2016.

- Grégory Batt gave invited talks at
  - Seminar at Systems Biology group, Clinical Pharmacology, Roche Pharma Research, on "Multi-scale modeling of TRAIL-induced apoptosis", March 2016, Basel
  - Seminar at Control Theory and Systems Biology Lab, Department of Biosystems Science and Engineering, ETHZ, on "Multi-scale modeling of TRAIL-induced apoptosis", March 2016, Basel
  - Journées Inria Cancer, Modeling dynamics of cell-to-cell variability in TRAIL-induced apoptosis explains fractional killing and predicts reversible resistance, March 2016, Paris
  - Third International Workshop on Synthesis of Complex Parameters, "What population reveals about individual cell identity: Single-cell parameter estimation of models of gene expression in yeast", April 2016, Eindhoven
  - Open University: systèmes hybrides et systèmes biologiques, "Predicting long-term effects of apoptosis-inducing drug treatments", May 2016, ENS Cachan, France
  - Journées scientifiques Inria, "What population reveals about individual cell identity: Single-cell parameter estimation of models of gene expression in yeast", June 2016, Rennes
  - Second Conference of the French Research Group on Synthetic Biology, "Balancing a genetic toggle switch by real-time control or periodic stimulations", June 2016, Bordeaux
  - Second Conference of the French Research Group on Symbolic Systems Biology, "Multi-scale modeling of TRAIL-induced apoptosis", July 2016, Lyon
  - Pasteur Quantitative Biology Symposium, "Balancing a genetic toggle switch by real-time control or periodic stimulations", Oct 2016, Paris
  - Bilille scientific day on Systems Biology, "Balancing a genetic toggle switch by real-time control or periodic stimulations", Nov 2016, Lille

- Jakob Ruess gave invited talks at
  - Dracula seminar, "Towards real-time in vivo mathematical biology at the level of single cells", Nov 2016, Lyon

9.1.4. Leadership within the Scientific Community

- Grégory Batt is a member of
  - the IEEE/CSS Technical Committee on Systems Biology,
  - the scientific board of the GDR de Biologie de Synthèse et des Systèmes
  - the GDR de Bioinformatique Moléculaire, in charge of the axis on Biological network modelling, systems biology and synthetic biology
  - co-animator of the French working group on Symbolic Systems Biology GT BIOSS

- François Fages is a member of
  - the Steering Committee of the International Conference on Computational Methods for Systems Biology since 2008,
  - the Scientific Council of the *Doctorate School "Frontières Du Vivant"* at *Center for Research and Interdisciplanirity*, Universities Paris Descartes and Paris Diderot, since 2010,
  - The Scientific Commmittee of the Summer School Ecole Thématique Modélisation Formelle des Réseaux de Régulation Biologique since 2010.

9.1.5. Scientific Expertise

François Fages

- is a member of the jury for the *Prix de thèse Gilles Kahn* of the *Société Informatique de France*, since 2015,
- reviewed one research project for the **Israel Science Foundation**.

Grégory Batt has been a member of the selection committee for Junior research scientists (CR2) at Inria Rennes - Bretagne Atlantique in 2016.

9.1.6. Research Administration

François Fages is member of the "Bureau du Comité des Projets" of Inria Saclay-IdF.

Sylvain Soliman is member of

- the "Commission Scientifique" of Inria Saclay-IdF
- and of the AAP Digiteo/Digicosme Ph.D. grant jury.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master: Grégory Batt (coordinator and teacher: 48h) and Jakob Ruess (24h), *Computational Biology*, M1, Master Approches Interdisciplinaires du Vivant (AIV).

Master: Grégory Batt (6h) and Denis Thieffry (coordinator), *Dynamical Modelling of Cellular Regulatory Networks*, M2, Interdisciplinary Master in Life Science at the Ecole Normale Supérieure, Paris.

Master/PhD: Grégory Batt (co-coordinator 80h, teacher 8h) and Jakob Ruess (8h), *Modeling and engineering of biological systems*, M2/PhD, Institut de Technologie et d'Innovation of Paris Sciences et Lettres (PSL-ITI), Paris.

Master: Chiara Fracassi (48h), *Experimental Methods in Biophysics*, M1, Master Approches Interdisciplinaires du Vivant (AIV).

Master: François Fages (coordinator module 48h, teaching 12h), Grégory Batt (12h), and Denis Thieffry (12h), C2-19 *Computational Methods for Systemic and Synthetic Biology*, Master Parisien de Recherche en Informatique (MPRI), Paris.

Doctorate: François Fages (6h), *Méthodes formelles pour la biologie des systèmes*, Ecole Thématique Modélisation Formelle des Réseaux de Régulation Biologique, Ile de Porquerolles, 6-10 June 2016

Master: Chiara Fracassi, *Dynamics of Living Systems*, 24h, M1, Master Approches Interdisciplinaires du Vivant (AIV).

Master: Thierry Martinez, *Développement logiciel*, 17h, M1, Ecole des Ponts et Chaussée, Champs-sur-Marne.

Master: Sylvain Soliman, C2-35-1 *Constraint Programming*, coordinator and teaching 24h, M2, Master Parisien de Recherche en Informatique (MPRI), Paris.

Master: Pauline Traynard, *Introduction to Linux and Programming with Python and R*, M1, M2, 30h, master IMaLiS du département de biologie de l'ENS,

9.2.2. Supervision

PhD : Pauline Traynard, *Model Building by Temporal Logic Constraint Solving: Investigation of the Coupling between the Cell Cycle and the Circadian Clock*, Université Paris Diderot, Paris (Oct 2012), Dir. François Fages and Denis Thieffry (ENS), 10 May 2016.

PhD : François Bertaux. *Cell-based multi-scale modeling for systems and synthetic biology: from stochastic gene expression in single cells to spatially organized cell populations*. Université Paris - Diderot, Dir. Dirk Drasdo (EPI MAMBA) and Grégory Batt, 15 May 2016.

PhD : Jean-Baptiste Lugagne, Université Paris Diderot, Paris (Oct 2012), Dir. Grégory Batt and Pascal Hersen (CNRS, MSC), 13 Dec 2016.

PhD in progress: Jonas Sénizergues, Université Paris Diderot, Paris (Oct 2015, until August 2016), Dir. François Fages and Sylvain Soliman.

PhD in progress (Sept 2016-): Virgile Andréani, Ecole Polytechnique, Paris , Dir. Grégory Batt and Lingchong You (Duke U.).

PhD in progress (Dec 2016-): Jean-Baptiste Caron, relais thèse Inria, Dir. Grégory Batt.

9.2.3. *Juries*

- HDR: Morgan Magnin, "Contributions à l'élaboration de connaissances qualitatives en bio-informatique", Ecole Centrale de Nantes, *François Fages, Reviewer*, 28 avril 2016.
- Ph.D.: Simona Catozzi, "Retroactivity in Signal Transduction", Univ. Nice Sophia-Antipolis, *François Fages, Examiner*, 15 Dec. 2016.
- Ph.D.: Alexandre Temperville, "Bases creuses en algèbre linéaire exacte et simplification algorithmique de modèles biologiques", Université de Lille, *François Fages, Reviewer*, 11 juillet 2016.
- Ph.D.: Ignacio Salas, "Packing Curved Objects with Interval Methods", Ecole des Mines de Nantes, *François Fages, Reviewer, Chairman of the Jury*, 29 avril 2016.
- Ph.D.: Louis Fippo Fitime, "Modélisation hybride, Analyse et Vérification Quantitative des grands réseaux de régulation biologique", École Centrale de Nantes, *Sylvain Soliman, Examiner*, November 28, 2016.
- Ph.D.: Adel Mezine, "Conduite d'expériences par apprentissage actif pour l'identification de systèmes dynamiques biologiques", Paris-Saclay University, *Grégory Batt, Reviewer*, October 11, 2016.

9.3. Popularization

François Fages

- wrote a book chapter "AI and Biological Modeling" [17], for an encyclopedic book entitled "A guided tour to Artificial Intelligence Research" to appear next year,
- participated in the **Colloque de Cérisy** "Sciences du vivant, sciences de l'information", 20 Septembre 2016, with a conference and an article [18] to appear in a book next year,
- wrote an article for ERCIM news [16],
- and received a schoolgirl for one afternoon in our research team.

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] F. BERTAUX.*Cell-based multi-scale modeling for systems and synthetic biology: from stochastic gene expression in single cells to spatially organized cell populations*, Université Pierre & Marie Curie - Paris 6, June 2016, https://tel.archives-ouvertes.fr/tel-01405430.

[2] J.-B. LUGAGNE.*Real-time control of a genetic toggle switch*, Université Paris 7, December 2016, https://tel.archives-ouvertes.fr/tel-01417700.

[3] S. SOLIMAN.*A structural perspective on the dynamics of biochemical systems*, Université Paris Sud - Orsay, December 2016, Habilitation à diriger des recherches, https://tel.archives-ouvertes.fr/tel-01403712.

[4] P. TRAYNARD.*Model Building by Temporal Logic Constraint Solving: Investigation of the Coupling between the Cell Cycle and the Circadian Clock*, Université Paris Diderot, May 2016, https://tel.archives-ouvertes.fr/tel-01404060.

Articles in International Peer-Reviewed Journal

[5] O. BÉTHOUX, A. LLAMOSI, S. TOUSSAINT.*Reinvestigation of <i>Protelytron permianum</i> (Insecta; Early Permian; USA) as an example for applying reflectance transformation imaging to insect imprint fossils*, in "Fossil Record", 2016, vol. 20, n° 1, p. 1 - 7 [*DOI :* 10.5194/FR-20-1-2016], https://hal.inria.fr/hal-01406082.

[6] A. LLAMOSI, A. GONZALEZ, C. VERSARI, E. CINQUEMANI, G. FERRARI-TRECATE, P. HERSEN, G. BATT.*What population reveals about individual cell identity: Single-cell parameter estimation of models of gene expression in yeast*, in "PLoS Computational Biology", February 2016, vol. 12, n° 2, e1004706 [*DOI :* 10.1371/JOURNAL.PCBI.1004706], https://hal.inria.fr/hal-01248298.

[7] F. NABLI, T. MARTINEZ, F. FAGES, S. SOLIMAN.*On Enumerating Minimal Siphons in Petri nets using CLP and SAT solvers: Theoretical and Practical Complexity*, in "Constraints", April 2016, vol. 21, n° 2, p. 251–276 [*DOI :* 10.1007/s10601-015-9190-1], https://hal.archives-ouvertes.fr/hal-01170962.

[8] P. TRAYNARD, A. FAURÉ, F. FAGES, D. THIEFFRY.*Logical model specification aided by model- checking techniques: application to the mammalian cell cycle regulation*, in "Bioinformatics", 2016, vol. 32, n° 17, p. i772-i780 [*DOI :* 10.1093/BIOINFORMATICS/BTW457], https://hal.archives-ouvertes.fr/hal-01378465.

[9] P. TRAYNARD, C. FEILLET, S. SOLIMAN, F. DELAUNAY, F. FAGES.*Model-based investigation of the circadian clock and cell cycle coupling in mouse embryonic fibroblasts: Prediction of RevErb-α up-regulation during mitosis*, in "BioSystems", November 2016, vol. 149, p. 59–69 [*DOI :* 10.1016/J.BIOSYSTEMS.2016.07.003], https://hal.archives-ouvertes.fr/hal-01378466.

International Conferences with Proceedings

[10] A. AGGOUN, A. RHIAT, F. FAGES.*Panorama of real-life applications in logistics embedding bin packing optimization algorithms, robotics and cloud computing technologies*, in "GOL'16 - 3rd IEEE International Conference on Logistics Operations Management", Morocco, Morocco, May 2016 [*DOI :* 10.1109/GOL.2016.7731693], https://hal.archives-ouvertes.fr/hal-01378469.

[11] F. FAGES, T. MARTINEZ, D. ROSENBLUETH, S. SOLIMAN.*Influence Systems vs Reaction Systems*, in "CMSB'16: Proceedings of the fourteenth international conference on Computational Methods in Systems Biology", Cambridge, UK, United Kingdom, E. BARTOCCI, P. LIO, N. PAOLETTI (editors), Lecture Notes in Computer Science, Springer-Verlag, September 2016, vol. 9859, p. 98-115 [*DOI :* 10.1007/978-3-319-45177-0_7], https://hal.archives-ouvertes.fr/hal-01378470.

[12] T. MARTINEZ, F. FAGES, A. AGGOUN.*A Stochastic Continuous Optimization Backend for MiniZinc with Applications to Geometrical Placement Problems*, in "Proceedings of the 13th International Confer-

ence on Integration of Artificial Intelligence and Operations Research Techniques in Constraint Programming, CPAIOR'16", Banff, Canada, C.-G. QUIMPER (editor), Springer, May 2016, vol. 9676, p. 262-278 [*DOI :* 10.1007/978-3-319-33954-2_19], https://hal.archives-ouvertes.fr/hal-01378468.

[13] T. MARTINEZ, F. FAGES, A. AGGOUN.*Un solveur d'optmization stochastique pour MiniZinc avec des applications à des problèmes de placements géométriques*, in "Douzièmes Journées Francophones de Programmation par Contraintes JFPC'16", 2016, https://hal.archives-ouvertes.fr/hal-01378467.

[14] S. K. PALANIAPPAN, M. PICHENÉ, G. BATT, E. FABRE, B. GENEST.*A Look-Ahead Simulation Algorithm for DBN Models of Biochemical Pathways*, in "Hybrid Systems Biology, 5th International Workshop, HSB 2016", Grenoble, France, E. CINQUEMANI, A. DONZÉ (editors), Lecture Notes in Computer Science, October 2016, vol. 9957, p. 3-15 [*DOI :* 10.1007/978-3-319-47151-8_1], https://hal.inria.fr/hal-01406115.

Research Reports

[15] D. FOURNIER, T. MARTINEZ, F. FAGES, D. MULARD.*Metro Energy Optimization through Rescheduling: Mathematical Models and Heuristic Algorithm Compared to MILP and CMA-ES* , Inria Saclay Ile de France, December 2016, https://hal.inria.fr/hal-01420311.

Scientific Popularization

[16] A. AGGOUN, N. BELDICEANU, G. CHABERT, F. FAGES.*Packing with Complex Shapes*, in "ERCIM News", 2016, n° 105, 32, https://hal.archives-ouvertes.fr/hal-01378471.

Other Publications

[17] F. FAGES.*AI in Biological Modelling*, November 2016, working paper or preprint, https://hal.inria.fr/hal-01409753.

[18] F. FAGES, G. LE GULUDEC.*Programmes biochimiques et algorithmes mixtes analogiques-digitaux dans la cellule*, October 2016, working paper or preprint, https://hal.inria.fr/hal-01409743.

[19] N. VASSELIN, F. FAGES.*Mixture Model-CMA-ES* , December 2016, internship report, https://hal.inria.fr/hal-01420342.

Project-Team M3DISIM

Mathematical and Mechanical Modeling with Data Interaction in Simulations for Medicine

IN COLLABORATION WITH: Laboratoire de Mécanique des Solides

IN PARTNERSHIP WITH:
Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Modeling and Control for Life Sciences

Table of contents

Project-Team M3DISIM

Creation of the Team: 2013 January 01, updated into Project-Team: 2016 June 01

Keywords:

Computer Science and Digital Science:

  6.1.1. - Continuous Modeling (PDE, ODE)
  6.1.2. - Stochastic Modeling (SPDE, SDE)
  6.1.4. - Multiscale modeling
  6.1.5. - Multiphysics modeling
  6.2.1. - Numerical analysis of PDE and ODE
  6.3.1. - Inverse problems
  6.3.2. - Data assimilation
  6.3.4. - Model reduction
  6.4. - Automatic control
  6.4.1. - Deterministic control
  6.4.2. - Stochastic control
  6.4.3. - Observability and Controlability
  6.4.4. - Stability and Stabilization

Other Research Topics and Application Domains:

  1.1.10. - Mathematical biology
  2.2.1. - Cardiovascular and respiratory diseases
  2.6.2. - Cardiac imaging
  2.6.3. - Biological Imaging

1. Members

Research Scientists
  Dominique Chapelle [Team leader, Inria, Senior Researcher, HDR]
  Radomir Chabiniok [Inria, Starting Research position]
  Sébastien Imperiale [Inria, Researcher]
  Philippe Moireau [Inria, Researcher]
  Fabrice Vallée [AP/HP, Researcher]

Faculty Members
  Jean-Marc Allain [Ecole Polytechnique, Associate Professor]
  Martin Genet [Ecole Polytechnique, Associate Professor]
  Patrick Le Tallec [Ecole Polytechnique, Professor, Part-time 50%, HDR]

Technical Staff
  Gautier Bureau [Inria]
  Sébastien Gilles [Inria]

PhD Students
  Jorge Albella Martinez [Univ. Santiago de Compostella, from Aug 2016]
  Bruno Burtschell [Inria, until Sep 2016]
  Federica Caforio [Inria]
  Ustim Khristenko [Ecole Polytechnique]

François Kimmig [Ecole Polytechnique, from Sep 2016]
Arthur Le Gall [AP/HP]
Florent Wijanto [Ecole Polytechnique]

Post-Doctoral Fellows
Jean-Sébastien Affagard [Ecole Polytechnique]
Antoine Tonnoir [Inria until Aug 2016]

Administrative Assistant
Hélèna Kutniak [Inria]

Others
Atte Aalto [Post-Doctoral Fellow and Research scientist visitor, until Nov 2016]
Matthieu Caruel [Univ. Paris XII, Research Scientist]
Jona Joachim [Univ. Paris V, Intern, from Feb 2016 until Sep 2016]
Alexandre Laurin [Inria, Post-Doctoral Fellow and Engineers]
Maeva Lopez Poncelas [Ecole Polytechnique, Intern, from Feb 2016 until Sep 2016]

2. Overall Objectives

2.1. Overall Objectives

The research carried out in the M3DISIM team has a rather global methodological perspective oriented towards biomechanics, encompassing mathematical modeling and analysis, inverse problems arising from model-data coupling, and the formulation and analysis of effective and reliable numerical procedures adapted to this overall program. We are also very keen on demonstrating the effectiveness and relevance of these methods in actual applications, usually by proof-of-concept studies carried out within various collaborations.

3. Research Program

3.1. Multi-scale modeling and coupling mechanisms for biomechanical systems, with mathematical and numerical analysis

Over the past decade, we have laid out the foundations of a multi-scale 3D model of the cardiac mechanical contraction responding to electrical activation. Several collaborations have been crucial in this enterprise, see below references. By integrating this formulation with adapted numerical methods, we are now able to represent the whole organ behavior in interaction with the blood during complete heart beats. This subject was our first achievement to combine a deep understanding of the underlying physics and physiology and our constant concern of proposing well-posed mathematical formulations and adequate numerical discretizations. In fact, we have shown that our model satisfies the essential thermo-mechanical laws, and in particular the energy balance, and proposed compatible numerical schemes that – in consequence – can be rigorously analyzed, see [5]. In the same spirit, we have recently formulated a poromechanical model adapted to the blood perfusion in the heart, hence precisely taking into account the large deformation of the mechanical medium, the fluid inertia and moving domain, and so that the energy balance between fluid and solid is fulfilled from the model construction to its discretization, see [6].

3.2. Inverse problems with actual data – Fundamental formulation, mathematical analysis and applications

A major challenge in the context of biomechanical modeling – and more generally in modeling for life sciences – lies in using the large amount of data available on the system to circumvent the lack of absolute modeling ground truth, since every system considered is in fact patient-specific, with possibly non-standard conditions associated with a disease. We have already developed original strategies for solving this particular type of inverse problems by adopting the observer stand-point. The idea we proposed consists in incorporating to the classical discretization of the mechanical system an estimator filter that can use the data to improve the quality of the global approximation, and concurrently identify some uncertain parameters possibly related to a diseased state of the patient, see [7], [8], [9]. Therefore, our strategy leads to a coupled model-data system solved similarly to a usual PDE-based model, with a computational cost directly comparable to classical Galerkin approximations. We have already worked on the formulation, the mathematical and numerical analysis of the resulting system – see [3] – and the demonstration of the capabilities of this approach in the context of identification of constitutive parameters for a heart model with real data, including medical imaging, see [1].

4. Application Domains

4.1. Clinical applications

After several validation steps – based on clinical and experimental data – we have reached the point of having validated the heart model in a pre-clinical context where we have combined direct and inverse modeling in order to bring predictive answers on specific patient states. For example, we have demonstrated the predictive ability of our model to set up pacemaker devices for a specific patient in cardiac resynchronization therapies, see [10]. We have also used our parametric estimation procedure to provide a quantitative characterization of an infarct in a clinical experiment performed with pigs, see [1].

5. Highlights of the Year

5.1. Highlights of the Year

- Official launch of M3DISIM as an Inria project-team (joint with Ecole Polytechnique / LMS) on June 1st
- Habilitation (HDR) of Philippe Moreau on November 28th

6. New Software and Platforms

6.1. FELiScE-branch HappyHeart

Finite Elements for Life SCiences and Engineering problems
KEYWORDS: Finite element modelling - Cardiac Electrophysiology - Cardiovascular and respiratory systems
SCIENTIFIC DESCRIPTION

FELISCE – standing for "Finite Elements for LIfe SCiences and Engineering" – is a new finite element code. One specific objective of this code is to provide in a unified software environment all the state-of-the-art tools needed to perform simulations of the complex cardiovascular models considered in the teams M3DISIM and REO – namely, involving fluid and solid mechanics, electrophysiology, and the various associated coupling phenomena.

FUNCTIONAL DESCRIPTION

In FELISCE we have prepared a branch called HappyHeart, which aims at providing a user-friendly interface able to deal efficiently with complex cardiovascular simulations. Started in 2013, the code is already quite large (about 55,000 lines of code in almost 700 different files) and its core is about to be complete. It includes among others full HPC functionalities, high-order finite elements, physics coupling and topology capabilities. Our purpose will then be to use the library to implement the sophisticated cardiovascular models of the team and couple them with Verdandi (data assimilation library) to provide patient-specific simulations.

- Participants: Gautier Bureau, Federica Caforio, Dominique Chapelle, Sébastien Gilles, Sébastien Imperiale, Philippe Moireau,

- Contact: Sébastien Gilles

- URL: http://felisce.gforge.inria.fr

6.2. HeartLab

KEYWORDS: Simulation - Health - Cardiac - Image analysis - Computational geometry
SCIENTIFIC DESCRIPTION

The heartLab software is a library designed to perform both simulation and estimation of the heart mechanical behavior (based on various types of measurements, e.g. images).

FUNCTIONAL DESCRIPTION

The heartLab software is a library designed to perform both simulation and estimation of the heart mechanical behavior (based on various types of measurements, e.g. images).

Also included are geometric data and tools in the code to define cardiac anatomical models compatible with the simulation requirements in terms of mesh quality, fiber direction data defined within each element, and the referencing necessary for handling boundary conditions and estimation, in particular. These geometries are analytical or come from computerized tomography (CT) or magnetic resonance (MR) image data of humans or animals.

- Participants: Gautier Bureau, Radomir Chabiniok, Dominique Chapelle and Philippe Moireau

- Contact: Philippe Moireau

- URL: https://raweb.inria.fr/rapportsactivite/RA2013/m3disim/uid14.html

6.3. Verdandi

KEYWORDS: HPC - Model - Software Components - Partial differential equation
FUNCTIONAL DESCRIPTION

Verdandi is a free and open-source (LGPL) library for data assimilation. It includes various such methods for coupling one or several numerical models and observational data. Mainly targeted at large systems arising from the discretization of partial differential equations, the library is devised as generic, which allows for applications in a wide range of problems (biology and medicine, environment, image processing, etc.). Verdandi also includes tools to ease the application of data assimilation, in particular in the management of observations or for a priori uncertainty quantification. Implemented in C++, the library may be used with models implemented in Fortran, C, C++ or Python.

- Participants: Gautier Bureau, Dominique Chapelle, Philippe Moireau

- Contact: Philippe Moireau

- URL: http://verdandi.gforge.inria.fr/

7. New Results

7.1. Mathematical and Mechanical modeling

7.1.1. *A 3D contact-mechanics model of the heart and thorax for seismocardiography*

Participants: Alexandre Laurin [correspondant], Sébastien Imperiale, Dominique Chapelle, Philippe Moireau.

The current interpretation of seismocardiogram fiducial points depends on their phenomenological association with the timing of events on simultaneous echocardiograms. Signal processing methods can be devised to acquire these timings automatically (see [21] and [22]). So far, no causal framework has been tested to explain this timing, nor their direction and amplitude. This work attempted to adapt a comprehensive 3D cardiac model to interact through contact with a model of the thoracic cage. The heart model was designed to represent multi-scale, multi-physics physiological processes such as the electrical activation, the mechanical contraction, as well as the system circulation. The objective is to link observed acceleration of the sternum to the underlying physiology, and offer a potential mechanical explanation for them. The modelling chain necessary to go from the heart model to a simulated SCG has been successfully implemented (see Figure 1). Furthermore, the complexity of the thoracic model has been substantially reduced, without deteriorating results, to improve the portability of the entire process. Once the relevant parameters of in-vivo thoraces will have been precisely identified, it will be possible to compute heart forces and the various cardiac events that cause them directly from SCG measurements. The subsequent aim is to apply the model to ageing and pathological physiologies.
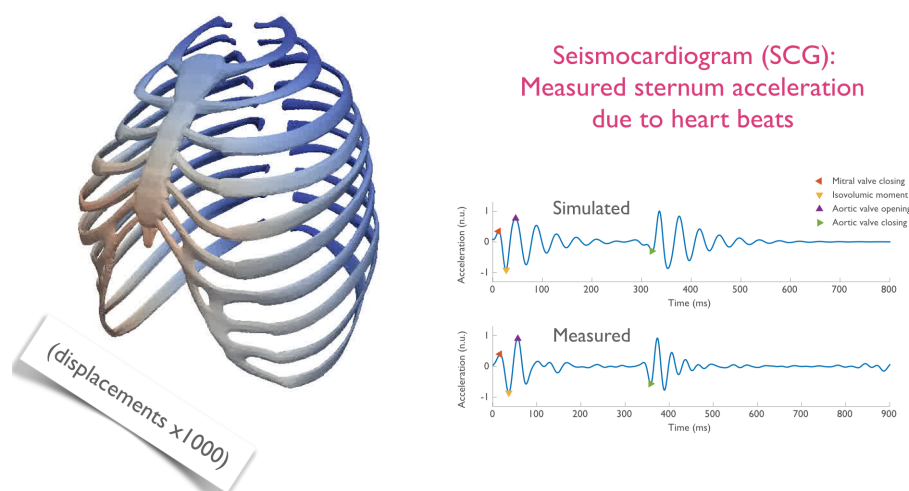
Figure 1. Simulation of the thorax deformation due to a heart beat and of the corresponding seismocardiogram

7.1.2. *Multi-scale modeling of muscle contraction*

Participants: François Kimmig [correspondant], Dominique Chapelle, Matthieu Caruel.

This works aims at proposing a multi-scale model of the muscular contraction that can be used in the context of cardiac simulation. The modeling will be based on the stochastic equations that describe muscular contraction at the molecular level. Asymptotic counterparts of the stochastic model will be considered in order to provide pertinent simplified models. The modeling elements will be confronted with experiments that will be performed on cardiac muscle cells by collaborators in the team of Professor Vincenzo Lombardi at the University of Florence.

In the framework of this collaboration, a chemicomechanical model is being implemented into CardiacLab, a simulation environment developed by the team. It will enrich the range of modeling tools of the team for the active contribution of muscle cells to the cardiac behavior.

7.1.3. *Multiphysics and multiscale modelling, data-model fusion and integration of organ physiology in the clinic: ventricular cardiac mechanics*

Participants: Radomir Chabiniok [correspondant], Philippe Moireau, Dominique Chapelle, Maxime Sermesant [Team Asclepios].

With heart and cardiovascular diseases continually challenging healthcare systems worldwide, translating basic research on cardiac (patho)physiology into clinical care is essential. Exacerbating this already extensive challenge is the complexity of the heart, relying on its hierarchical structure and function to maintain cardiovascular flow. Computational modelling has been proposed and actively pursued as a tool for accelerating research and translation. Allowing exploration of the relationships between physics, multiscale mechanisms and function, computational modelling provides a platform for improving our understanding of the heart. Further integration of experimental and clinical data through data assimilation and parameter estimation techniques is bringing computational models closer to use in routine clinical practice. This work published in [17] reviews developments in computational cardiac modelling and how their integration with medical imaging data is providing new pathways for translational cardiac modelling.

7.1.4. *Eidolon: visualization and computational framework for multi-modal biomedical data analysis*

Participant: Radomir Chabiniok [correspondant].

Biomedical research, combining multi-modal image and geometry data, presents unique challenges for data visualization, processing, and quantitative analysis. Medical imaging provides rich information, from anatomical to deformation, but extracting this to a coherent picture across image modalities with preserved quality is not trivial. Addressing these challenges and integrating visualization with image and quantitative analysis results in Eidolon, a platform which can adapt to rapidly changing research workflows. In the paper [26] we outline Eidolon, a software environment aimed at addressing these challenges, and discuss the novel integration of visualization and analysis components. These capabilities are demonstrated through the example of cardiac strain analysis, showing the Eidolon supports and enhances the workflow.

7.1.5. *Mathematical and numerical modeling of elastic waves propagation in the heart*

Participants: Federica Caforio [correspondant], Dominique Chapelle, Sébastien Imperiale.

The objective of this work is to develop a rigorous mathematical and numerical background for the extension and dissemination of elastography imaging modalities, applied to the cardiac setting. The problems treated concern the topics of mathematical modelling, numerical analysis and scientific computing. More precisely, the plan is to define a linearised model for the propagation of elastic waves in the heart, to study approximations of these models and define adapted numerical methods for the discretisation of the resulting partial differential equations.

7.2. Numerical Methods

7.2.1. *Effective and energy-preserving time discretization for a general nonlinear poromechanical formulation*

Participants: Bruno Burtschell, Dominique Chapelle [correspondant], Philippe Moireau.

In this work, we consider a general nonlinear poromechanical model, formulated based on fundamental thermodynamics principle, suitable for representing the coupling of rapid internal fluid flows with large deformations of the solid, and compatible with a wide class of constitutive behavior. The objective of the present work is to propose for this model a time discretization scheme of the partitioned type, to allow the use of existing time schemes - and possibly separate solvers - for each component of the model, i.e. for the fluid

and the solid. To that purpose, we adapt and extend an earlier proposed approach devised for fluid-structure interaction in an Arbitrary Lagrangian-Eulerian framework. We then establish an energy estimate for the resulting time scheme, in a form that is consistent with the underlying energy principle in the poromechanical formulation, up to some numerical dissipation effects and some perturbations that we have carefully identified and assessed. In addition, we provide some numerical illustrations of our numerical strategy with test problems that present typical features of large strains and rapid fluid flows, and also a case of singular transition related to total drainage. An example of challenging application envisioned for this model and associated numerical coupling scheme concerns the perfusion of the heart. This work has resulted in the publication [15].

7.2.2. *Delayed feedback control method for calculating space-time periodic solutions of viscoelastic problems*

Participants: Ustim Khristenko, Patrick Le Tallec.

We are interested in fast techniques for calculating a periodic solution to viscoelastic evolution problems with a space-time periodic condition. In order to avoid the invertion of very large matrices, such a solution is often computed as an asymptotic limit of the initial value problem with arbitrary initial data. We have developed a control method, accelerating the convergence to the periodic state. The main idea is to modify our problem by introducing a feedback control term, based on a periodicity error.

First, an abstract evolution problem has been studied. From the analytic solution of the modified (controlled) problem, an efficient control has been constructed, optimizing the spectrum of the problem. The proposed control term can be mechanically interpreted, and its efficiency increases with the relaxation time.

In order to confirm numerically the theoretical results, a finite element simulation has been carried out on a full 3D model for a steady rolling of a viscoelastic tyre with periodic sculpture. It has demonstrated that the controlled solution converges indeed faster than the non-controlled one, and that the efficiency of the method increases with the problem's relaxation time, that is when the memory of the underlying problem is large.

7.2.3. *Construction and analysis of an adapted spectral finite element method to convective acoustic equations*

Participant: Sébastien Imperiale [correspondant].

This work addresses the construction of a non spurious mixed spectral finite element (FE) method to problems in the field of computational aeroacous-tics. Based on a computational scheme for the conservation equations of linear acoustics, the extension towards convected wave propagation is investigated. In aeroacoustic applications, the mean flow effects can have a significant impact on the generated sound field even for smaller Mach numbers. For those convective terms, the initial spectral FE discretization leads to non-physical, spurious solutions. Therefore, a regularization procedure is proposed and qualitatively investigated by means of discrete eigenvalues analysis of the discrete operator in space. A study of convergence and an application of the proposed scheme to simulate the flow induced sound generation in the process of human phonation underlines stability and validity. This work has resulted in the publication [19].

7.2.4. *Space/time convergence analysis of a class of conservative schemes for linear wave equations*

Participants: Juliette Chabassier [MAGIQUE 3D team], Sébastien Imperiale [correspondant].

This work concerns the space/time convergence analysis of conservative two-steps time discretizations for linear wave equations. Explicit and implicit, second and fourth order schemes are considered, while the space discretization is given and satisfies minimal hypotheses. The convergence analysis is done using energy techniques and holds if the time step is upper-bounded by a quantity depending on space discretization parameters. In addition to showing the convergence for recently introduced fourth order schemes, the novelty of this work consists in the independency of the convergence estimates with respect to the difference between the time step and its greatest admissible value. This work has resulted in the publication [16].

7.3. Inverse Problems

7.3.1. *Front observer for data assimilation of electroanatomical mapping data for a numerical atrial model*

Participants: Antoine Gérard [Carmen team], Annabelle Collin [Monc team], Jason Bayer [Carmen team], Philippe Moireau [correspondant], Yves Coudière [Carmen team].

The purpose of our work is to personalize an atrial model of the propagation of the action potential, based on electrical catheter data with the help of the data assimilation approach introduced in [Collin & Al, Journal of Computational Physics 2015]. The originality of the approach is to indroduce a Luenberger observer of a surface atrial model of the propagation which can pursue - like in classical Kalman filtering approach - the actual activation front reconstructed from catheter data. Moreover, this approach may account for the breakthrough of new activation fronts at anytime with an additional topological gradient term. In the present work, we adapt this approach to the bilayer surface atrial model of th propagation of action potentials [Labarthe & Al, Europace 2014], and evaluated for the first time on a real patient's dataset. First, the model was geometrically fit to the patient's data. A fiber architecture was added to the geometry. Then an initial electrophysiological state was guessed, and the model was run with the Luenberger filter for some catheter data acquired during a CARTO procedure. All along the simulation, the filter corrects the action potential so as to track CARTO local activation times, while preserving a biophysical behavior. With this technique, we are able to reconstruct smooth activation maps over the whole atrial surfaces. This promising technique may also allow to reconstruct velocity fields and directions, phase map and possibly give information on repolarization. This work results from a collaborative project carried out during a training session at CEMRACS 2016 in Marseille, Luminy. This work has resulted in the publication [28].

7.3.2. *Iterative observer-based state and parameter estimation for linear systems*

Participant: Atte Aalto [correspondant].

In this work we propose an iterative method for joint state and parameter estimation using measurements on a finite time interval for systems that are backward output stabilizable. Since this time interval is fixed, errors in initial state may have a big impact on the parameter estimate. We propose to use the back and forth nudging (BFN) method for estimating the system's initial state and a Gauss–Newton step between BFN iterations for estimating the system parameters. Taking advantage of results on the optimality of the BFN method, we show that for systems with skew-adjoint generators, the initial state and parameter estimate minimizing an output error cost functional is an attractive fixed point for the proposed method. We treat both linear source estimation and bilinear parameter estimation problems.

7.3.3. *Estimation from moments measurements for amyloid depolymerisation*

Participants: Aurora Armiento [Mamba team], Marie Doumic [Mamba team], Philippe Moireau [correspondant].

Estimating reaction rates and size distributions of protein polymers is an important step for understanding the mechanisms of protein misfolding and aggregation, a key feature for amyloid diseases. This study aims at setting this framework problem when the experimental measurements consist in the time-dynamics of a moment of the population (*i.e.* for instance the total polymerised mass, as in Thioflavine T measurements, or the second moment measured by Static Light Scattering). We propose a general methodology, and we solve the problem theoretically and numerically in the case of a depolymerising system. We then apply our method to experimental data of degrading oligomers, and conclude that smaller aggregates of ovPrP protein should be more stable than larger ones. This has an important biological implication, since it is commonly admitted that small oligomers constitute the most cytotoxic species during prion misfolding process. This work has resulted in the publication [14].

7.3.4. *Analysis of an observers strategy for initial state reconstruction in unbounded domains*
Participants: Antoine Tonnoir [correspondant], Sonia Fliss [Poems team], Sébastien Imperiale, Philippe Moireau.

In this work, we are interested in the problem of recovering a compactly supported initial state of the wave equation in unbounded domain (such as the whole plane, a waveguide...). To this purpose, we assume that the velocity is known in a bounded observation region surrounding the support of the initial state. We consider an iterative algorithm of reconstruction based on back and forth nudging and prove the exponential convergence of this algorithm and its robustness with respect to noisy measures, at the continuous level. We also study the effect of the discretization process on the convergence of the algorithm.

7.4. Experiments and Clinical applications

7.4.1. *Characterization of mechanical properties of soft tissues*
Participants: Jean-Marc Allain [correspondant], Jean-Sebastien Affagard, Maeva Lopez Poncelas.

Soft tissues - such as skin - have complex mechanical properties: large strains, anisotropy, etc.. Identifying constitutive properties incorporating microstructure effects is very important for applications in medicine (surgery and other therapies) and industry (anti-ageing cosmetics, etc.). This characterization, however, requires complex experiments. We have developed a novel biaxial traction experimental method for mice skin, relying on a sensitivity analysis for determining optimal experimental parameters, including in particular sample size and most informative loading paths. This protocol has already been used on multiple samples, and 3 distinct constitutive laws of increasing complexity have been characterized (Master's internship of Maeva Lopez).

Another originality in our approach is to place our setup under a microscope to monitor the microstructure evolution during the test. These rich measurements allow detailed comparisons of classical models (such as Holzapfel's) with our data.

7.4.2. *Non-invasive model-based assessment of passive left-ventricular myocardial stiffness in healthy subjects and in patients with non-ischemic dilated cardiomyopathy*
Participant: Radomir Chabiniok [correspondant].

Patient-specific modelling has emerged as a tool for studying heart function, demonstrating the potential to provide non-invasive estimates of tissue passive stiffness. However, reliable use of model-derived stiffness requires sufficient model accuracy and unique estimation of model parameters. In this work we present personalised models of cardiac mechanics, focusing on improving model accuracy, while ensuring unique parametrisation. The influence of principal model uncertainties on accuracy and parameter identifiability was systematically assessed in a group of patients with dilated cardiomyopathy and healthy volunteers. For all cases, we examined three circumferentially symmetric fibre distributions and two epicardial boundary conditions. Our results demonstrated the ability of data-derived boundary conditions to improve model accuracy and highlighted the influence of the assumed fibre distribution on both model fidelity and stiffness estimates. The model personalisation pipeline – based strictly on non-invasive data – produced unique parameter estimates and satisfactory model errors for all cases, supporting the selected model assumptions. The thorough analysis performed enabled the comparison of passive parameters between volunteers and dilated cardiomyopathy patients, illustrating elevated stiffness in diseased hearts.

7.4.3. *Age-related changes in intraventricular kinetic energy: a physiological or pathological adaptation*
Participant: Radomir Chabiniok [correspondant].

Aging has important deleterious effects on the cardiovascular system. In this work we sought to compare intraventricular kinetic energy (KE) in healthy subjects of varying ages with subjects with ventricular

dysfunction to understand if changes in energetic momentum may predispose individuals to heart failure. Four-dimensional flow MRI was acquired in 35 healthy subjects (age: 1 -67 yr) and 10 patients with left ventricular (LV) dysfunction (age: 28-79 yr). Healthy subjects were divided into age quartiles (1st quartile: 16 yr, 2nd quartile: 17-32 yr, 3rd quartile: 33-48 yr, and 4th quartile: 49 - 64 yr). KE was measured in the LV throughout the cardiac cycle and indexed to ventricular volume. In healthy subjects, two large peaks corresponding to systole and early diastole occurred during the cardiac cycle. A third smaller peak was seen during late diastole in eight adults. Systolic KE (P 0.182) and ejection fraction (P 0.921) were preserved through all age groups. Older adults showed a lower early peak diastolic KE compared with children (P 0.0001) and young adults (P 0.025). Subjects with LV dysfunction had reduced ejection fraction (P 0.001) and compared with older healthy adults exhibited a similar early peak diastolic KE (P 0.142) but with the addition of an elevated KE in diastasis (P 0.029). In healthy individuals, peak diastolic KE progressively decreases with age, whereas systolic peaks remain constant. Peak diastolic KE in the oldest subjects is comparable to those with LV dysfunction. Unique age-related changes in ventricular diastolic energetics might be physiological or herald subclinical pathology. This work has resulted in the publication [24].

7.4.4. *Patient-specific computational analysis of ventricular mechanics in pulmonary arterial hypertension*

Participant: Martin Genet [correspondant].

Patient-specific biventricular computational models associated with a normal subject and a pulmonary arterial hypertension (PAH) patient were developed to investigate the disease effects on ventricular mechanics. These models were developed using geometry reconstructed from magnetic resonance (MR) images, and constitutive descriptors of passive and active mechanics in cardiac tissues. Model parameter values associated with ventricular mechanical properties and myofiber architecture were obtained by fitting the models with measured pressure–volume loops and circumferential strain calculated from MR images using a hyperelastic warping method. Results show that the peak right ventricle (RV) pressure was substantially higher in the PAH patient (65 mmHg versus 20 mmHg), who also has a significantly reduced ejection fraction (EF) in both ventricles (left ventricle (LV): 39% versus 66% and RV: 18% versus 64%). Peak systolic circumferential strain was comparatively lower in both the left ventricle (LV) and RV free wall (RVFW) of the PAH patient (LV: -6.8% versus -13.2% and RVFW: -2.1% versus -9.4%). Passive stiffness, contractility, and myofiber stress in the PAH patient were all found to be substantially increased in both ventricles, whereas septum wall in the PAH patient possessed a smaller curvature than that in the LV free wall. Simulations using the PAH model revealed an approximately linear relationship between the septum curvature and the transseptal pressure gradient at both early-diastole and end-systole. These findings suggest that PAH can induce LV remodeling, and septum curvature measurements may be useful in quantifying transseptal pressure gradient in PAH patients. This work has resulted in the publication [25].

8. Partnerships and Cooperations

8.1. European Initiatives

8.1.1. *FP7 & H2020 Projects*

8.1.1.1. *VP2HF*

Title: Computer model derived indices for optimal patient-specific treatment selection and planning in Heart Failure

Programm: FP7

Duration: October 2013 - March 2017

Coordinator: King's College London (UK)

See also: http://vp2hf.eu/

Inria contact: Dominique Chapelle

Abstract: Heart failure (HF) is one of the major health issues in Europe affecting 6 million patients and growing substantially because of the ageing population and improving survival following myocardial infarction. The poor short to medium term prognosis of these patients means that treatments such as cardiac re-synchronisation therapy and mitral valve repair can have substantial impact. However, these therapies are ineffective in up to 50% of the treated patients and involve significant morbidity and substantial cost. The primary aim of VP2HF is to bring together image and data processing tools with statistical and integrated biophysical models mainly developed in previous VPH projects, into a single clinical workflow to improve therapy selection and treatment optimisation in HF.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the organizing committees
Philippe Moireau

P. Moireau, Member of the CEMRACS-2016 organizing committee

M. Genet, Member of the GIENS-2017 organizing committee

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Editorial Boards

D. Chapelle, Member of the editorial board of journal *Computers & Structures*

D. Chapelle, Member of the editorial board of journal *ESAIM: M2AN*

9.1.2.2. Reviewer - Reviewing Activities

J-M. Allain, Reviewer for "Journal of the Royal Society Interface", "Physica D" and "Journal of the Mechanical Behavior of Biomedical Materials".

R. Chabiniok, reviewer for "Journal of Biomechanical Engineering and Computational" and "Mathematical Methods in Medicine".

M. Genet, reviewer for "Journal of Elasticity".

S. Imperiale, reviewer for "Journal of Computational Physics" and "Journal of Differential Equations".

9.1.3. Invited Talks

J.-M. Allain, "Caractérisation in vitro de tissus mous à l'échelle microscopique", Colloque Mecamat, Aussois.

J.-M. Allain, "Multiscale characterization of skin biomechanics", Workshop constitutive behaviour of soft tissue, Manchester, UK.

R. Chabiniok, "Biophysical modeling of cardiac function for clinical applications" at University Southwestern, Dallas, Texas (Seminar series of Biomedical Engineering Department, and at regular clinical echocardiography meeting of Dept. of Pediatrics, UT Southwestern Medical Center).

D. Chapelle, seminar at CEMRACS-16.

M. Genet, "Modélisation et simulation en biomécanique cardiaque", Département de Génie Mécanique, École Normale Supérieure de Cachan.

9.1.4. Leadership within the Scientific Community

J.-M. Allain, Member of Society of Experimental Mechanics and of Biophysical Society

J.-M. Allain, Member of the Academic Council of Université Paris-Sacaly, France

D. Chapelle, Member of the board of directors of the VPH Institute

9.1.5. Research Administration

J.-M. Allain, Responsability of the teaching experimental center (mechanics), 32h, Ecole Polytechnique, France

J.-M. Allain, Scientific Advisory Board, chair BioMecAM, ENSAM, Paris, France

D. Chapelle, VP research of Inria Saclay-Ile-de-France

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Bachelor: J.-M. Allain, co-supervision of the new program for the Polytechnique Bachelor, 15h, Ecole Polytechnique, France

Bachelor: F. Caforio, "Math 255 – Differential calculus for physics (mathematical analysis in two and three dimensions)", 42h, (L2), Université d'Orsay, France

Bachelor: M. Genet, "MEC431 – Modélisation et Simulation en Mécanique Industrielle", 32h, (L3), École Polytechnique, France

Bachelor: M. Genet, "MEC431 – Mécanique des Milieux Continus", 16h, (L3), École Polytechnique, France

Bachelor: S. Imperiale, "MA102 – Analyse pour les EDP", 24h, (L3), ENSTA ParisTech, France

Bachelor: S. Imperiale, "MA104 – Analyse complexe", 12h, (L3), ENSTA ParisTech, France

Bachelor: P. Moireau, "MA103 – Introduction aux EDP et à leur approximation numérique", 14h, (L3), ENSTA ParisTech, France

Master : J.-M. Allain, "Computational fluid dynamics", 36h, (M1), Ecole Polytechnique, France

Master : J.-M. Allain, "Cellular Motility", 32h, (M2), Ecole Polytechnique, France

Master: D. Chapelle, "Biomechanical Modeling of Active Tissues", 23h, (M2), Université Paris-Saclay, France

Master : M. Genet, "MEC551 – Plasticité & Rupture", 18h, (M1), École Polytechnique, France

Master: S. Imperiale, "MA2610 Calcul Scientifique – Mécanique des solides", 6h, (M1), Central/Supelec, France

Master: S. Imperiale, "Simnum – Programmation C++", 18h, (M1), ENSTA ParisTech, France

Master: S. Imperiale, "MAP-Ann1 – La méthode des éléments finis", 12h, (M1), ENSTA ParisTech, France

Master: P. Moireau, "MAP-Ann1 – La méthode des éléments finis", 21h, (M1), ENSTA ParisTech, France

Master: P. Moireau, "MAP 431 – Analyse variationnelle des équations aux dérivées partielles", (M1), Ecole Polytechnique, France

Master: P. Moireau, "Biomechanical Modeling of Active Tissues", 12h, (M2), Université Paris-Saclay, France

Master: P. Moireau, "Méthodes et problèmes inverses en dynamique des populations", 24h, (M2), UPMC, France

9.2.2. Supervision

HdR : Philippe Moireau, Observers for data assimilation – Applications to cardiac modeling, Université Paris-Saclay, November 28th

PhD : Bruno Burtschell, Mechanical modeling and numerical methods for poromechanics – Application to myocardium perfusion, Université Paris-Saclay, September 30th, supervisors: D. Chapelle and P. Moireau

PhD in progress : Aurora Armiento, Inverse problems and data assimilation methods applied to protein depolymerisation, started: Nov 2013, supervisors: M. Doumic and P. Moireau

PhD in progress : Federica Caforio, "Modélisation mathématique et numérique de la propagation d'ondes élastique dans le coeur", started: Nov 2015, supervisors: D. Chapelle and S. Imperiale

PhD in progress : Florent Wijanto, Modélisation multi-échelle des fibres de collagènes, started: Sept 2015, supervisors:, Jean-Marc Allain and Mathieu Carruel

PhD in progress : Arthur Le Gall, "Application of biomechanical heart modeling in hemodynamic monitoring of increased risk patients during aneshesia using clinical data", started: Nov 2016, supervisors: Dominique Chapelle, Etienne Gayat, Radomir Chabiniok

PhD in progress : François Kimmig, "Multi-scale modeling of muscle contraction", started: Sept 2016, supervisors: Dominique Chapelle, Matthieu Caruel

9.2.3. *Juries*

J.-M. Allain, Reviewer for Laure Laforgue's PhD Thesis, LiPhy Grenoble.

P. Moireau, Reviewer for Stefano Pagani's PhD Thesis, Politecnico di Milano.

9.3. Popularization

D. Chapelle, Debate on "Data sciences and personalized medicine" at Cité des Sciences (also on web-TV), Oct 9th

D. Chapelle, Roundtable in workshop "Mathématiques Oxygène du Numérique" (UPMC, Oct 21st)

10. Bibliography

Major publications by the team in recent years

[1] R. CHABINIOK, P. MOIREAU, P.-F. LESAULT, A. RAHMOUNI, J.-F. DEUX, D. CHAPELLE.*Estimation of tissue contractility from cardiac cine-MRI using a biomechanical heart model*, in "Biomechanics and Modeling in Mechanobiology", 2012, vol. 11, n° 5, p. 609-630 [*DOI : 10.1007/S10237-011-0337-8*], http://hal.inria.fr/hal-00654541.

[2] D. CHAPELLE, K. BATHE.*The Finite Element Analysis of Shells - Fundamentals - Second Edition*, Computational Fluid and Solid Mechanics, Springer, 2011, 410 [*DOI : 10.1007/978-3-642-16408-8*], http://hal.inria.fr/hal-00654533.

[3] D. CHAPELLE, N. CÎNDEA, P. MOIREAU.*Improving convergence in numerical analysis using observers - The wave-like equation case*, in "Mathematical Models and Methods in Applied Sciences", 2012, vol. 22, n° 12 [*DOI : 10.1142/S0218202512500406*], http://hal.inria.fr/inria-00621052.

[4] D. CHAPELLE, M. FRAGU, V. MALLET, P. MOIREAU.*Fundamental principles of data assimilation underlying the Verdandi library: applications to biophysical model personalization within euHeart*, in "Medical & Biological Engineering & Computing", 2013, vol. 51, p. 1221–1233 [*DOI : 10.1007/S11517-012-0969-6*], https://hal.inria.fr/hal-00760887.

[5] D. CHAPELLE, P. LE TALLEC, P. MOIREAU, M. SORINE.*An energy-preserving muscle tissue model: formulation and compatible discretizations*, in "International Journal for Multiscale Computational Engineering", 2012, vol. 10, n⁰ 2, p. 189-211 [*DOI :* 10.1615/INTJMULTCOMPENG.2011002360], http://hal.inria.fr/hal-00678772.

[6] D. CHAPELLE, P. MOIREAU.*General coupling of porous flows and hyperelastic formulations – From thermodynamics principles to energy balance and compatible time schemes*, in "European Journal of Mechanics - B/Fluids", 2014, vol. 46, p. 82-96, Updated version of previously published research report [*DOI :* 10.1016/J.EUROMECHFLU.2014.02.009], https://hal.inria.fr/inria-00520612.

[7] P. MOIREAU, D. CHAPELLE, P. LE TALLEC.*Joint state and parameter estimation for distributed mechanical systems*, in "Computer Methods in Applied Mechanics and Engineering", 2008, vol. 197, n⁰ 6-8, p. 659-677 [*DOI :* 10.1016/J.CMA.2007.08.021], http://hal.archives-ouvertes.fr/hal-00175623.

[8] P. MOIREAU, D. CHAPELLE, P. LE TALLEC.*Filtering for distributed mechanical systems using position measurements: perspectives in medical imaging*, in "Inverse Problems", 2009, vol. 25, n⁰ 3, 035010 [*DOI :* 10.1088/0266-5611/25/3/035010], http://hal.archives-ouvertes.fr/hal-00358914.

[9] P. MOIREAU, D. CHAPELLE.*Reduced-order Unscented Kalman Filtering with application to parameter identification in large-dimensional systems*, in "ESAIM - Control Optimisation and Calculus of Variations", 2010, Published online - See also erratum DOI:10.1051/cocv/2011001 [*DOI :* 10.1051/COCV/2010006], http://hal.inria.fr/inria-00550104.

[10] M. SERMESANT, R. CHABINIOK, P. CHINCHAPATNAM, T. MANSI, F. BILLET, P. MOIREAU, J.-M. PEYRAT, K. C. WONG, J. RELAN, K. S. RHODE, M. GINKS, P. LAMBIASE, H. DELINGETTE, M. SORINE, C. A. RINALDI, D. CHAPELLE, R. RAZAVI, N. AYACHE.*Patient-Specific Electromechanical Models of the Heart for Prediction of the Acute Effects of Pacing in CRT: a First Validation*, in "Medical Image Analysis", January 2012, vol. 16, n⁰ 1, p. 201-215 [*DOI :* 10.1016/J.MEDIA.2011.07.003], http://hal.inria.fr/inria-00616191.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] B. BURTSCHELL.*Mechanical modeling and numerical methods for poromechanics - Application to myocardium perfusion*, Ecole Polytechnique (EDX), September 2016, https://tel.archives-ouvertes.fr/tel-01396502.

[12] P. MOIREAU.*Observers for data assimilation - Applications to cardiac modeling*, Université Paris Saclay ; Université Paris Sud - Paris XI, November 2016, Habilitation à diriger des recherches, https://tel.archives-ouvertes.fr/tel-01404866.

Articles in International Peer-Reviewed Journal

[13] A. AALTO.*Output error minimizing back and forth nudging method for initial state recovery*, in "Systems and Control Letters", June 2016, vol. 94, p. 111-117, This is the preprint version of the article. The final, published version is available on the journal's website [*DOI :* 10.1016/J.SYSCONLE.2016.06.002], https://hal.inria.fr/hal-01216075.

[14] A. ARMIENTO, M. DOUMIC, P. MOIREAU, H. REZAEI.*Estimation from Moments Measurements for Amyloid Depolymerisation*, in "Journal of Theoretical Biology", March 2016 [*DOI :* 10.1016/J.JTBI.2016.02.037], https://hal.archives-ouvertes.fr/hal-01248255.

[15] B. BURTSCHELL, D. CHAPELLE, P. MOIREAU.*Effective and energy-preserving time discretization for a general nonlinear poromechanical formulation*, in "Computers and Structures", 2016 [*DOI :* 10.1016/J.COMPSTRUC.2016.10.022], https://hal.inria.fr/hal-01395508.

[16] J. CHABASSIER, S. IMPERIALE.*Space/Time convergence analysis of a class of conservative schemes for linear wave equations*, in "Comptes Rendus Mathématique", December 2016, https://hal.inria.fr/hal-01421882.

[17] R. CHABINIOK, V. Y. WANG, M. HADJICHARALAMBOUS, L. ASNER, J. LEE, M. SERMESANT, E. KUHL, A. A. YOUNG, P. MOIREAU, M. P. NASH, D. CHAPELLE, D. NORDSLETTEN.*Multiphysics and multiscale modelling, data-model fusion and integration of organ physiology in the clinic: ventricular cardiac mechanics*, in "Interface Focus", 2016, vol. 6, n° 2 [*DOI :* 10.1098/RSFS.2015.0083], https://hal.inria.fr/hal-01277684.

[18] M. HADJICHARALAMBOUS, L. ASNER, R. CHABINIOK, E. SAMMUT, J. WONG, D. PERESSUTTI, E. KERFOOT, A. D. KING, J. LEE, R. RAZAVI, N. SMITH, G. CARR-WHITE, D. NORDSLETTEN.*Non-invasive Model-Based Assessment of Passive Left-Ventricular Myocardial Stiffness in Healthy Subjects and in Patients with Non-ischemic Dilated Cardiomyopathy*, in "Annals of Biomedical Engineering", August 2016, https://hal.inria.fr/hal-01419936.

[19] A. HÜPPE, G. COHEN, S. IMPERIALE, M. KALTENBACHER.*Construction and analysis of an adapted spectral finite element method to convective acoustic equations*, in "Communications in Computational Physics", June 2016 [*DOI :* 10.4208/CICP.250515.161115A], https://hal.inria.fr/hal-01363149.

[20] M. LANDAJUELA, M. VIDRASCU, D. CHAPELLE, M. A. FERNÁNDEZ.*Coupling schemes for the FSI forward prediction challenge: comparative study and validation*, in "International Journal for Numerical Methods in Biomedical Engineering", June 2016 [*DOI :* 10.1002/CNM.2813], https://hal.inria.fr/hal-01239931.

[21] A. LAURIN, F. KHORSOW, A. BLABER, K. TAVAKOLIAN.*Accurate and consistent automatic seismocardiogram annotation without concurrent ECG*, in "Physiological Measurement", September 2016, vol. 37, n° 9 [*DOI :* 10.1088/0967-3334/37/9/1588], https://hal.inria.fr/hal-01356419.

[22] N. MITROU, A. LAURIN, T. DICK, J. INSKIP.*A peak detection method for identifying phase in physiological signals*, in "Biomedical Signal Processing and Control", 2017, vol. 31, p. 452 - 462 [*DOI :* 10.1016/J.BSPC.2016.07.001], https://hal.inria.fr/hal-01380875.

[23] E. SCHENONE, A. COLLIN, J.-F. GERBEAU.*Numerical simulation of electrocardiograms for full cardiac cycles in healthy and pathological conditions*, in "International Journal for Numerical Methods in Biomedical Engineering", May 2016, vol. 32, n° 5, In press (in International Journal for Numerical Methods in Biomedical Engineering) [*DOI :* 10.1002/CNM.2744], https://hal.inria.fr/hal-01184744.

[24] J. WONG, R. CHABINIOK, A. DEVECCHI, N. DEDIEU, E. SAMMUT, T. SCHAEFFTER, R. RAZAVI.*Age-related changes in intraventricular kinetic energy: a physiological or pathological adaptation?*, in "AJP - Heart and Circulatory Physiology", 2016, vol. 310, p. 747 - 755 [*DOI :* 10.1152/AJPHEART.00075.2015], https://hal.archives-ouvertes.fr/hal-01386703.

[25] C. XI, C. LATNIE, J. L. TAN, S. T. WALL, M. GENET, L. ZHONG, L. C. LEE, X. ZHAO.*Patient-Specific Computational Analysis of Ventricular Mechanics in Pulmonary Arterial Hypertension*, in "Journal of Biomechanical Engineering", November 2016, vol. 138, n⁰ 11 [*DOI :* 10.1115/1.4034559], https://hal.archives-ouvertes.fr/hal-01427545.

Conferences without Proceedings

[26] E. KERFOOT, L. FOVARGUE, S. RIVOLO, W. SHI, D. RUECKERT, D. NORDSLETTEN, J. LEE, R. CHABINIOK, R. RAZAVI.*Eidolon: Visualization and Computational Framework for Multi-Modal Biomedical Data Analysis*, in "7th International Conference on Medical Imaging and Augmented Reality (MIAR 2016)", Berne, Switzerland, August 2016, https://hal.archives-ouvertes.fr/hal-01346357.

Other Publications

[27] A. AALTO.*Iterative observer-based state and parameter estimation for linear systems*, September 2016, Manuscript submitted for publication, https://hal.inria.fr/hal-01370430.

[28] A. GÉRARD, A. COLLIN, J. BAYER, A. FRONTERA, P. MOIREAU, Y. COUDIÈRE.*Front observer for data assimilation of electroanatomical mapping data for a numerical atrial model*, September 2016, Liryc Workshop, Poster, https://hal.inria.fr/hal-01400776.

Project-Team MEXICO

Modeling and Exploitation of Interaction and Concurrency

IN COLLABORATION WITH: Laboratoire specification et vérification (LSV)

IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Cachan

RESEARCH CENTER
Saclay - Île-de-France

THEME
Proofs and Verification

Table of contents

Project-Team MEXICO

Creation of the Team: 2009 March 01, updated into Project-Team: 2011 January 01

Keywords:

Computer Science and Digital Science:

      2.3. - Embedded and cyber-physical systems

      2.4. - Verification, reliability, certification

      4.5. - Formal methods for security

      6.4.3. - Observability and Controlability

      7.1. - Parallel and distributed algorithms

      7.2. - Discrete mathematics, combinatorics

      7.3. - Optimization

      7.4. - Logic in Computer Science

      7.9. - Graph theory

      7.11. - Performance evaluation

Other Research Topics and Application Domains:

      1.1.2. - Molecular biology

      1.1.3. - Cellular biology

      1.1.11. - Systems biology

      1.1.12. - Synthetic biology

      6.3.1. - Web

      6.3.3. - Network Management

      7.1. - Traffic management

      7.2.1. - Smart vehicles

1. Members

Research Scientists

    Stefan Haar [Team leader, Inria, Senior Researcher, HDR]

    Benedikt Bollig [CNRS, Researcher]

    Matthias Fuegger [CNRS, Researcher]

Faculty Members

    Beatrice Berard [Univ. Paris VI, Professor, Inria *delegation* from Sep 2016]

    Thomas Chatain [ENS Cachan, Associate Professor]

    Paul Gastin [ENS Cachan, Professor, HDR]

    Serge Haddad [ENS Cachan, Professor, HDR]

    Claudine Picaronny [ENS Cachan, Associate Professor]

    Stefan Schwoon [ENS Cachan, Associate Professor]

PhD Students

    Yann Duplouy [Inst. de Recherche Technologique SystemX]

    Engel Lefaucheux [ENS Cachan]

    Marie Fortin [ENS Cachan, from Sep 2016]

    Hugues Mandon [Inria, from Oct 2016]

    Simon Theissing [Inria, until Aug 2016, granted by Institut de Recherche SystemX]

Administrative Assistant
   Thida Iem [Inria]
Other
   Clara Scherbaum [Inria, Internship student (ERASMUS) from Aachen University, Germany, from Mar 2016 until Jul 2016]

2. Overall Objectives

2.1. Scientific Objectives

2.1.1. Introduction

In the increasingly networked world, reliability of applications becomes ever more critical as the number of users of, e.g., communication systems, web services, transportation etc., grows steadily. Management of networked systems, in a very general sense of the term, therefore is a crucial task, but also a difficult one.

MExICo strives to take advantage of distribution by orchestrating cooperation between different agents that observe local subsystems, and interact in a localized fashion.

The need for applying formal methods in the analysis and management of complex systems has long been recognized. It is with much less unanimity that the scientific community embraces methods based on asynchronous and distributed models. Centralized and sequential modeling still prevails.

However, we observe that crucial applications have increasing numbers of users, that networks providing services grow fast both in the number of participants and the physical size and degree of spatial distribution. Moreover, traditional *isolated* and *proprietary* software products for local systems are no longer typical for emerging applications.

In contrast to traditional centralized and sequential machinery for which purely functional specifications are efficient, we have to account for applications being provided from diverse and non-coordinated sources. Their distribution (e.g. over the Web) must change the way we verify and manage them. In particular, one cannot ignore the impact of quantitative features such as delays or failure likelihoods on the functionalities of composite services in distributed systems.

We thus identify three main characteristics of complex distributed systems that constitute research challenges:

- *Concurrency* of behavior;
- *Interaction* of diverse and semi-transparent components; and
- management of *Quantitative* aspects of behavior.

2.1.2. Concurrency

The increasing size and the networked nature of communication systems, controls, distributed services, etc. confront us with an ever higher degree of parallelism between local processes. This field of application for our work includes telecommunication systems and composite web services. The challenge is to provide sound theoretical foundations and efficient algorithms for management of such systems, ranging from controller synthesis and fault diagnosis to integration and adaptation. While these tasks have received considerable attention in the *sequential* setting, managing *non-sequential* behavior requires profound modifications for existing approaches, and often the development of new approaches altogether. We see concurrency in distributed systems as an opportunity rather than a nuisance. Our goal is to *exploit* asynchronicity and distribution as an advantage. Clever use of adequate models, in particular *partial order semantics* (ranging from Mazurkiewicz traces to event structures to MSCs) actually helps in practice. In fact, the partial order vision allows us to make causal precedence relations explicit, and to perform diagnosis and test for the dependency between events. This is a conceptual advantage that interleaving-based approaches cannot match. The two key features of our work will be *(i)* the exploitation of concurrency by using asynchronous models with partial order semantics, and *(ii)* distribution of the agents performing management tasks.

2.1.3. Interaction

Systems and services exhibit non-trivial *interaction* between specialized and heterogeneous components. A coordinated interplay of several components is required; this is challenging since each of them has only a limited, partial view of the system's configuration. We refer to this problem as *distributed synthesis* or *distributed control*. An aggravating factor is that the structure of a component might be semi-transparent, which requires a form of *grey box management*.

2.1.4. Quantitative Features

Besides the logical functionalities of programs, the *quantitative* aspects of component behavior and interaction play an increasingly important role.

- *Real-time* properties cannot be neglected even if time is not an explicit functional issue, since transmission delays, parallelism, etc, can lead to time-outs striking, and thus change even the logical course of processes. Again, this phenomenon arises in telecommunications and web services, but also in transport systems.

- In the same contexts, *probabilities* need to be taken into account, for many diverse reasons such as unpredictable functionalities, or because the outcome of a computation may be governed by race conditions.

- Last but not least, constraints on *cost* cannot be ignored, be it in terms of money or any other limited resource, such as memory space or available CPU time.

2.1.5. Evolution and Perspectives

Since the creation of *MExICo*, the weight of *quantitative* aspects in all parts of our activities has grown, be it in terms of the models considered (weighted automata and logics), be it in transforming verification or diagnosis verdict into probabilistic statements (probabilistic diagnosis, statistical model checking), or within the recently started SystemX cooperation on supervision in multi-modal transport systems. This trend is certain to continue over the next couple of years, along with the growing importance of diagnosis and control issues.

In another development, the theory and use of partial order semantics has gained momentum in the past four years, and we intend to further strengthen our efforts and contacts in this domain to further develop and apply partial-order based deduction methods.

As concerns the study of interaction, our progress has been thus far less in the domain of *distributed* approaches than in the analysis of *system composition*, such as in networks of untimed or timed automata. While continuing this line of study, we also intend to turn more strongly towards distributed *algorithms*, namely in terms of parametrized verification methods.

3. Research Program

3.1. Concurrency

Participants: Benedikt Bollig, Thomas Chatain, Paul Gastin, Stefan Haar, Serge Haddad, Stefan Schwoon.

Concurrency; Semantics; Automatic Control ; Diagnosis ; Verification

Concurrency: Property of systems allowing some interacting processes to be executed in parallel.

Diagnosis: The process of deducing from a partial observation of a system aspects of the internal states or events of that system; in particular, *fault diagnosis* aims at determining whether or not some non-observable fault event has occurred.

Conformance Testing: Feeding dedicated input into an implemented system IS and deducing, from the resulting output of I, whether I respects a formal specification S.

3.1.1. Introduction

It is well known that, whatever the intended form of analysis or control, a *global* view of the system state leads to overwhelming numbers of states and transitions, thus slowing down algorithms that need to explore the state space. Worse yet, it often blurs the mechanics that are at work rather than exhibiting them. Conversely, respecting concurrency relations avoids exhaustive enumeration of interleavings. It allows us to focus on 'essential' properties of non-sequential processes, which are expressible with causal precedence relations. These precedence relations are usually called causal (partial) orders. Concurrency is the explicit absence of such a precedence between actions that do not have to wait for one another. Both causal orders and concurrency are in fact essential elements of a specification. This is especially true when the specification is constructed in a distributed and modular way. Making these ordering relations explicit requires to leave the framework of state/interleaving based semantics. Therefore, we need to develop new dedicated algorithms for tasks such as conformance testing, fault diagnosis, or control for distributed discrete systems. Existing solutions for these problems often rely on centralized sequential models which do not scale up well.

3.1.2. Diagnosis

Participants: Benedikt Bollig, Stefan Haar, Serge Haddad, Stefan Schwoon.

Fault Diagnosis for discrete event systems is a crucial task in automatic control. Our focus is on *event oriented* (as opposed to *state oriented*) model-based diagnosis, asking e.g. the following questions:
given a - potentially large - *alarm pattern* formed of observations,

-   what are the possible *fault scenarios* in the system that *explain* the pattern ?

-   Based on the observations, can we deduce whether or not a certain - invisible - fault has actually occurred ?

Model-based diagnosis starts from a discrete event model of the observed system - or rather, its relevant aspects, such as possible fault propagations, abstracting away other dimensions. From this model, an extraction or unfolding process, guided by the observation, produces recursively the explanation candidates.

In asynchronous partial-order based diagnosis with Petri nets [51], [52], [56], one unfolds the *labelled product* of a Petri net model $\mathcal{N}$ and an observed alarm pattern $\mathcal{A}$, also in Petri net form. We obtain an acyclic net giving partial order representation of the behaviors compatible with the alarm pattern. A recursive online procedure filters out those runs *(configurations)* that explain *exactly* $\mathcal{A}$. The Petri-net based approach generalizes to dynamically evolving topologies, in dynamical systems modeled by graph grammars, see [35]

3.1.2.1. Observability and Diagnosability

Diagnosis algorithms have to operate in contexts with low observability, i.e., in systems where many events are invisible to the supervisor. Checking *observability* and *diagnosability* for the supervised systems is therefore a crucial and non-trivial task in its own right. Analysis of the relational structure of occurrence nets allows us to check whether the system exhibits sufficient visibility to allow diagnosis. Developing efficient methods for both verification of *diagnosability checking* under concurrency, and the *diagnosis* itself for distributed, composite and asynchronous systems, is an important field for *MExICo*.

3.1.2.2. Distribution

Distributed computation of unfoldings allows one to factor the unfolding of the global system into smaller *local* unfoldings, by local supervisors associated with sub-networks and communicating among each other. In [52], [37], elements of a methodology for distributed computation of unfoldings between several supervisors, underwritten by algebraic properties of the category of Petri nets have been developed. Generalizations, in particular to Graph Grammars, are still do be done.

Computing diagnosis in a distributed way is only one aspect of a much vaster topic, that of *distributed diagnosis* (see [48], [60]). In fact, it involves a more abstract and often indirect reasoning to conclude whether or not some given invisible fault has occurred. Combination of local scenarios is in general not sufficient: the global system may have behaviors that do not reveal themselves as faulty (or, dually, non-faulty) on any local supervisor's domain (compare [34], [40]). Rather, the local diagnosers have to join all *information* that is available to them locally, and then deduce collectively further information from the combination of their views. In particular, even the *absence* of fault evidence on all peers may allow to deduce fault occurrence jointly, see [64], [65]. Automatizing such procedures for the supervision and management of distributed and locally monitored asynchronous systems is a long-term goal to which *MExICo* hopes to contribute.

3.1.3. *Contextual nets*

Participant: Stefan Schwoon.

Assuring the correctness of concurrent systems is notoriously difficult due to the many unforeseeable ways in which the components may interact and the resulting state-space explosion. A well-established approach to alleviate this problem is to model concurrent systems as Petri nets and analyse their unfoldings, essentially an acyclic version of the Petri net whose simpler structure permits easier analysis [50].

However, Petri nets are inadequate to model concurrent read accesses to the same resource. Such situations often arise naturally, for instance in concurrent databases or in asynchronous circuits. The encoding tricks typically used to model these cases in Petri nets make the unfolding technique inefficient. Contextual nets, which explicitly do model concurrent read accesses, address this problem. Their accurate representation of concurrency makes contextual unfoldings up to exponentially smaller in certain situations. An abstract algorithm for contextual unfoldings was first given in [36]. In recent work, we further studied this subject from a theoretical and practical perspective, allowing us to develop concrete, efficient data structures and algorithms and a tool (Cunf) that improves upon existing state of the art. This work led to the PhD thesis of César Rodríguez in 2014 .

Contextual unfoldings deal well with two sources of state-space explosion: concurrency and shared resources. Recently, we proposed an improved data structure, called *contextual merged processes* (CMP) to deal with a third source of state-space explosion, i.e. sequences of choices. The work on CMP [66] is currently at an abstract level. In the short term, we want to put this work into practice, requiring some theoretical groundwork, as well as programming and experimentation.

Another well-known approach to verifying concurrent systems is *partial-order reduction*, exemplified by the tool SPIN. Although it is known that both partial-order reduction and unfoldings have their respective strengths and weaknesses, we are not aware of any conclusive comparison between the two techniques. Spin comes with a high-level modeling language having an explicit notion of processes, communication channels, and variables. Indeed, the reduction techniques implemented in Spin exploit the specific properties of these features. On the other side, while there exist highly efficient tools for unfoldings, Petri nets are a relatively general low-level formalism, so these techniques do not exploit properties of higher language features. Our work on contextual unfoldings and CMPs represents a first step to make unfoldings exploit richer models. In the long run, we wish raise the unfolding technique to a suitable high-level modelling language and develop appropriate tool support.

3.1.4. *Dynamic and parameterized concurrent systems*

Participants: Benedikt Bollig, Paul Gastin.

In the past few years, our research has focused on concurrent systems where the architecture, which provides a set of processes and links between them, is *static* and *fixed in advance*. However, the assumption that the set of processes is fixed somehow seems to hinder the application of formal methods in practice. It is not appropriate in areas such as mobile computing or ad-hoc networks. In concurrent programming, it is actually perfectly natural to design a program, and claim its correctness, independently of the number of processes that participate in its execution. There are, essentially, two kinds of systems that fall into this category. When the process architecture is static but unknown, it is a parameter of the system; we then call a system

parameterized. When, on the other hand, the process architecure is generated at runtime (i.e., process creation is a communication primitive), we say that a system is *dynamic*. Though parameterized and dynamic systems have received increasing interest in recent years, there is, by now, no canonical approach to modeling and verifying such systems. Our research program aims at the development of *a theory of parameterized and dynamic concurrent systems*. More precisely, our goal is a *unifying* theory that lays algebraic, logical, and automata-theoretic foundations to support and facilitate the study of parameterized and dynamic concurrent systems. Such theories indeed exist in non-parameterized settings where the number of processes and the way they are connected are fixed in advance. However, parameterized and dynamic systems lack such foundations and often restict to very particular models with specialized verification techniques.

3.1.5. *Testing*

Participants: Benedikt Bollig, Paul Gastin, Stefan Haar.

3.1.5.1. Introduction

The gap between specification and implementation is at the heart of research on formal testing. The general *conformance testing problem* can be defined as follows: Does an implementation $\mathcal{M}'$ conform a given specification $\mathcal{M}$? Here, both $\mathcal{M}$ and $\mathcal{M}'$ are assumed to have input and output channels. The formal model $\mathcal{M}$ of the specification is entirely known and can be used for analysis. On the other hand, the implementation $\mathcal{M}'$ is unknown but interacts with the environment through observable input and output channels. So the behavior of $\mathcal{M}'$ is partially controlled by input streams, and partially observable via output streams. The Testing problem consists in computing, from the knowledge of $\mathcal{M}$, *input streams* for $\mathcal{M}'$ such that observation of the resulting output streams from $\mathcal{M}'$ allows to determine whether $\mathcal{M}'$ conforms to $\mathcal{M}$ as intended.

In this project, we focus on distributed or asynchronous versions of the conformance testing problem. There are two main difficulties. First, due to the distributed nature of the system, it may not be possible to have a unique global observer for the outcome of a test. Hence, we may need to use *local* observers which will record only *partial views* of the execution. Due to this, it is difficult or even impossible to reconstruct a coherent global execution. The second difficulty is the lack of global synchronization in distributed asynchronous systems. Up to now, models were described with I/O automata having a centralized control, hence inducing global synchronizations.

3.1.5.2. Asynchronous Testing

Since 2006 and in particular during his sabbatical stay at the University of Ottawa, Stefan Haar has been working with Guy-Vincent Jourdan and Gregor v. Bochmann of UOttawa and Claude Jard of IRISA on asynchronous testing. In the synchronous (sequential) approach, the model is described by an I/O automaton with a centralized control and transitions labeled with individual input or output actions. This approach has known limitations when inputs and outputs are distributed over remote sites, a feature that is characteristic of , e.g., web computing. To account for concurrency in the system, they have developed in [58], [41] asynchronous conformance testing for automata with transitions labeled with (finite) partial orders of I/O. Intuitively, this is a "big step" semantics where each step allows concurrency but the system is synchronized before the next big step. This is already an important improvement on the synchronous setting. The non-trivial challenge is now to cope with fully asynchronous specifications using models with decentralized control such as Petri nets.

3.1.5.3. Near Future

Completion of asynchronous testing in the setting without any big-step synchronization, and an improved understanding of the relations and possible interconnections between local (i.e. distributed) and asynchronous (centralized) testing. This has been the objective of the *TECSTES* project (2011-2014), funded by a DIGITEO *DIM/LSC* grant, and which involved Hernán Ponce de Léon and Stefan Haar of *MExICo*, and Delphine Longuet at LRI, University Paris-Sud/Orsay. We have extended several well known conformance (ioco style) relations for sequential models to models that can handle concurrency (labeled event structures). Two semantics (interleaving and partial order) were presented for every relation. With the interleaving semantics, the relations we obtained boil down to the same relations defined for labeled transition systems, since they focus on sequences of actions. The only advantage of using labeled event structures as a specification formalism for testing remains in the conciseness of the concurrent model with respect to a sequential one.

As far as testing is concerned, the benefit is low since every interleaving has to be tested. By contrast, under the partial order semantics, the relations we obtain allow to distinguish explicitly implementations where concurrent actions are implemented concurrently, from those where they are interleaved, i.e. implemented sequentially. Therefore, these relations will be of interest when designing distributed systems, since the natural concurrency between actions that are performed in parallel by different processes can be taken into account. In particular, the fact of being unable to control or observe the order between actions taking place on different processes will not be considered as an impediment for testing. We have developed 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.

A mid-to long term goal (which may or may not be addressed by *MExICo* depending on the availability of staff for this subject) is the comprehensive formalization of testing and testability in asynchronous systems with distributed architecture and test protocols.

3.2. Interaction

Participants: Benedikt Bollig, Thomas Chatain, Paul Gastin, Stefan Haar, Serge Haddad.

3.2.1. *Introduction*

Systems and services exhibit non-trivial *interaction* between specialized and heterogeneous components. This interplay is challenging for several reasons. On one hand, a coordinated interplay of several components is required, though each has only a limited, partial view of the system's configuration. We refer to this problem as *distributed synthesis* or *distributed control*. An aggravating factor is that the structure of a component might be semi-transparent, which requires a form of *grey box management*.

Interaction, one of the main characteristics of systems under consideration, often involves an environment that is not under the control of cooperating services. To achieve a common goal, the services need to agree upon a strategy that allows them to react appropriately regardless of the interactions with the environment. Clearly, the notions of opponents and strategies fall within *game theory*, which is naturally one of our main tools in exploring interaction. We will apply to our problems techniques and results developed in the domains of distributed games and of games with partial information. We will consider also new problems on games that arise from our applications.

3.2.2. *Distributed Control*

Participants: Benedikt Bollig, Thomas Chatain, Paul Gastin, Stefan Haar.

Program synthesis, as introduced by Church [47] aims at deriving directly an implementation from a specification, allowing the implementation to be correct by design. When the implementation is already at hand but choices remain to be resolved at run time then the problem becomes controller synthesis. Both program and controller synthesis have been extensively studied for sequential systems. In a distributed setting, we need to synthesize a distributed program or distributed controllers that interact locally with the system components. The main difficulty comes from the fact that the local controllers/programs have only a partial view of the entire system. This is also an old problem largely considered undecidable in most settings [63], [59], [62], [53], [55].

Actually, the main undecidability sources come from the fact that this problem was addressed in a synchronous setting using global runs viewed as sequences. In a truly distributed system where interactions are asynchronous we have recently obtained encouraging decidability results [54], [45]. This is a clear witness where concurrency may be exploited to obtain positive results. It is essential to specify expected properties directly in terms of causality revealed by partial order models of executions (MSCs or Mazurkiewicz traces). We intend to develop this line of research with the ambitious aim to obtain decidability for all natural systems and specifications. More precisely, we will identify natural hypotheses both on the architecture of our distributed system and on the specifications under which the distributed program/controller synthesis problem is decidable. This should open the way to important applications, e.g., for distributed control of embedded systems.

3.2.3. *Adaptation and Grey box management*

Participants: Stefan Haar, Serge Haddad.

Contrary to mainframe systems or monolithic applications of the past, we are experiencing and using an increasing number of services that are performed not by one provider but rather by the interaction and cooperation of many specialized components. As these components come from different providers, one can no longer assume all of their internal technologies to be known (as it is the case with proprietary technology). Thus, in order to compose e.g. orchestrated services over the web, to determine violations of specifications or contracts, to adapt existing services to new situations etc, one needs to analyze the interaction behavior of *boxes* that are known only through their public interfaces. For their semi-transparent-semi-opaque nature, we shall refer to them as **grey boxes**. While the concrete nature of these boxes can range from vehicles in a highway section to hotel reservation systems, the tasks of *grey box management* have universal features allowing for generalized approaches with formal methods. Two central issues emerge:

- Abstraction: From the designer point of view, there is a need for a trade-off between transparency (no abstraction) in order to integrate the box in different contexts and opacity (full abstraction) for security reasons.

- Adaptation: Since a grey box gives a partial view about the behavior of the component, even if it is not immediately useable in some context, the design of an adaptator is possible. Thus the goal is the synthesis of such an adaptator from a formal specification of the component and the environment.

Our work on direct modeling and handling of "grey boxes" via modal models (see [49]) was halted when Dorsaf El-Hog stopped her PhD work to leave academia, and has not resumed for lack of staff. However, it should be noted that semi-transparent system management in a larger sense remains an active field for the team, witness in particular our work on diagnosis and testing.

3.3. Management of Quantitative Behavior

Participants: Benedikt Bollig, Thomas Chatain, Paul Gastin, Stefan Haar, Serge Haddad.

3.3.1. *Introduction*

Besides the logical functionalities of programs, the *quantitative* aspects of component behavior and interaction play an increasingly important role.

- *Real-time* properties cannot be neglected even if time is not an explicit functional issue, since transmission delays, parallelism, etc, can lead to time-outs striking, and thus change even the logical course of processes. Again, this phenomenon arises in telecommunications and web services, but also in transport systems.

- In the same contexts, *probabilities* need to be taken into account, for many diverse reasons such as unpredictable functionalities, or because the outcome of a computation may be governed by race conditions.

- Last but not least, constraints on *cost* cannot be ignored, be it in terms of money or any other limited resource, such as memory space or available CPU time.

Traditional mainframe systems were proprietary and (essentially) localized; therefore, impact of delays, unforeseen failures, etc. could be considered under the control of the system manager. It was therefore natural, in verification and control of systems, to focus on *functional* behavior entirely.

With the increase in size of computing system and the growing degree of compositionality and distribution, quantitative factors enter the stage:

- calling remote services and transmitting data over the web creates *delays*;

- remote or non-proprietary components are not "deterministic", in the sense that their behavior is uncertain.

Time and *probability* are thus parameters that management of distributed systems must be able to handle; along with both, the *cost* of operations is often subject to restrictions, or its minimization is at least desired. The mathematical treatment of these features in distributed systems is an important challenge, which *MExICo* is addressing; the following describes our activities concerning probabilistic and timed systems. Note that cost optimization is not a current activity but enters the picture in several intended activities.

3.3.2. *Probabilistic distributed Systems*

Participants: Stefan Haar, Serge Haddad, Claudine Picaronny.

3.3.2.1. *Non-sequential probabilistic processes*

Practical fault diagnosis requires to select explanations of *maximal likelihood*. For partial-order based diagnosis, this leads therefore to the question what the probability of a given partially ordered execution is. In Benveniste et al. [39], [32], we presented a model of stochastic processes, whose trajectories are partially ordered, based on local branching in Petri net unfoldings; an alternative and complementary model based on Markov fields is developed in [57], which takes a different view on the semantics and overcomes the first model's restrictions on applicability.

Both approaches abstract away from real time progress and randomize choices in *logical* time. On the other hand, the relative speed - and thus, indirectly, the real-time behavior of the system's local processes - are crucial factors determining the outcome of probabilistic choices, even if non-determinism is absent from the system.

In another line of research [43] we have studied the likelihood of occurrence of non-sequential runs under random durations in a stochastic Petri net setting. It remains to better understand the properties of the probability measures thus obtained, to relate them with the models in logical time, and exploit them e.g. in *diagnosis*.

3.3.2.2. *Distributed Markov Decision Processes*
Participant: Serge Haddad.

Distributed systems featuring non-deterministic and probabilistic aspects are usually hard to analyze and, more specifically, to optimize. Furthermore, high complexity theoretical lower bounds have been established for models like partially observed Markovian decision processes and distributed partially observed Markovian decision processes. We believe that these negative results are consequences of the choice of the models rather than the intrinsic complexity of problems to be solved. Thus we plan to introduce new models in which the associated optimization problems can be solved in a more efficient way. More precisely, we start by studying connection protocols weighted by costs and we look for online and offline strategies for optimizing the mean cost to achieve the protocol. We have been cooperating on this subject with the SUMO team at Inria Rennes; in the joint work [33]; there, we strive to synthesize for a given MDP a control so as to guarantee a specific stationary behavior, rather than - as is usually done - so as to maximize some reward.

3.3.3. *Large scale probabilistic systems*

Addressing large-scale probabilistic systems requires to face state explosion, due to both the discrete part and the probabilistic part of the model. In order to deal with such systems, different approaches have been proposed:

- Restricting the synchronization between the components as in queuing networks allows to express the steady-state distribution of the model by an analytical formula called a product-form [38].

- Some methods that tackle with the combinatory explosion for discrete-event systems can be generalized to stochastic systems using an appropriate theory. For instance symmetry based methods have been generalized to stochastic systems with the help of aggregation theory [46].

- At last simulation, which works as soon as a stochastic operational semantic is defined, has been adapted to perform statistical model checking. Roughly speaking, it consists to produce a confidence interval for the probability that a random path fulfills a formula of some temporal logic [67] .

We want to contribute to these three axes: (1) we are looking for product-forms related to systems where synchronization are more involved (like in Petri nets), see [2]; (2) we want to adapt methods for discrete-event systems that require some theoretical developments in the stochastic framework and, (3) we plan to address some important limitations of statistical model checking like the expressiveness of the associated logic and the handling of rare events.

3.3.4. *Real time distributed systems*

Nowadays, software systems largely depend on complex timing constraints and usually consist of many interacting local components. Among them, railway crossings, traffic control units, mobile phones, computer servers, and many more safety-critical systems are subject to particular quality standards. It is therefore becoming increasingly important to look at networks of timed systems, which allow real-time systems to operate in a distributed manner.

Timed automata are a well-studied formalism to describe reactive systems that come with timing constraints. For modeling distributed real-time systems, networks of timed automata have been considered, where the local clocks of the processes usually evolve at the same rate [61] [44]. It is, however, not always adequate to assume that distributed components of a system obey a global time. Actually, there is generally no reason to assume that different timed systems in the networks refer to the same time or evolve at the same rate. Any component is rather determined by local influences such as temperature and workload.

3.3.4.1. *Implementation of Real-Time Concurrent Systems*
Participants: Thomas Chatain, Stefan Haar, Serge Haddad.

This was one of the tasks of the ANR ImpRo.

Formal models for real-time systems, like timed automata and time Petri nets, have been extensively studied and have proved their interest for the verification of real-time systems. On the other hand, the question of using these models as specifications for designing real-time systems raises some difficulties. One of those comes from the fact that the real-time constraints introduce some artifacts and because of them some syntactically correct models have a formal semantics that is clearly unrealistic. One famous situation is the case of Zeno executions, where the formal semantics allows the system to do infinitely many actions in finite time. But there are other problems, and some of them are related to the distributed nature of the system. These are the ones we address here.

One approach to implementability problems is to formalize either syntactical or behavioral requirements about what should be considered as a reasonable model, and reject other models. Another approach is to adapt the formal semantics such that only realistic behaviors are considered.

These techniques are preliminaries for dealing with the problem of implementability of models. Indeed implementing a model may be possible at the cost of some transformation, which make it suitable for the target device. By the way these transformations may be of interest for the designer who can now use high-level features in a model of a system or protocol, and rely on the transformation to make it implementable.

We aim at formalizing and automating translations that preserve both the timed semantics and the concurrent semantics. This effort is crucial for extending concurrency-oriented methods for logical time, in particular for exploiting partial order properties. In fact, validation and management - in a broad sense - of distributed systems is not realistic *in general* without understanding and control of their real-time dependent features; the link between real-time and logical-time behaviors is thus crucial for many aspects of *MExICo*'s work.

3.3.5. *Weighted Automata and Weighted Logics*
Participants: Benedikt Bollig, Paul Gastin.

Time and probability are only two facets of quantitative phenomena. A generic concept of adding weights to qualitative systems is provided by the theory of weighted automata [31]. They allow one to treat probabilistic or also reward models in a unified framework. Unlike finite automata, which are based on the Boolean semiring, weighted automata build on more general structures such as the natural or real numbers (equipped with the usual addition and multiplication) or the probabilistic semiring. Hence, a weighted automaton associates with

any possible behavior a weight beyond the usual Boolean classification of "acceptance" or "non-acceptance". Automata with weights have produced a well-established theory and come, e.g., with a characterization in terms of rational expressions, which generalizes the famous theorem of Kleene in the unweighted setting. Equipped with a solid theoretical basis, weighted automata finally found their way into numerous application areas such as natural language processing and speech recognition, or digital image compression.

What is still missing in the theory of weighted automata are satisfactory connections with verification-related issues such as (temporal) logic and bisimulation that could lead to a general approach to corresponding satisfiability and model-checking problems. A first step towards a more satisfactory theory of weighted systems was done in [42]. That paper, however, does not give definite answers to all the aforementioned problems. It identifies directions for future research that we will be tackling.

4. Application Domains

4.1. Telecommunications

Participants: Stefan Haar, Serge Haddad.

MExICo's research is motivated by problems of *system management* in several domains, such as:

- In the domain of service oriented computing, it is often necessary to insert some Web service into an existing orchestrated business process, e.g. to replace another component after failures. This requires to ensure, often actively, conformance to the interaction protocol. One therefore needs to synthesize adaptators for every component in order to steer its interaction with the surrounding processes.

- Still in the domain of telecommunications, the supervision of a network tends to move from out-of-band technology, with a fixed dedicated supervision infrastructure, to in-band supervision where the supervision process uses the supervised network itself. This new setting requires to revisit the existing supervision techniques using control and diagnosis tools.

Currently, we have no active cooperation on these subjects.

4.2. Transport Systems

Participants: Stefan Haar, Serge Haddad, Yann Duplouy, Simon Theissing.

We participate in the IRT System X's system of systems program TMM, in two projects:

- project MIC (terminated in November 2016) on multi-modal transport systems with academic partners UPMC, IFSTTAR and CEA, and several industrial partners including Alstom (project leader), COSMO and Renault. Transportation operators in an urban area need to plan, supervise and steer different means of transportation with respect to several criteria:
  - Maximize capacity;
  - guarantee punctuality and robustness of service;
  - minimize energy consumption.

  The systems must achieve these objectives not only under ideal conditions, but also be robust to perturbations (such as a major cultural or sport event creating additional traffic), modifications of routes (roadwork, accidents, demonstrations, ...) and tolerant to technical failures. Therefore, systems must be enabled to raise appropriate alarms upon detection of anomalies, diagnose the type of anomaly and select the appropriate response. While the above challenges belong already to the tasks of individual operators in the unimodal setting, the rise of and increasing demand for multi-modal transports forces to achieve these planning, optimization and control goals not in isolation, but in a cooperative manner, across several operators. The research task here is first to analyze the transportation system regarding the available means, capacities and structures, and so as to identify the impacting factors and interdependencies of the system variables. Based on this analysis, the

task is to derive and implement robust planning, with tolerance to technical faults; diagnosis and control strategies that are optimal under several, possibly different, criteria (average case vs worst case performance, energy efficiency, etc.) and allow to adapt to changes e.g. from nominal mode to reduced mode, sensor failures, etc.

- the project SVA (Simulation pour la Sécurité du Véhicule Autonome), where the PhD Thesis of Yann Duplouy targets the application of formal methods to the development of embedded systems for autonomous vehicles.

4.3. Biological Systems

Participants: Thomas Chatain, Stefan Haar, Serge Haddad, Stefan Schwoon.

We have begun in 2014 to examine concurrency issues in systems biology, and are currently enlarging the scope of our research's applications in this direction. To see the context, note that in recent years, a considerable shift of biologists' interest can be observed, from the mapping of static genotypes to gene expression, i.e. the processes in which genetic information is used in producing functional products. These processes are far from being uniquely determined by the gene itself, or even jointly with static properties of the environment; rather, regulation occurs throughout the expression processes, with specific mechanisms increasing or decreasing the production of various products, and thus modulating the outcome. These regulations are central in understanding cell fate (how does the cell differenciate ? Do mutations occur ? etc), and progress there hinges on our capacity to analyse, predict, monitor and control complex and variegated processes. We have applied Petri net unfolding techniques for the efficient computation of attractors in a regulatory network; that is, to identify strongly connected reachability components that correspond to stable evolutions, e.g. of a cell that differentiates into a specific functionality (or mutation). This constitutes the starting point of a broader research with Petri net unfolding techniques in regulation. In fact, ,he use of ordinary Petri nets for capturing regulatory network (RN) dynamics overcomes the limitations of traditional RN models : those impose e.g. Monotonicity properties in the influence that one factor had upon another, i.e. always increasing or always decreasing, and were thus unable to cover all actual behaviours (see [75]). Rather, we follow the more refined model of boolean networks of automata, where the local states of the different factors jointly detemine which state transitions are possible. For these connectors, ordinary PNs constitute a first approximation, improving greatly over the literature but leaving room for improvement in terms of introducing more refined logical connectors. Future work thus involves transcending this class of PN models. Via unfoldings, one has access – provided efficient techniques are available – to all behaviours of the model, rather than over-or under-approximations as previously. This opens the way to efficiently searching in particular for determinants of the cell fate : which attractors are reachable from a given stage, and what are the factors that decide in favor of one or the other attractor, etc. The list of potential applications in biology and medicine of such a methodology would be too long to reproduce here.

5. Highlights of the Year

5.1. Highlights of the Year

Diagnosis, Anti-alignments and Coverability

DIAGNOSIS

Several new advances were obtained, concerning Diagnosis in Infinite-State Probabilistic Systems, Approximate Diagnosability of Stochastic Systems, and Diagnosability of Repairable Faults; see the 'New Results' section for a detailed description.

ANTI-ALIGNMENTS IN CONFORMANCE CHECKING – THE DARK SIDE OF PROCESS MODELS

Conformance checking techniques asses the suitability of a process model in representing an underlying process, observed through a collection of real executions. These techniques suffer from the well-known state space explosion problem, hence handling process models exhibiting large or even infinite state spaces remains a challenge. One important metric in conformance checking is to asses the precision of the model with respect to the observed executions, i.e., characterize the ability of the model to produce behavior unrelated to the one observed. By avoiding the computation of the full state space of a model, current techniques only provide estimations of the precision metric, which in some situations tend to be very optimistic, thus hiding real problems a process model may have. In [15], [25] we present the notion of anti-alignment as a concept to help unveiling traces in the model that may deviate significantly from the observed behavior. Using anti-alignments, current estimations can be improved, e.g., in precision checking. We show how to express the problem of finding anti-alignments as the satisfiability of a Boolean formula, and provide a tool which can deal with large models efficiently. In [19], [20], a novel approach to measure precision and generalization is presented, which relies on the notion of anti-alignments. We propose metrics for precision and generalization that resemble the leave-one-out cross-validation techniques, where individual traces of the log are removed and the computed anti-alignment assess the model's capability to describe precisely or generalize the observed behavior.

APPROACHING THE COVERABILITY PROBLEM CONTINUOUSLY

The coverability problem for Petri nets plays a central role in the verification of concurrent shared-memory programs. However, its high EXPSPACE-complete complexity poses a challenge when encountered in real-world instances. In [13], we develop a new approach to this problem which is primarily based on applying forward coverability in continuous Petri nets as a pruning criterion inside a backward coverability framework. A cornerstone of our approach is the efficient encoding of a recently developed polynomial-time algorithm for reachability in continuous Petri nets into SMT. We demonstrate the effectiveness of our approach on standard benchmarks from the literature, which shows that our approach decides significantly more instances than any existing tool and is in addition often much faster, in particular on large instances.

6. New Software and Platforms

6.1. DarkSider

FUNCTIONAL DESCRIPTION

DarkSider computes anti-alignments between a Petri net model and a log of observed traces, as described in [15], [25].
- Participant: Thomas Chatain
- Contact: Thomas Chatain
- URL: http://www.lsv.ens-cachan.fr/~chatain/darksider/

6.2. COSMOS

FUNCTIONAL DESCRIPTION

COSMOS is a statistical model checker for the Hybrid Automata Stochastic Logic (HASL). HASL employs Linear Hybrid Automata (LHA), a generalization of Deterministic Timed Automata (DTA), to describe accepting execution paths of a Discrete Event Stochastic Process (DESP), a class of stochastic models which includes, but is not limited to, Markov chains. As a result HASL verification turns out to be a unifying framework where sophisticated temporal reasoning is naturally blended with elaborate reward-based analysis. COSMOS takes as input a DESP (described in terms of a Generalized Stochastic Petri Net), an LHA and an expression Z representing the quantity to be estimated. It returns a confidence interval estimation of Z, recently, it has been equipped with functionalities for rare event analysis. COSMOS is written in C++
- Participants: Benoît Barbot, Hilal Djafri, Paolo Ballarini, Marie Duflot-Kremer and Serge Haddad
- Contact: Hilal Djafri
- URL: http://www.lsv.ens-cachan.fr/~barbot/cosmos/

6.3. CosyVerif

<small>FUNCTIONAL DESCRIPTION</small>

CosyVerif is a platform dedicated to the formal specification and verification of dynamic systems. It allows to specify systems using several formalisms (such as automata and Petri nets), and to run verification tools on these models.

- Participants: Serge Haddad, Fabrice Kordon, Laure Petrucci and Alban Linard
- Partners: LIP6 - LIPN (Laboratoire d'Informatique de l'Université Paris Nord) - LSV
- Contact: Serge Haddad
- URL: http://www.cosyverif.org/

6.4. Mole

<small>FUNCTIONAL DESCRIPTION</small>

Mole computes, given a safe Petri net, a finite prefix of its unfolding. It is designed to be compatible with other tools, such as PEP and the Model-Checking Kit, which are using the resulting unfolding for reachability checking and other analyses. The tool Mole arose out of earlier work on Petri nets.

- Participant: Stefan Schwoon
- Contact: Stefan Schwoon
- URL: http://www.lsv.ens-cachan.fr/~schwoon/tools/mole/

7. New Results

7.1. Analyzing Timed Systems Using Tree Automata

Timed systems, such as timed automata, are usually analyzed using their operational semantics on timed words. The classical region abstraction for timed automata reduces them to (untimed) finite state automata with the same time-abstract properties, such as state reachability. In [10], we propose a new technique to analyze such timed systems using finite tree automata instead of finite word automata. The main idea is to consider timed behaviors as graphs with matching edges capturing timing constraints. Such graphs can be interpreted in trees opening the way to tree automata based techniques which are more powerful than analysis based on word automata. The technique is quite general and applies to many timed systems. In this paper, as an example, we develop the technique on timed pushdown systems, which have recently received considerable attention. Further, we also demonstrate how we can use it on timed automata and timed multi-stack pushdown systems (with boundedness restrictions).

7.2. Interrupt Timed Automata with Auxiliary Clocks and Parameters

Interrupt Timed Automata (ITA) are an expressive timed model, introduced to take into account interruptions according to levels. Due to this feature, this formalism is incomparable with Timed Automata. However several decidability results related to reachability and model checking have been obtained. In , we add auxiliary clocks to ITA, thereby extending its expressive power while preserving decidability of reachability. Moreover, we define a parametrized version of ITA, with polynomials of parameters appearing in guards and updates. While parametric reasoning is particularly relevant for timed models, it very often leads to undecidability results. We prove that various reachability problems, including robust reachability, are decidable for this model, and we give complexity upper bounds for a fixed or variable number of clocks, levels and parameters.

7.3. One-Counter Automata with Counter Observability

In a one-counter automaton (OCA), one can produce a letter from some finite alphabet, increment and decrement the counter by one, or compare it with constants up to some threshold. It is well-known that universality and language inclusion for OCAs are undecidable. In [14], we consider OCAs with counter observability: Whenever the automaton produces a letter, it outputs the current counter value along with it. Hence, its language is now a set of words over an infinite alphabet. We show that universality and inclusion for that model are PSPACE-complete, thus no harder than the corresponding problems for finite automata. In fact, by establishing a link with visibly one-counter automata, we show that OCAs with counter observability are effectively determinizable and closed under all boolean operations.

7.4. Diagnosis in Infinite-State Probabilistic Systems

In a recent work, we introduced four variants of diagnosability (FA, IA, FF, IF) in (finite) probabilistic systems (pLTS) depending whether one considers (1) finite or infinite runs and (2) faulty or all runs. We studied their relationship and established that the corresponding decision problems are PSPACE-complete. A key ingredient of the decision procedures was a characterisation of diagnosability by the fact that a random run almost surely lies in an open set whose specification only depends on the qualitative behaviour of the pLTS. In [12], we investigate similar issues for infinite pLTS. We first show that this characterisation still holds for FF-diagnosability but with a $G\delta$ set instead of an open set and also for IF-and IA-diagnosability when pLTS are finitely branching. We also prove that surprisingly FA-diagnosability cannot be characterised in this way even in the finitely branching case. Then we apply our characterisations for a partially observable probabilistic extension of visibly pushdown automata (POpVPA), yielding EXPSPACE procedures for solving diagnosability problems. In addition, we establish some computational lower bounds and show that slight extensions of POpVPA lead to undecidability.

7.5. Accurate Approximate Diagnosability of Stochastic Systems

Diagnosis of partially observable stochastic systems prone to faults was introduced in the late nineties. Diagnosability, i.e. the existence of a diagnoser, may be specified in different ways: (1) exact diagnosability (called A-diagnosability) requires that almost surely a fault is detected and that no fault is erroneously claimed while (2) approximate diagnosability (called ε-diagnosability) allows a small probability of error when claiming a fault and (3) accurate approximate diagnosability (called AA-diagnosability) requires that this error threshold may be chosen arbitrarily small. In [11], we mainly focus on approximate diagnoses. We first refine the almost sure requirement about finite delay introducing a uniform version and showing that while it does not discriminate between the two versions of exact diagnosability this is no more the case in approximate diagnosis. Then we establish a complete picture for the decidability status of the diagnosability problems: (uniform) ε-diagnosability and uniform AA-diagnosability are undecidable while AA-diagnosability is decidable in PTIME, answering a longstanding open question.

7.6. Diagnosability of Repairable Faults

The diagnosis problem for discrete event systems consists in deciding whether some fault event occurred or not in the system, given partial observations on the run of that system. Diagnosability checks whether a correct diagnosis can be issued in bounded time after a fault, for all faulty runs of that system. This problem appeared two decades ago and numerous facets of it have been explored, mostly for permanent faults. It is known for example that diagnosability of a system can be checked in polynomial time, while the construction of a diagnoser is exponential. In [21], we examine the case of transient faults, that can appear and be repaired. Diagnosability in this setting means that the occurrence of a fault should always be detected in bounded time, but also before the fault is repaired. Checking this notion of diagnosability is proved to be PSPACE-complete. It is also shown that faults can be reliably counted provided the system is diagnosable for faults and for repairs.

7.7. Optimal constructions for active diagnosis

The task of diagnosis consists in detecting, without ambiguity, occurrence of faults in a partially observed system. Depending on the degree of observability, a discrete event system may be diagnosable or not. Active diagnosis aims at controlling the system in order to make it diagnosable. Solutions have already been proposed for the active diagnosis problem, but their complexity remains to be improved. In [8], we solve 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.

7.8. Verification of parameterized communicating automata via split-width

In [16] study verification problems for distributed systems communicating via unbounded FIFO channels. The number of processes of the system as well as the communication topology are not fixed a priori. Systems are given by parameterized communicating automata (PCAs) which can be run on any communication topology of bounded degree, with arbitrarily many processes. Such systems are Turing powerful so we concentrate on under-approximate verification. We extend the notion of split-width to behaviors of PCAs. We show that emptiness, reachability and model-checking problems of PCAs are decidable when restricted to behaviors of bounded split-width. Reachability and emptiness are EXPTIME-complete, but only polynomial in the size of the PCA. We also describe several concrete classes of bounded split-width, for which we prove similar results.

7.9. Cyclic Ordering through Partial Orders

The orientation problem for ternary cyclic order relations has been attacked in the literature from combinatorial perspectives, through rotations, and by connection with Petri nets. In [7], we propose a two-fold characterization of orientable cyclic orders in terms of symmetries of partial orders as well as in terms of separating sets (cuts). The results are inspired by properties of non-sequential discrete processeses, but also apply to dense structures of any cardinality.

7.10. Predicting Traffic Load in Public Transportation Networks

This work is part of an ongoing effort to understand the dynamics of passenger loads in modern, multimodal transportation networks (TNs) and to mitigate the impact of perturbations, under the restrictions that the precise number of passengers in some point of the TN that intend to reach a certain destination (i.e. their distribution over different trip profiles) is unknown. In [29], we introduce an approach based on a stochastic hybrid automaton model for a TN that allows to compute how such probabilistic load vectors are propagated through the TN. In [23], [30], develop a computation strategy for forecasting the network's load a certain time in the future.

In [22], [28], we continue our work on perturbation analysis of multimodal transportation networks (TNs) by means of a stochastic hybrid automaton (SHA) model. We focus here on the approximate computation , in particular on the major bottleneck consisting in the high dimensionality of systems of stochastic differential balance equations (SDEs) that define the continuous passenger-flow dynamics in the different modes of the SHA model. In fact, for every pair of a mode and a station, one system of coupled SDEs relates the passenger loads of all discrete points such as platforms considered in this station, and all vehicles docked to it, to the passenger flows in between. In general, such an SDE system has many dimensions, which makes its numerical computation and thus the approximate computation of the SHA model intractable. We show how these systems can be canonically replaced by lower-dimensional ones, by decoupling the passenger flows inside every mode from one another. We prove that the resulting approximating passenger-flow dynamics converges to the original one, if the replacing set of balance equations set up for all decoupled passenger flows communicate their results among each other in vanishing time intervals.

For more information about the whole project, see [27].

7.11. Unfolding of Parametric Logical Regulatory Networks

In systems biology, models of cellular regulatory processes such as gene regulatory networks or signalling pathways are crucial to understanding the behaviour of living cells. Available biological data are however often insufficient for full model specification. In [18], we focus on partially specified models where the missing information is abstracted in the form of parameters. We introduce a novel approach to analysis of parametric logical regulatory networks addressing both sources of combinatoric explosion native to the model. First, we introduce a new compact representation of admissible parameters using Boolean lattices. Then, we define the unfolding of parametric regulatory networks. The resulting structure provides a partial- order reduction of concurrent transitions, and factorises the common transitions among the concrete models. A comparison is performed against state-of-the-art approaches to parametric model analysis.

7.12. Relationship between the Reprogramming Determinants of Boolean Networks and their Interaction Graph

In [24], we address the formal characterization of tar- gets triggering cellular trans-differentiation in the scope of Boolean net- works with asynchronous dynamics. Given two fixed points of a Boolean network, we are interested in all the combinations of mutations which allow to switch from one fixed point to the other, either possibly, or in- evitably. In the case of existential reachability, we prove that the set of nodes to (permanently) flip are only and necessarily in certain connected components of the interaction graph. In the case of inevitable reachabil- ity, we provide an algorithm to identify a subset of possible solutions.

7.13. D-SPACES: An Implementation of Declarative Semantics for Spatially Structured Information

We introduce in [17] D-SPACES, an implementation of constraint systems with space and extrusion operators. Constraint systems are algebraic models that allow for a semantic language-like representation of information in systems where the concept of space is a primary structural feature. We give this information mainly an epistemic interpretation and consider various agents as entities acting upon it. D-SPACES is coded as a c++11 library providing implementations for constraint systems, space functions and extrusion functions. The interfaces to access each implementation are minimal and thoroughly documented. D-SPACES also provides property-checking methods as well as an implementation of a specific type of constraint systems (a boolean algebra). This last implementation serves as an entry point for quick access and proof of concept when using these models. Furthermore, we offer an illustrative example in the form of a small social network where users post their beliefs and utter their opinions.

7.14. Belief, Knowledge, Lies and Other Utterances in an Algebra for Space and Extrusion

The notion of constraint system (cs) is central to declarative formalisms from concurrency theory such as process calculi for concurrent constraint programming (ccp). Constraint systems are often represented as lattices: their elements, called constraints, represent partial information and their order corresponds to entailment. Recently a notion of n-agent spatial cs was introduced to represent information in concurrent constraint programs for spatially distributed multi-agent systems. From a computational point of view a spatial constraint system can be used to specify partial information holding in a given agent's space (local information). From an epistemic point of view a spatial cs can be used to specify information that a given agent considers true (beliefs). Spatial constraint systems, however, do not provide a mechanism for specifying the mobility of information/processes from one space to another. Information mobility is a fundamental aspect of concurrent systems. In [6] we develop the theory of spatial constraint systems with operators to specify information and processes moving from a space to another. We shall investigate the properties of

this new family of constraint systems and illustrate their applications. From a computational point of view the new operators provide for process/information extrusion, a central concept in formalisms for mobile communication. From an epistemic point of view extrusion corresponds I to a notion we shall call utterance; a piece of information that an agent communicate to others but that may be inconsistent with the agent's beliefs. Utterances can then be used to express instances of epistemic notions such as hoaxes or intentional lies which are common place in social media. Spatial constraint system can express the epistemic notion of belief by means of space functions that specify local information. We shall also show that spatial constraint can also express the epistemic notion of knowledge by means of a derived spatial operator that specifies global information.

7.15. Goal-Driven Unfolding of Petri Nets

Unfoldings provide an efficient way to avoid the state-space explosion due to interleavings of concurrent transitions when exploring the runs of a Petri net. The theory of adequate orders allows one to define finite prefixes of unfoldings which contain all the reachable markings. In this paper we are interested in reachability of a single given marking, called the goal. In [26], We propose an algorithm for computing a finite prefix of the unfolding of a 1-safe Petri net that preserves all minimal configurations reaching this goal. Our algorithm combines the unfolding technique with on-the-fly model reduction by static analysis aiming at avoiding the exploration of branches which are not needed for reaching the goal. We present some experimental results.

8. Partnerships and Cooperations

8.1. National Initiatives

We will be participating in the ANR Project ALGORECELL that starts in 2017.

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

Serge Haddad is participating in the ERC EQualIS, 'Enhancing the Quality of Interacting Systems', directed by Patricia Bouyer.

8.3. International Initiatives

8.3.1. Inria Associate Teams Not Involved in an Inria International Labs

8.3.1.1. LifeForm

Title: Life Sciences need formal Methods !

International Partner (Institution - Laboratory - Researcher):

Newcastle University (United Kingdom) - School of Computing Science - Victor Khomenko

Start year: 2016

See also: http://projects.lsv.ens-cachan.fr/LifeForm/

This project extends an existing cooperation between the MEXICO team and Newcastle University on partial-order based formal methods for concurrent systems. We enlarge the partnership to bioinformatics and synthetic biology. The proposal addresses addresses challenges concerning formal specification, verification, monitoring and control of synthetic biological systems, with use cases conducted in the Center for Synthetic Biology and the Bioeconomy (CSBB) in Newcastle. A main challenge is to create a solid modelling framework based on Petri-net type models that allow for causality analysis and rapid state space exploration for verification, monitoring and control purposes; a potential extension to be investigated concerns the study of attractors and cell reprogramming in Systems Biology.

8.3.2. Participation in Other International Programs

UMI with CMI, India, starting in 2017; currently LIA INFORMEL, see below.

8.4. International Research Visitors

8.4.1. Visits of International Scientists

- Visits by Victor Khomenko and Maciej Koutny within the LifeForm associated team

8.4.2. Internships

- **Juraj Kolčák** from Masaryk University, Brno, Czech Republic, on *Efficient Analysis of Boolean Networks under Parameter Uncertainty*, Spring/summer of 2016 (Master's thesis research); director: Stefan Haar
- **Clara Scherbaum** from Aachen University, Germany, on *Computing Cut Sets for Petri Nets*, Spring 2016, LSV (ENS Cachan),
- **Hugues Mandon**: Algorithms for cellular reprogramming.

8.4.3. Visits to International Teams

8.4.3.1. Research Stays Abroad

Paul Gastin is visiting IIT Bombay and Chennay Mathematical Institute, India, from October 10, 2016 to March 10, 2017.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. LIA INFORMEL

The Indo-French Formal Methods Lab is an International Associated Laboratory (LIA) fostering the scientific collaboration between India and France in the domain of formal methods and applications to the verification of complex systems. Our research focuses on theoretical foundations of games, automata, and logics, three important tools in formal methods. We study applications to the verification of safety-critical systems, with an emphasis on quantitative aspects (time, cost, energy, etc.), concurrency, control, and security protocols. The Laboratory was founded in 2012 by a consortium of researchers from the French Centre for Scientific Research (CNRS), Ecole Normale Supérieure de Cachan (ENS Cachan), Université Bordeaux 1, the Institute of Mathematical Sciences Chennai (IMSc), the Chennai Mathematical Institute (CMI), and the Indian Institute of Science Bangalore (IISc). It is directed by Paul Gastin (ENS Cachan, MExICo team) and Madhavan Mukund (CMI). The LIA has been scientifically extremely active and productive since its creation. The LIA has supported numerous scientific exchanges and joint research papers, see here. Among many other activities, the LIA organised another edition of the ACTS workshop.

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

- Thomas Chatain was a member of the program committee of *(ACSD 2016)*.
- Matthias Függer was a member of the PCs of DDECS'16 and ASYNC'16.
- Stefan Haar was a member of the PCs of *13th International Workshop on Discrete Event SystemsWODES 2016*, the *16th International Conference on Applications of Concurrency to Systems Design(ACSD 2016)*, *Int. WS on Petri Nets and Software Engineering PNSE 2016*, ATAED Workshop on Analysis of Event Data 2016, and *IEEE Int. Conf. on Emerging Technologies and Factory Automation(ETFA)* 2016.
- Serge Haddad was a member of the PC of the 10th International Workshop on Verification and Evaluation of Computer and Communication Systems (VECOS 2016),Tunis, Tunisia.

- Stefan Schwoon was a member of the PC of the 37th International Conference on Applications and Theory of Petri Nets and Concurrency (PN 2016).
- Claudine Picaronny was a PC member for the Eighth International Conference on Advances in System Simulation (SIMUL'16)

9.1.2.2. Reviewer

- Matthias Függer was a reviewer for ICALP, ASYNC, DISC, DDECS, and IPDPS.
- Stefan HAAR was a reviewer for MFCS 2016.
- Stefan Schwoon acted as a reviewer for the following conferences taking place in 2016 : TACAS, ACSD, CONCUR, FSTTCS.

9.1.3. Journal

9.1.3.1. Member of Editorial Boards

- Stefan Haar is an associate editor of the *Journal of Discrete Event Dynamic Systems: Theory and Applications*, and a guest editor (with R. Meyer) of the upcoming special issue on ACSD 2015 in *ACM Transactions on Embedded Computing Systems (TECS)*.

9.1.3.2. Reviewer - Reviewing Activities

- Matthias Függer was a reviewer for the Journal *Energies*.
- Stefan Haar was a reviewer for *LMCS*, *MSCS*, *IEEE Transactions on Automatic Control* and *Journal of Discrete Event Dynamic Systems*.
- Stefan Schwoon acted as a reviewer for the following journals in 2016 : Fundamenta Informaticae, Transactions on Software Engineering.

9.1.4. Invited Talks

- Serge Haddad gave the following invited talks:
    – at the Joint AFSEC/ANR PACS workshop on May 26, 2016, Paris, France, on "Polynomial Interrupt Timed Automata";
    – at the VECOS 2016 conference, Tunis, Tunisia, on October 6, 2016, "Active Diagnosis";
    – at IDC 2016 (10th International Symposium on Intelligent Distributed Computing), October 11, 2016, Paris, France, on "Fault Diagnosis in Probabilistic Systems".
- Benedikt Bollig gave an invited tutorial at Highlights, Brussels, Belgium, 2016, on Automata and Logics for Distributed Systems

9.1.5. Research Administration

- Paul Gastin is one of the directors of the LIA INFORMEL.
- Stefan Haar is the head of the *SCILEX* axis within the *DIGICOSME* Labex. He was the Inria center of Saclay's correspondent for european partnerships until the summer of 2017, when he stepped down from this position to accept the presidency of Inria's COST-GTRI (international relations working group).
- Serge Haddad was a member of the recruitment committee for a professorship at INSA Toulouse.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Serge Haddad and Paul Gastin are professors at ENS Cachan (now ENS Paris-Saclay), Claudine Picaronny, Thomas Chatain and Stefan Schwoon are associate professors of the same university. Serge Haddad is the head of the Computer Science Department, and Stefan Schwoon is in charge of the L3 class. Claudine Picaronny is a co-director of the ENS Paris-Saclay's Mathematics department and a member of the juries of 'l'agrégation interne de Mathématiques' and of the second 'concours de Mathématiques' of ENS Cachan; she is also the coordinator of the mathematics/computer science examination of E3A, parts MP and MC.

Master : Benedikt Bollig, Non-sequential Theory of Distributed Systems, 36, M2, MPRI, ENS Cachan, France.

9.2.2. Supervision

Defended theses:

- PhD ([3])by Salim Perchy , 'Opinions, Lies and Knowledge. An Algebraic Approach to Mobility of Information and Processes', E´cole Polytechnique, defended October 4, supervised by Stefan Haar and Franck Valencia (COMETE team).
- PhD by Simon Theissing [4], 'Supervision for Multimodal Tranport Systems', ENS Cachan, defended December 5, supervised by Stefan Haar.

PhD in progress:

- Tymofii PROKOPENKO, Ecole Polytechnique since Oct 1, 'Privacy', jointly supervised by Catuscia Palamidessi (COMETE team) and Serge Haddad;
- Engel Lefaucheux, ENS Paris-Saclay since 2015, 'controlling information in probabilistic systems', jointly supervised by Nathalie Bertrand (SUMO team) and Serge Haddad
- Yann Duplouy, ENS Paris-Saclay since 2015,'application of formal methods to the development of embedded systems for autonomous vehicles', supervised by Béatrice Bérard and Serge Haddad. Marie Fortin (ENS Paris-Saclay since Oct 1); 'Tree-automata techniques for the analysis of distributed systems', co-supervised by Benedikt Bollig and Paul Gastin.
- Hugues Mandon (ENS Paris-Saclay since Oct 1, Digicosme Grant), Computational Models and Algorithms for the Prediction of Cell Reprogramming Strategies; supervised by Stefan Haar, co-supervision by Loic Paulevé (LRI).
- Robert Najvirt (TU Wien, Austrian FWF SIC project), *realistic delay models with applications in high-speed and low-power circuits*, co-supervised by Matthias Függer and Andreas Steininger.
- Martin Perner (TU Wien, Austrian FWF SIC project), *clock generation on-chip and formalisms suitable to prove correct VLSI circuits*, co-supervised by Matthias Függer and Ulrich Schmid.
- Juergen Maier (TU Wien, Austrian FWF SIC project), *on realistic delay models with applications in high-speed and low-power circuits, with focus on noise and high-order models*, co-supervised by Matthias Függer and with Ulrich Schmid.

9.2.3. Juries

- Benedikt bollig was
  – reviewer and jury member of the PhD thesis Logics on Data Words: Expressivity, Satisfiability, Model Checking by Ahmet Kara (Supervisor: Thomas Schwentick), Universität Dortmund, Germany, 2016, and
  – Reviewer of the PhD thesis Probabilistic Logic, Probabilistic Regular Expressions, and Constraint Temporal Logic by Thomas Weidner (Supervisor: Manfred Droste), Universität Leipzig, Germany, 2016
- Thomas Chatain was a member of the jury for the PhD defense of María Martos-Salgado, Universidad Complutense de Madrid, in January 2016.
- In addition to the juries of the two supervised students, Stefan Haar was the president of the jury for the PhD of Hassan Ibrahim, on 'SAT-based Diagnosability and Predictability Analysis in Centralized and Distributed Discrete Event Systems' at Université Paris-Sud on December 16.
- Serge Haddad was
  – a member of the juries for the PhD of Amira Methni on 'Méthodes de vérification de logiciel système critique", on July 7, 2016, at CNAM,
  – the president of the PhD jury for Hadrien Bride on "Verifying Modal Specifications of Workflow Nets" on October 24, 2016, at Université de Franche-Comté, and

– a member of the HdR jury for Yann Thierry-Mieg, "From Symbolic Verification To
Domain Specific Languages", on December 7, 2016, at Université Paris 6.

9.3. Popularization

- Stefan Haar gave a talk entitled 'Post hoc sed non propter hoc, or: why you should care about
causality', in the Seminar@SystemX series of IRT SystemX on September 14, 2016.

10. Bibliography

Major publications by the team in recent years

[1] B. BARBOT, S. HADDAD, C. PICARONNY.*Coupling and Importance Sampling for Statistical Model Checking*,
in "Proceedings of the 18th International Conference on Tools and Algorithms for Construction and Analysis
of Systems (TACAS'12)", Tallinn, Estonia, C. FLANAGAN, B. KÖNIG (editors), Lecture Notes in Computer
Science, Springer, March 2012, vol. 7214, p. 331-346, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/
BHP-tacas12.pdf.

[2] S. HADDAD, J. MAIRESSE, H.-T. NGUYEN.*Synthesis and Analysis of Product-form Petri Nets*, in "Funda-
menta Informaticae", 2013, vol. 122, n° 1-2, p. 147-172, https://hal.archives-ouvertes.fr/hal-00925774.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[3] S. PERCHY.*Opinions, Lies and Knowledge. An Algebraic Approach to Mobility of Information and Processes*,
Université Paris-Saclay, October 2016, https://hal.inria.fr/tel-01413970.

[4] S. THEISSING.*Supervision in Multimodal Transportation Systems*, Université Paris-Saclay, December 2016,
https://hal.inria.fr/tel-01419126.

Articles in International Peer-Reviewed Journal

[5] É. ANDRÉ, T. CHATAIN, C. RODRIGUEZ.*Preserving Partial Order Runs in Parametric Time Petri
Nets*, in "ACM Transactions on Embedded Computing Systems (TECS)", 2016, vol. 16, 25
[*DOI : * 10.1145/3012283], https://hal.inria.fr/hal-01425696.

[6] M. GUZMAN, S. HAAR, S. PERCHY, C. RUEDA, F. VALENCIA.*Belief, Knowledge, Lies and Other Utterances
in an Algebra for Space and Extrusion*, in "Journal of Logical and Algebraic Methods in Programming",
September 2016 [*DOI : * 10.1016/J.JLAMP.2016.09.001], https://hal.inria.fr/hal-01257113.

[7] S. HAAR.*Cyclic Ordering through Partial Orders **, in "Journal of Multivalued-Logic and Soft Computing",
September 2016, vol. 27, n° 2-3, p. 209-228, https://hal.inria.fr/hal-01360144.

[8] S. HAAR, S. HADDAD, T. MELLITI, S. SCHWOON.*Optimal constructions for active diagnosis*, in "Journal
of Computer and System Sciences", 2017, vol. 83, n° 1, p. 101-120 [*DOI : * 10.1016/J.JCSS.2016.04.007],
https://hal.archives-ouvertes.fr/hal-01408047.

[9] F. KORDON, H. GARAVEL, L. M. HILLAH, E. PAVIOT-ADET, L. JEZEQUEL, C. RODRÍGUEZ, F. HULIN-HUBARD.*MCC'2015 – The Fifth Model Checking Contest*, in "Transactions on Petri Nets and Other Models of Concurrency", 2016, vol. 9930, p. 262-273 [*DOI :* 10.1007/978-3-662-53401-4_12], https://hal.inria.fr/hal-01361274.

International Conferences with Proceedings

[10] S. AKSHAY, P. GASTIN, S. N. KRISHNA.*Analyzing Timed Systems Using Tree Automata*, in "27th International Conference on Concurrency Theory (CONCUR 2016)", Québec City, Canada, Proceedings of the 27th International Conference on Concurrency Theory (CONCUR 2016), Leibniz-Zentrum für Informatik, 2016, vol. 59, p. 27:1–27:14 [*DOI :* 10.4230/LIPIcs.CONCUR.2016.27], https://hal.archives-ouvertes.fr/hal-01407942.

[11] N. BERTRAND, S. HADDAD, E. LEFAUCHEUX.*Accurate approximate diagnosability of stochastic systems*, in "10th International Conference on Language and Automata Theory and Applications", Prague, Czech Republic, Springer, March 2016, https://hal.inria.fr/hal-01220954.

[12] N. BERTRAND, S. HADDAD, E. LEFAUCHEUX.*Diagnosis in Infinite-State Probabilistic Systems* , in "27th International Conference on Concurrency Theory (Concur 2016)", Québec city, Canada, 27th International Conference on Concurrency Theory (Concur 2016), August 2016 [*DOI :* 10.4230/LIPIcs.CONCUR.2016.37], https://hal.inria.fr/hal-01373354.

[13] M. BLONDIN, A. FINKEL, C. HAASE, S. HADDAD.*Approaching the Coverability Problem Continuously*, in "22nd International Conference on Tools and Algorithms for Construction and Analysis of Systems (TACAS 2016)", Eindhoven, Netherlands, Proceedings of the 22nd International Conference on Tools and Algorithms for Construction and Analysis of Systems (TACAS 2016), 2016, vol. LNCS 9636, p. 480-496, https://hal.archives-ouvertes.fr/hal-01408044.

[14] B. BOLLIG.*One-Counter Automata with Counter Observability*, in "36th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2016)", Chennai, India, Proceedings of the 36th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2016), December 2016 [*DOI :* 10.4230/LIPIcs.FSTTCS.2016], https://hal.archives-ouvertes.fr/hal-01407932.

[15] T. CHATAIN, J. CARMONA.*Anti-Alignments in Conformance Checking - The Dark Side of Process Models*, in "37th International Conference on Applications and Theory of Petri Nets (PETRI NETS 2016)", Torún, Poland, Proceedings of the 37th International Conference on Applications and Theory of Petri Nets (PETRI NETS 2016), 2016, vol. LNCS 9698, p. 240-258, https://hal.archives-ouvertes.fr/hal-01408043.

[16] M. FORTIN, P. GASTIN.*Verification of parameterized communicating automata via split-width*, in "19th International Conference on Foundations of Software Science and Computation Structures (FoSSaCS 2016)", Eindhoven, Netherlands, Proceedings of the 19th International Conference on Foundations of Software Science and Computation Structures (FoSSaCS 2016), 2016, vol. LNCS 9634, p. 197-213, https://hal.archives-ouvertes.fr/hal-01408041.

[17] S. HAAR, S. PERCHY, F. VALENCIA.*D-SPACES: Implementing Declarative Semantics for Spatially Structured Information*, in "11th International Conference on Semantic Computing", San Diego, California, United States, IEEE ICSC 2017, IEEE, January 2017, vol. 11, https://hal.inria.fr/hal-01328189.

[18] J. KOLČÁK, D. ŠAFRÁNEK, S. HAAR, L. PAULEVÉ.*Unfolding of Parametric Logical Regulatory Networks*, in "The Seventh International Workshop on Static Analysis and Systems Biology (SASB 2016)", Edimbourg, United Kingdom, Electronic Notes in Theoretical Computer Science, Elsevier, September 2016, https://hal. archives-ouvertes.fr/hal-01354109.

[19] B. F. VAN DONGEN, J. CARMONA, T. F. CHATAIN.*A Unified Approach for Measuring Precision and Generalization Based on Anti-alignments*, in "14th International Conference on Business Process Management (BPM'16)", Rio de Janeiro, Brazil, Proceedings of the 14th International Conference on Business Process Management (BPM'16), September 2016, p. 39 - 56 [*DOI :* 10.1007/978-3-319-45348-4_3], https://hal. archives-ouvertes.fr/hal-01406850.

[20] B. VAN DONGEN, J. CARMONA, T. CHATAIN.*Alignment-based Quality Metrics in Conformance Checking*, in "7th International Workshop on Enterprise Modelling and Information Systems Architectures (EMISA'16)", Vienna, Austria, Proceedings of the 7th Int. Workshop on Enterprise Modelling and Information Systems Architectures (EMISA'16), October 2016, https://hal.archives-ouvertes.fr/hal-01406813.

Conferences without Proceedings

[21] E. FABRE, L. HÉLOUËT, E. LEFAUCHEUX, H. MARCHAND.*Diagnosability of Repairable Faults*, in "13th International Workshop on Discrete Event Systems", Xi'an, China, 2016, p. 256-262, (Version Longue), https://hal.inria.fr/hal-01302562.

[22] S. HAAR, S. THEISSING.*Decoupling Passenger Flows for Improved Load Prediction*, in "13th International Conference on Quantitative Evaluation of SysTems (QEST 2016)", Québec City, Canada, August 2016, https:// hal.inria.fr/hal-01330136.

[23] S. HAAR, S. THEISSING.*Predicting Traffic Load in Public Transportation Networks*, in "2016 American Control Conference", Boston, United States, July 2016, https://hal.inria.fr/hal-01329632.

[24] H. MANDON, S. HAAR, L. PAULEVÉ.*Relationship between the Reprogramming Determinants of Boolean Networks and their Interaction Graph*, in "Fifth International Workshop on Hybrid Systems Biology (HSB 2016)", Grenoble, France, E. CINQUEMANI, A. DONZÉ (editors), Lecture Notes in Computer Science, Springer International Publishing, October 2016, vol. 9957, p. 113-127 [*DOI :* 10.1007/978-3-319-47151-8_8], https://hal.archives-ouvertes.fr/hal-01354079.

Research Reports

[25] T. CHATAIN, J. CARMONA.*Anti-Alignments in Conformance Checking – The Dark Side of Process Models*, LSV, ENS Cachan, CNRS, Inria, Université Paris-Saclay, Cachan (France) ; Universitat Politècnica de Catalunya, Barcelona (Spain), January 2016, https://hal.inria.fr/hal-01267015.

Other Publications

[26] T. CHATAIN, L. PAULEVÉ.*Goal-Driven Unfolding of Petri Nets*, October 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01392203.

[27] S. HAAR, S. THEISSING.*A Passenger-centric Multi-agent System Model for Multimodal Public Transportation*, May 2016, working paper or preprint, https://hal.inria.fr/hal-01322956.

[28] S. HAAR, S. THEISSING.*Decoupling Passenger Flows for Improved Load Prediction*, March 2016, working paper or preprint, https://hal.inria.fr/hal-01294498.

[29] S. HAAR, S. THEISSING.*Forecasting Passenger Loads in Transportation Networks*, January 2016, working paper or preprint, https://hal.inria.fr/hal-01259585.

[30] S. HAAR, S. THEISSING.*Predicting traffic load in public transportation networks*, March 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01286476.

References in notes

[31] W. KUICH, H. VOGLER, M. DROSTE (editors). *Handbook of Weighted Automata*, EATCS Monographs in Theoretical Computer Science, Springer, 2009.

[32] S. ABBES, A. BENVENISTE, S. HAAR.*A Petri net model for distributed estimation*, in "Proc. MTNS 2004, Sixteenth International Symposium on Mathematical Theory of Networks and Systems, Louvain (Belgium), ISBN 90-5682-517-8", 2004.

[33] S. AKSHAY, N. BERTRAND, S. HADDAD, L. HELOUET.*The steady-state control problem for Markov decision processes*, in "Qest 2013", Buenos Aires, Argentina, K. R. JOSHI, M. SIEGLE, M. STOELINGA, P. R. D'ARGENIO (editors), Springer, September 2013, vol. 8054, p. 290-304, https://hal.inria.fr/hal-00879355.

[34] R. ALUR, K. ETESSAMI, M. YANNAKAKIS.*Realizability and Verification of MSC Graphs*, in "Theor. Comput. Sci.", 2005, vol. 331, n° 1, p. 97–114.

[35] P. BALDAN, TH. CHATAIN, S. HAAR, B. KÖNIG.*Unfolding-based Diagnosis of Systems with an Evolving Topology*, in "Information and Computation", October 2010, vol. 208, n° 10, p. 1169-1192, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/BCHK-icomp10.pdf.

[36] P. BALDAN, A. CORRADINI, B. König, S. SCHWOON.*McMillan's complete prefix for contextual nets*, in "Transactions on Petri Nets and Other Models of Concurrency", November 2008, vol. 1, p. 199–220, Volume 5100 of Lecture Notes in Computer Science.

[37] P. BALDAN, S. HAAR, B. KOENIG.*Distributed Unfolding of Petri Nets*, in "Proc.FOSSACS 2006", LNCS, Springer, 2006, vol. 3921, p. 126-141, Extended version: Technical Report CS-2006-1. Department of Computer Science, University Ca' Foscari of Venice.

[38] F. BASKETT, K. M. CHANDY, R. R. MUNTZ, F. G. PALACIOS.*Open, Closed, and Mixed Networks of Queues with Different Classes of Customers*, in "J. ACM", April 1975, vol. 22, p. 248–260, http://doi.acm.org/10.1145/321879.321887.

[39] A. BENVENISTE, É. FABRE, S. HAAR.*Markov Nets: Probabilistic Models for distributed and concurrent Systems*, in "IEEE Transactions on Automatic Control", 2003, vol. 48 (11), p. 1936-1950, Extended version: IRISA Research Report 1538.

[40] P. BHATEJA, P. GASTIN, M. MUKUND, K. NARAYAN KUMAR.*Local testing of message sequence charts is difficult*, in "Proceedings of the 16th International Symposium on Fundamentals of Computation Theory (FCT'07)", Budapest, Hungary, E. CSUHAJ-VARJÚ, Z. ÉSIK (editors), Lecture Notes in Computer Science,

Springer, August 2007, vol. 4639, p. 76-87 [*DOI :* 10.1007/978-3-540-74240-1_8], http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/BGMN-fct07.pdf.

[41] G. V. BOCHMANN, S. HAAR, C. JARD, G.-V. JOURDAN.*Testing Systems Specified as Partial Order Input/Output Automata*, in "Proc. TESTCOM/Fates 08, 20th IFIP International Conference on Testing of Communicating Systems and 8th International Workshop on Formal Approaches to Testing of Software", LNCS, Springer, 2008, vol. 5047, p. 169-183.

[42] B. BOLLIG, P. GASTIN.*Weighted versus Probabilistic Logics*, in "Proceedings of the 13th International Conference on Developments in Language Theory (DLT'09)", Stuttgart, Germany, V. DIEKERT, D. NOWOTKA (editors), Lecture Notes in Computer Science, Springer, June-July 2009, vol. 5583, p. 18-38 [*DOI :* 10.1007/978-3-642-02737-6_2], http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/BG-dlt09.pdf.

[43] A. BOUILLARD, S. HAAR, S. ROSARIO.*Critical paths in the Partial Order Unfolding of a Stochastic Petri Net*, in "Proceedings of the 7th International Conference on Formal Modelling and Analysis of Timed Systems (FORMATS'09)", Budapest, Hungary, J. OUAKNINE, F. VAANDRAGER (editors), Lecture Notes in Computer Science, Springer, September 2009, vol. 5813, p. 43-57 [*DOI :* 10.1007/978-3-642-04368-0_6], http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/BHR-formats09.pdf.

[44] P. BOUYER, S. HADDAD, P.-A. REYNIER.*Timed Unfoldings for Networks of Timed Automata*, in "Proceedings of the 4th International Symposium on Automated Technology for Verification and Analysis (ATVA'06)", Beijing, ROC, S. GRAF, W. ZHANG (editors), Lecture Notes in Computer Science, Springer, October 2006, vol. 4218, p. 292-306, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/BHR-atva06.pdf.

[45] TH. CHATAIN, P. GASTIN, N. SZNAJDER.*Natural Specifications Yield Decidability for Distributed Synthesis of Asynchronous Systems*, in "Proceedings of the 35th International Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM'09)", Špindlerův Mlýn, Czech Republic, M. NIELSEN, A. KUČERA, P. BRO MILTERSEN, C. PALAMIDESSI, P. TŮMA, F. VALENCIA (editors), Lecture Notes in Computer Science, Springer, January 2009, vol. 5404, p. 141-152 [*DOI :* 10.1007/978-3-540-95891-8_16], http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/CGS-sofsem09.pdf.

[46] G. CHIOLA, C. DUTHEILLET, G. FRANCESCHINIS, S. HADDAD.*Stochastic Well-Formed Colored Nets and Symmetric Modeling Applications*, in "IEEE Transactions on Computers", November 1993, vol. 42, nᵒ 11, p. 1343-1360, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PS/CDFH-toc93.ps.

[47] A. CHURCH.*Logic, arithmetics, and automata*, in "Proc. of Int. Congr. of Mathematicians", 1962, p. 23–35.

[48] R. DEBOUK, D. TENEKETZIS.*Coordinated decentralized protocols for failure diagnosis of discrete-event systems*, in "Journal of Discrete Event Dynamical Systems: Theory and Application", 2000, vol. 10, p. 33–86.

[49] D. EL HOG-BENZINA, S. HADDAD, R. HENNICKER.*Process Refinement and Asynchronous Composition with Modalities*, in "Proceedings of the 2nd International Workshop on Abstractions for Petri Nets and Other Models of Concurrency (APNOC'10)", Braga, Portugal, N. SIDOROVA, A. SEREBRENIK (editors), June 2010, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/EHH-apnoc10.pdf.

[50] J. ESPARZA, K. HELJANKO.*Unfoldings - A Partial-Order Approach to Model Checking*, EATCS Monographs in Theoretical Computer Science, Springer, 2008.

[51] É. FABRE, A. BENVENISTE, C. JARD, S. HAAR.*Diagnosis of Asynchronous Discrete Event Systems, a Net Unfolding Approach*, in "IEEE Trans. Aut. Control", 2003, vol. 48 (5), p. 714-727.

[52] É. FABRE, A. BENVENISTE, C. JARD, S. HAAR.*Distributed monitoring of concurrent and asynchronous systems*, in "Discrete Event Dynamic Systems: theory and application", 2005, vol. 15 (1), p. 33-84, Preliminary version: Proc. CONCUR 2003, LNCS 2761, pp.1–28, Springer.

[53] B. FINKBEINER, S. SCHEWE.*Uniform distributed synthesis*, in "Proc. of the 20th IEEE Annual Symposium on Logic in Computer Science (LICS'05)", IEEE Computer Society Press, 2005, p. 321–330.

[54] P. GASTIN, B. LERMAN, M. ZEITOUN.*Distributed games with causal memory are decidable for series-parallel systems*, in "Proceedings of the 24th Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS'04)", Chennai, India, K. LODAYA, M. MAHAJAN (editors), Lecture Notes in Computer Science, Springer, December 2004, vol. 3328, p. 275-286, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/GLZ-fsttcs04.pdf.

[55] P. GASTIN, N. SZNAJDER, M. ZEITOUN.*Distributed synthesis for well-connected architectures*, in "Proceedings of the 26th Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS'06)", Kolkata, India, N. GARG, S. ARUN-KUMAR (editors), Lecture Notes in Computer Science, Springer, December 2006, vol. 4337, p. 321-332 [*DOI :* 10.1007/11944836_30], http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/GSZ-fsttcs2006.pdf.

[56] S. HAAR, A. BENVENISTE, É. FABRE, C. JARD.*Partial Order Diagnosability Of Discrete Event Systems Using Petri Net Unfoldings*, in "42nd IEEE Conference on Decision and Control (CDC)", 2003.

[57] S. HAAR.*Probabilistic Cluster Unfoldings*, in "Fundamenta Informaticae", 2003, vol. 53 (3-4), p. 281-314.

[58] S. HAAR, C. JARD, G.-V. JOURDAN.*Testing Input/Output Partial Order Automata*, in "Proc. TESTCOM/FATES", LNCS, Springer, 2007, vol. 4581, p. 171-185, LNCS 4581.

[59] O. KUPFERMAN, M. Y. VARDI.*Synthesizing Distributed Systems*, in "Proc. of the 16th IEEE Annual Symposium on Logic in Computer Science (LICS'01)", IEEE Computer Society Press, 2001.

[60] S. LAFORTUNE, Y. WANG, T.-S. YOO.*Diagnostic Decentralisé Des Systèmes A Evénements Discrets*, in "Journal Europeen des Systèmes Automatisés (RS-JESA)", August 2005, vol. 99, n$^\text{o}$ 99, p. 95–110.

[61] K. G. LARSEN, P. PETTERSSON, W. YI.*Compositional and symbolic model-checking of real-time systems*, in "Proc. of RTSS 1995", IEEE Computer Society, 1995, p. 76-89.

[62] S. MOHALIK, I. WALUKIEWICZ.*Distributed Games*, in "Proc. of the 23rd Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS'03)", LNCS, Springer, 2003, vol. 2914, p. 338–351.

[63] A. PNUELI, R. ROSNER.*Distributed reactive systems are hard to synthesize*, in "Proc. of the 31st Annual IEEE Symposium on Foundations of Computer Science (FOCS'90)", IEEE Computer Society Press, 1990, vol. II, p. 746–757.

[64] L. Ricker, K. Rudie.*Know Means No: Incorporating Knowledge into Discrete-Event Control Systems*, in "IEEE Transactions on Automatic Control", September 2000, vol. 45, n° 9, p. 1656–1668.

[65] L. Ricker, K. Rudie.*Knowledge Is a Terrible Thing to Waste: Using Inference in Discrete-Event Control Problems*, in "IEEE Transactions on Automatic Control", MarchSeptember 2007, vol. 52, n° 3, p. 428–441.

[66] C. Rodríguez, S. Schwoon, V. Khomenko.*Contextual Merged Processes*, in "34th International Conference on Applications and Theory of Petri Nets (ICATPN'13)", Italy, Lecture Notes in Computer Science, Springer, 2013, vol. 7927, p. 29-48 [*DOI :* 10.1007/978-3-642-38697-8_3], https://hal.archives-ouvertes.fr/hal-00926202.

[67] H. L. S. Younes, R. G. Simmons.*Statistical probabilistic model checking with a focus on time-bounded properties*, in "Inf. Comput.", September 2006, vol. 204, p. 1368–1409 [*DOI :* 10.1016/J.IC.2006.05.002], http://dl.acm.org/citation.cfm?id=1182767.1182770.

Project-Team PARIETAL

Modelling brain structure, function and variability based on high-field MRI data.

IN COLLABORATION WITH: CEA Neurospin

IN PARTNERSHIP WITH:
Centre CEA-Saclay

RESEARCH CENTER
Saclay - Île-de-France

THEME
Computational Neuroscience and Medecine

Table of contents

Project-Team PARIETAL

Creation of the Project-Team: 2009 July 01

Keywords:

Computer Science and Digital Science:

3.3. - Data and knowledge analysis
3.3.2. - Data mining
3.3.3. - Big data analysis
3.4. - Machine learning and statistics
3.4.1. - Supervised learning
3.4.2. - Unsupervised learning
3.4.4. - Optimization and learning
3.4.5. - Bayesian methods
3.4.6. - Neural networks
3.4.7. - Kernel methods
3.4.8. - Deep learning
5.3.2. - Sparse modeling and image representation
5.3.3. - Pattern recognition
5.9.1. - Sampling, acquisition
5.9.2. - Estimation, modeling
5.9.6. - Optimization tools
6.2.4. - Statistical methods
6.2.6. - Optimization
8.2. - Machine learning
8.3. - Signal analysis

Other Research Topics and Application Domains:

1.3. - Neuroscience and cognitive science
1.3.1. - Understanding and simulation of the brain and the nervous system
1.3.2. - Cognitive science
2.2.6. - Neurodegenerative diseases
2.6.1. - Brain imaging

The Parietal team is localized in two places: in the Alan Turing building of Inria Saclay, and the Neurospin building of CEA Saclay.

1. Members

Research Scientists
Bertrand Thirion [Team leader, Inria, Research Scientist,Senior Researcher, HDR]
Philippe Ciuciu [CEA, Research Scientist, Senior Researcher, HDR]
Shan Liu [Inria, Research Scientist, Researcher]
Gaël Varoquaux [Inria, Research Scientist, Researcher]

Faculty Member
Matthieu Kowalski [Associate Professor,Univ. Paris XI, Faculty Member, until Aug 2016]

Technical Staff

    Kamalaker Reddy Dadi [CEA]

    Loïc Estève [Inria]

    Ana Luisa Grilo Pinho [Inria]

PhD Students

    Patricio Cerda Reyes [Inria, from Oct 2016]

    Jérome Dockès [Inria, from Apr 2016]

    Elvis Dohmatob [Inria, granted by Fondation Cooper. Scient. Campus Paris Saclay-DIGITEO]

    Loubna El Gueddari [Min. Ens. Sup. Recherche, from Oct 2016]

    Andres Hoyos Idrobo [Inria]

    Carole Lazarus [CEA]

    Arthur Mensch [Univ. Paris XI]

Post-Doctoral Fellows

    Darya Chyzhyk [Inria, from Aug 2016]

    Joke Durnez [Inria, from Nov 2016]

    Daria La Rocca [CEA]

    Andre Monteiro Manoel [CEA, from Dec 2016]

    Mehdi Rahim [Inria]

Administrative Assistants

    Régine Bricquet [Inria, until Jun 2016]

    Tiffany Caristan [Inria, from Jun 2016]

Others

    Alexandre Abraham [Inria, until Sep 2016]

    Moritz Boos [Inria, intern, from Jul 2016 until Oct 2016]

    Olivier Grisel [Inria, Engineers]

    Guillaume Lemaitre [Inria, Post-Doctoral Fellow, from Dec 2016]

    Joao Loula Guimaraes de Campos [Inria, intern, from Jun 2016]

2. Overall Objectives

2.1. Overall Objectives

The Parietal team focuses on mathematical methods for modeling and statistical inference based on neuroimaging data, with a particular interest in machine learning techniques and applications of human functional imaging. This general theme splits into four research axes:

- Modeling for neuroimaging population studies,

- Encoding and decoding models for cognitive imaging,

- Statistical and machine learning methods for large-scale data,

- Compressed-sensing for MRI.

Parietal is also strongly involved in open-source software development in scientific Python (machine learning) and for neuroimaging applications.

3. Research Program

3.1. Inverse problems in Neuroimaging

Many problems in neuroimaging can be framed as forward and inverse problems. For instance, brain population imaging is concerned with the *inverse problem* that consists in predicting individual information (behavior, phenotype) from neuroimaging data, while the corresponding *forward problem* boils down to explaining neuroimaging data with the behavioral variables. Solving these problems entails the definition of two terms: a loss that quantifies the goodness of fit of the solution (does the model explain the data well enough ?), and a regularization scheme that represents a prior on the expected solution of the problem. These priors can be used to enforce some properties on the solutions, such as sparsity, smoothness or being piece-wise constant.

Let us detail the model used in typical inverse problem: Let $\mathbf{X}$ be a neuroimaging dataset as an $(n_{subjects}, n_{voxels})$ matrix, where $n_{subjects}$ and n_{voxels} are the number of subjects under study, and the image size respectively, $\mathbf{Y}$ a set of values that represent characteristics of interest in the observed population, written as $(n_{subjects}, n_{features})$ matrix, where $n_{features}$ is the number of characteristics that are tested, and β an array of shape $(n_{voxels}, n_{features})$ that represents a set of pattern-specific maps. In the first place, we may consider the columns $\mathbf{Y}_1, .., \mathbf{Y}_{n_{features}}$ of Y independently, yielding $n_{features}$ problems to be solved in parallel:

$$\mathbf{Y}_i = \mathbf{X}\beta_i + \epsilon_i, \forall i \in \{1, .., n_{features}\},$$

where the vector contains β_i is the ith row of β. As the problem is clearly ill-posed, it is naturally handled in a regularized regression framework:

$$\widehat{\beta_i} = \operatorname{argmin}_{\beta_i} \|\mathbf{Y}_i - \mathbf{X}\beta_i\|^2 + \Psi(\beta_i), \tag{10}$$

where Ψ is an adequate penalization used to regularize the solution:

$$\Psi(\beta; \lambda_1, \lambda_2, \eta_1, \eta_2) = \lambda_1 \|\beta\|_1 + \lambda_2 \|\beta\|_2 + \eta_1 \|\nabla\beta\|_{2,1} + \eta_2 \|\nabla\beta\|_{2,2} \tag{11}$$

with λ_1, λ_2, η_1, $\eta_2 \geq 0$ (this formulation particularly highlights the fact that convex regularizers are norms or quasi-norms). In general, only one or two of these constraints is considered (hence is enforced with a non-zero coefficient):

- When $\lambda_1 > 0$ only (LASSO), and to some extent, when $\lambda_1, \lambda_2 > 0$ only (elastic net), the optimal solution β is (possibly very) sparse, but may not exhibit a proper image structure; it does not fit well with the intuitive concept of a brain map.
- Total Variation regularization (see Fig. 1) is obtained for ($\eta_1 > 0$ only), and typically yields a piece-wise constant solution. It can be associated with Lasso to enforce both sparsity and sparse variations.
- Smooth lasso is obtained with ($\eta_2 > 0$ and $\lambda_1 > 0$ only), and yields smooth, compactly supported spatial basis functions.

Note that, while the qualitative aspect of the solutions are very different, the predictive power of these models is often very close.

The performance of the predictive model can simply be evaluated as the amount of variance in $\mathbf{Y}_i$ fitted by the model, for each $i \in \{1, .., n_{features}\}$. This can be computed through cross-validation, by *learning* $\widehat{\beta_i}$ on some part of the dataset, and then estimating $\|\mathbf{Y}_i - \mathbf{X}\widehat{\beta_i}\|^2$ using the remainder of the dataset.
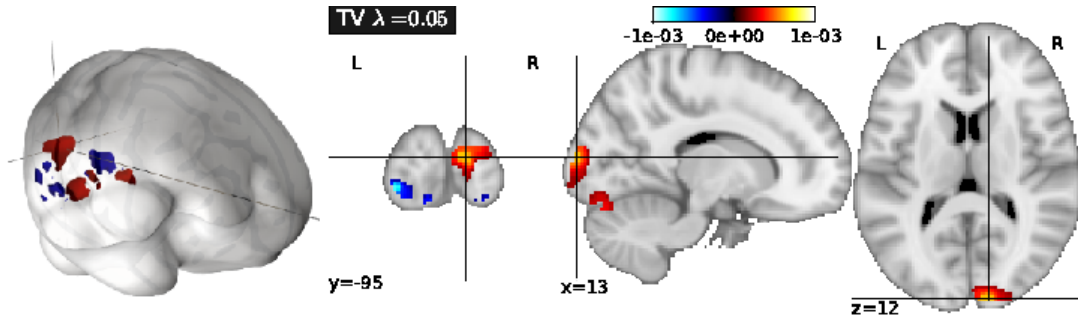
Figure 1. Example of the regularization of a brain map with total variation in an inverse problem. The problem here is to predict the spatial scale of an object presented as a stimulus, given functional neuroimaging data acquired during the presentation of an image. Learning and test are performed across individuals. Unlike other approaches, Total Variation regularization yields a sparse and well-localized solution that also enjoys high predictive accuracy.

This framework is easily extended by considering

- *Grouped penalization*, where the penalization explicitly includes a prior clustering of the features, i.e. voxel-related signals, into given groups. This amounts to enforcing structured priors on the problem solution.

- *Combined penalizations*, i.e. a mixture of simple and group-wise penalizations, that allow some variability to fit the data in different populations of subjects, while keeping some common constraints.

- *Logistic and hinge regression*, where a non-linearity is applied to the linear model so that it yields a probability of classification in a binary classification problem.

- *Robustness to between-subject variability* to avoid the learned model overly reflecting a few outlying particular observations of the training set. Note that noise and deviating assumptions can be present in both $\mathbf{Y}$ and $\mathbf{X}$

- *Multi-task learning*: if several target variables are thought to be related, it might be useful to constrain the estimated parameter vector β to have a shared support across all these variables.
  For instance, when one of the variables $\mathbf{Y}_i$ is not well fitted by the model, the estimation of other variables $\mathbf{Y}_j, j \neq i$ may provide constraints on the support of β_i and thus, improve the prediction of $\mathbf{Y}_i$.

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon, \tag{12}$$

then

$$\widehat{\beta} = \mathrm{argmin}_{\beta=(\beta_i), i=1..n_f} \sum_{i=1}^{n_f} \|\mathbf{Y_i} - \mathbf{X}\beta_\mathbf{i}\|^2 + \lambda \sum_{j=1}^{n_{voxels}} \sqrt{\sum_{i=1}^{n_f} \beta_\mathbf{i,j}^\mathbf{2}} \tag{13}$$

3.2. Multivariate decompositions

Multivariate decompositions provide a way to model complex data such as brain activation images: for instance, one might be interested in extracting an *atlas of brain regions* from a given dataset, such as regions exhibiting similar activity during a protocol, across multiple protocols, or even in the absence of protocol (during resting-state). These data can often be factorized into spatial-temporal components, and thus can be estimated through *regularized Principal Components Analysis* (PCA) algorithms, which share some common steps with regularized regression.

Let $\mathbf{X}$ be a neuroimaging dataset written as an $(n_{subjects}, n_{voxels})$ matrix, after proper centering; the model reads

$$\mathbf{X} = \mathbf{A}\mathbf{D} + \epsilon, \tag{14}$$

where $\mathbf{D}$ represents a set of n_{comp} spatial maps, hence a matrix of shape (n_{comp}, n_{voxels}), and $\mathbf{A}$ the associated subject-wise loadings. While traditional PCA and independent components analysis are limited to reconstructing components $\mathbf{D}$ within the space spanned by the column of $\mathbf{X}$, it seems desirable to add some constraints on the rows of $\mathbf{D}$, that represent spatial maps, such as sparsity, and/or smoothness, as it makes the interpretation of these maps clearer in the context of neuroimaging. This yields the following estimation problem:

$$\min_{\mathbf{D},\mathbf{A}} \|\mathbf{X} - \mathbf{A}\mathbf{D}\|^2 + \Psi(\mathbf{D}) \text{ s.t. } \|\mathbf{A}_i\| = 1 \ \forall i \in \{1..n_{features}\}, \tag{15}$$

where $(\mathbf{A}_i)$, $i \in \{1..n_{features}\}$ represents the columns of $\mathbf{A}$. Ψ can be chosen such as in Eq. (2) in order to enforce smoothness and/or sparsity constraints.

The problem is not jointly convex in all the variables but each penalization given in Eq (2) yields a convex problem on $\mathbf{D}$ for $\mathbf{A}$ fixed, and conversely. This readily suggests an alternate optimization scheme, where $\mathbf{D}$ and $\mathbf{A}$ are estimated in turn, until convergence to a local optimum of the criterion. As in PCA, the extracted components can be ranked according to the amount of fitted variance. Importantly, also, estimated PCA models can be interpreted as a probabilistic model of the data, assuming a high-dimensional Gaussian distribution (probabilistic PCA).

Utlimately, the main limitations to these algorithms is the cost due to the memory requirements: holding datasets with large dimension and large number of samples (as in recent neuroimaging cohorts) leads to inefficient computation. To solve this issue, online method are particularly attractive.

3.3. Covariance estimation

Another important estimation problem stems from the general issue of learning the relationship between sets of variables, in particular their covariance. Covariance learning is essential to model the dependence of these variables when they are used in a multivariate model, for instance to study potential interactions between variables. Covariance learning is necessary to model latent interactions in high-dimensional observation spaces, e.g. when considering multiple contrasts or functional connectivity data.

The difficulties are two-fold: on the one hand, there is a shortage of data to learn a good covariance model from an individual subject, and on the other hand, subject-to-subject variability poses a serious challenge to the use of multi-subject data. While the covariance structure may vary from population to population, or depending on the input data (activation versus spontaneous activity), assuming some shared structure across problems, such as their sparsity pattern, is important in order to obtain correct estimates from noisy data. Some of the most important models are:

- **Sparse Gaussian graphical models**, as they express meaningful conditional independence relationships between regions, and do improve conditioning/avoid overfit.

- **Decomposable models**, as they enjoy good computational properties and enable intuitive interpretations of the network structure. Whether they can faithfully or not represent brain networks is still an open question.
- **PCA-based regularization of covariance** which is powerful when modes of variation are more important than conditional independence relationships.

Adequate model selection procedures are necessary to achieve the right level of sparsity or regularization in covariance estimation; the natural evaluation metric here is the out-of-samples likelihood of the associated Gaussian model. Another essential remaining issue is to develop an adequate statistical framework to test differences between covariance models in different populations. To do so, we consider different means of parametrizing covariance distributions and how these parametrizations impact the test of statistical differences across individuals.
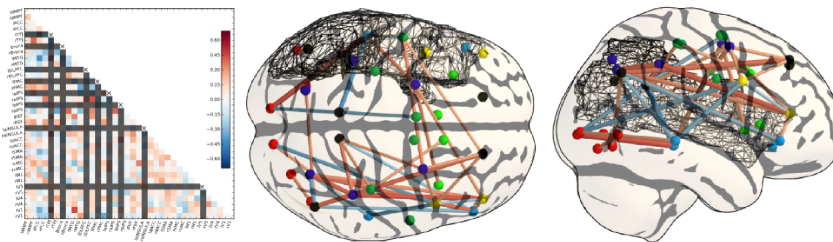
Figure 2. Example of functional connectivity analysis: The correlation matrix describing brain functional connectivity in a post-stroke patient (lesion volume outlined as a mesh) is compared to a group of control subjects. Some edges of the graphical model show a significant difference, but the statistical detection of the difference requires a sophisticated statistical framework for the comparison of graphical models.

4. Application Domains

4.1. Cognitive neuroscience

4.1.1. *Macroscopic Functional cartography with functional Magnetic Resonance Imaging (fMRI)*

The brain as a highly structured organ, with both functional specialization and a complex newtork organization. While most of the knowledge historically comes from lesion studies and animal electophysiological recordings, the development of non-invasive imaging modalities, such as fMRI, has made it possible to study routinely high-level cognition in humans since the early 90's. This has opened major questions on the interplay between mind and brain , such as: How is the function of cortical territories constrained by anatomy (connectivity) ? How to assess the specificity of brain regions ? How can one characterize reliably inter-subject differences ?

4.1.2. *Analysis of brain Connectivity*

Functional connectivity is defined as the interaction structure that is underlies brain function. Since the beginning of fMRI, it has been observed that remote regions sustain high correlation in their spontaneous activity, i.e. in the absence of a driving task. This means that the signals observed during resting-state define a signature of the connectivity of brain regions. The main interest of retsing-state fMRI is that it provides easy-to-acquire functional markers that have recently been proved to be very powerful for population studies.

4.1.3. Modeling of brain processes (MEG)

While fMRI has been very useful in defining the function of regions at the mm scale, Magneto-encephalography (MEG) provides the other piece of the puzzle, namely temporal dynamics of brain activity, at the ms scale. MEG is also non-invasive. It makes it possible to keep track of precise schedule of mental operations and their interactions. It also opens the way toward a study of the rythmic activity of the brain. On the other hand, the localization of brain activity with MEG entails the solution of a hard inverse problem.

5. New Software and Platforms

5.1. Mayavi

FUNCTIONAL DESCRIPTION

Mayavi is the most used scientific 3D visualization Python software. Mayavi can be used as a visualization tool, through interactive command line or as a library. It is distributed under Linux through Ubuntu, Debian, Fedora and Mandriva, as well as in PythonXY and EPD Python scientific distributions. Mayavi is used by several software platforms, such as PDE solvers (fipy, sfepy), molecule visualization tools and brain connectivity analysis tools (connectomeViewer).

- Contact: Gaël Varoquaux
- URL: http://mayavi.sourceforge.net/

5.2. Nilearn

NeuroImaging with scikit learn
KEYWORDS: Health - Neuroimaging - Medical imaging
FUNCTIONAL DESCRIPTION

NiLearn is the neuroimaging library that adapts the concepts and tools of scikit-learn to neuroimaging problems. As a pure Python library, it depends on scikit-learn and nibabel, the main Python library for neuroimaging I/O. It is an open-source project, available under BSD license. The two key components of NiLearn are i) the analysis of functional connectivity (spatial decompositions and covariance learning) and ii) the most common tools for multivariate pattern analysis. A great deal of efforts has been put on the efficiency of the procedures both in terms of memory cost and computation time.

- Participants: Gaël Varoquaux, Bertrand Thirion, Loïc Estève, Alexandre Abraham, Michael Eickenberg, Alexandre Gramfort, Fabian Pedregosa Izquierdo, Elvis Dohmatob and Virgile Fritsch
- Contact: Bertrand Thirion
- URL: http://nilearn.github.io/

5.3. PyHRF

KEYWORDS: FMRI - Statistic analysis - Neurosciences - IRM - Brain - Health - Medical imaging
FUNCTIONAL DESCRIPTION

As part of fMRI data analysis, PyHRF provides a set of tools for addressing the two main issues involved in intra-subject fMRI data analysis : (i) the localization of cerebral regions that elicit evoked activity and (ii) the estimation of the activation dynamics also referenced to as the recovery of the Hemodynamic Response Function (HRF). To tackle these two problems, PyHRF implements the Joint Detection-Estimation framework (JDE) which recovers parcel-level HRFs and embeds an adaptive spatio-temporal regularization scheme of activation maps.

- Participants: Thomas Vincent, Solveig Badillo, Lotfi Chaari, Christine Bakhous, Florence Forbes, Philippe Ciuciu, Laurent Risser, Thomas Perret and Aina Frau Pascual
- Partners: CEA - NeuroSpin
- Contact: Florence Forbes
- URL: http://pyhrf.org

5.4. Scikit-learn

KEYWORDS: Classification - Learning - Clustering - Regession - Medical imaging
SCIENTIFIC DESCRIPTION

Scikit-learn is a Python module integrating classic machine learning algorithms in the tightly-knit scientific Python world. It aims to provide simple and efficient solutions to learning problems, accessible to everybody and reusable in various contexts: machine-learning as a versatile tool for science and engineering.
FUNCTIONAL DESCRIPTION

Scikit-learn can be used as a middleware for prediction tasks. For example, many web startups adapt Scikitlearn to predict buying behavior of users, provide product recommendations, detect trends or abusive behavior (fraud, spam). Scikit-learn is used to extract the structure of complex data (text, images) and classify such data with techniques relevant to the state of the art.

Easy to use, efficient and accessible to non datascience experts, Scikit-learn is an increasingly popular machine learning library in Python. In a data exploration step, the user can enter a few lines on an interactive (but non-graphical) interface and immediately sees the results of his request. Scikitlearn is a prediction engine . Scikit-learn is developed in open source, and available under the BSD license.

- Participants: Olivier Grisel, Gaël Varoquaux, Bertrand Thirion, Michael Eickenberg, Loïc Estève, Alexandre Gramfort, Arthur Mensch
- Partners: CEA - Logilab - Nuxeo - Saint Gobain - Telecom Paris - Tinyclues
- Contact: Olivier Grisel
- URL: http://scikit-learn.org

6. New Results

6.1. Dictionary Learning for Massive Matrix Factorization

Sparse matrix factorization is a popular tool to obtain interpretable data decompositions, which are also effective to perform data completion or denoising. Its applicability to large datasets has been addressed with online and randomized methods, that reduce the complexity in one of the matrix dimension, but not in both of them. In this paper, we tackle very large matrices in both dimensions. We propose a new factoriza-tion method that scales gracefully to terabyte-scale datasets, that could not be processed by previous algorithms in a reasonable amount of time. We demonstrate the efficiency of our approach on massive functional Magnetic Resonance Imaging (fMRI) data, and on matrix completion problems for recommender systems, where we obtain significant speed-ups compared to state-of-the art coordinate descent methods.

See Fig. 3 for an illustration and [22] for more information.

6.2. Learning brain regions via large-scale online structured sparse dictionary-learning

We propose a multivariate online dictionary-learning method for obtaining de-compositions of brain images with structured and sparse components (aka atoms). Sparsity is to be understood in the usual sense: the dictionary atoms are constrained to contain mostly zeros. This is imposed via an 1-norm constraint. By "structured", we mean that the atoms are piece-wise smooth and compact, thus making up blobs, as opposed to scattered patterns of activation. We propose to use a Sobolev (Laplacian) penalty to impose this type of structure. Combining the two penalties, we obtain decompositions that properly delineate brain structures from functional images. This non-trivially extends the online dictionary-learning work of Mairal et al. (2010), at the price of only a factor of 2 or 3 on the overall running time. Just like the Mairal et al. (2010) reference method, the online nature of our proposed algorithm allows it to scale to arbitrarily sized datasets. Experiments on brain data show that our proposed method extracts structured and denoised dictionaries that are more intepretable and better capture inter-subject variability in small medium, and large-scale regimes alike, compared to state-of-the-art models.
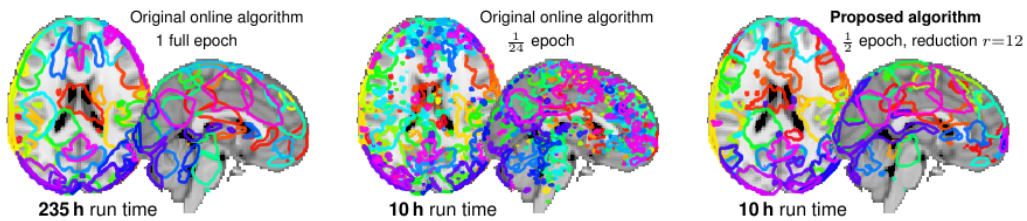
Figure 3. Brain atlases: outlines of each map obtained with dictionary learning. Left: the reference algorithm on the full dataset. Middle: the reference algorithm on a twentieth of the dataset. Right: the proposed algorithm with a similar run time: half the dataset and a compression factor of 9. Compared to a full run of the baseline algorithm, the figure explore two possible strategies to decrease computation time: processing less data (middle), or our approach (right). Our approach achieves a result closer to the gold standard in a given time budget. See [22] for more information.
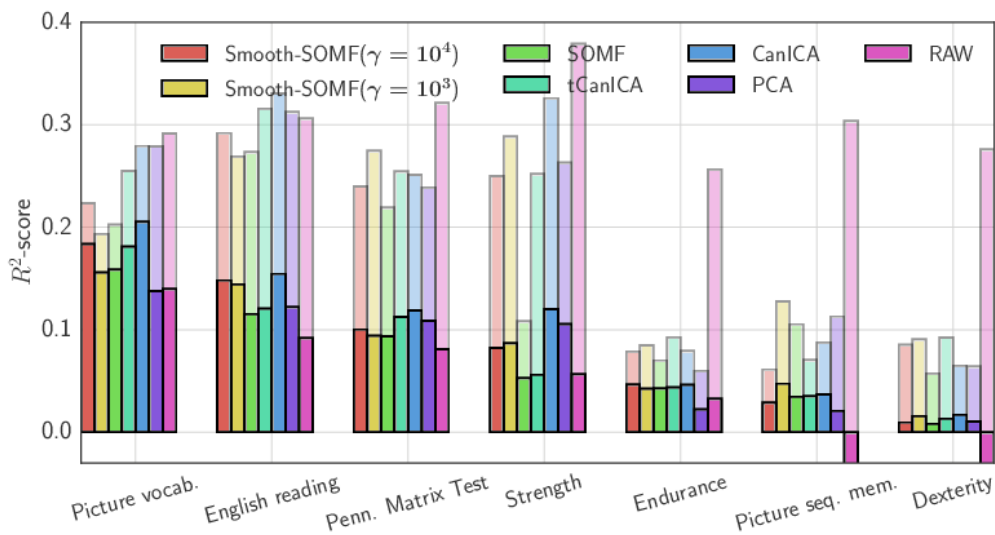
Figure 4. Predicting behavioral variables of the Human Connectome Project dataset using subject-level brain activity maps and various intermediate representations obtained with variants of dictionary learning. Bold bars represent performance on test set while faint bars in the background represent performance on train set. See [19] for more information.

See Fig. 4 for an illustration and [19] for more information.

6.3. Social-sparsity brain decoders: faster spatial sparsity

Spatially-sparse predictors are good models for brain decoding: they give accurate predictions and their weight maps are interpretable as they focus on a small number of regions. However, the state of the art, based on total variation or graph-net, is computationally costly. Here we introduce sparsity in the local neighborhood of each voxel with social-sparsity, a structured shrinkage operator. We find that, on brain imaging classification problems, social-sparsity performs almost as well as total-variation models and better than graph-net, for a fraction of the computational cost. It also very clearly outlines predictive regions. We give details of the model and the algorithm.
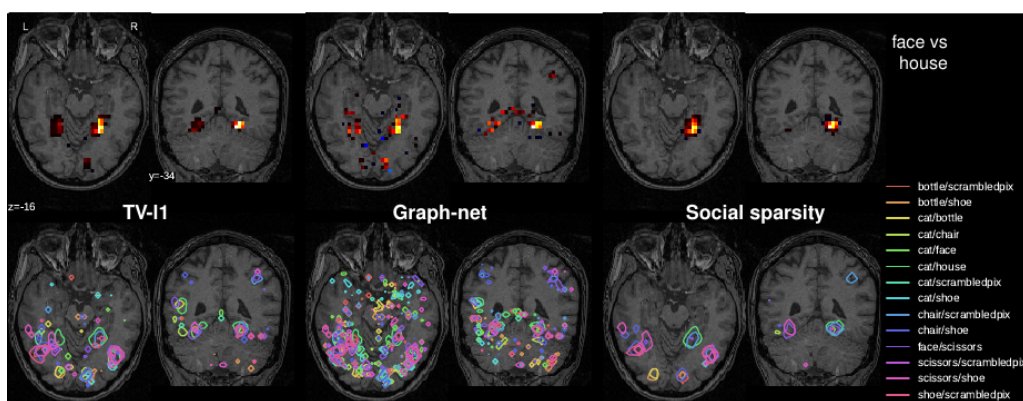
Figure 5. Decoder maps for the object-classification task – Top: weight maps for the face-versus-house task. Overall, the maps segment the right and left parahippocampal place area (PPA), a well-known place-specific regions, although the left PPA is weak in TV-l1, spotty in graph-net, and absent in social sparsity. Bottom: outlines at 0.01 of the other tasks. Beyond the PPA, several known functional regions stand out such as primary or secondary visual areas around the prestriate cortex as well as regions in the lateral occipital cortex, responding to structured objects. Note that the graphnet outlines display scattered small regions even thought the value of the contours is chosen at 0.01, well above numerical noise. See [32] for more information.

See Fig. 5 for an illustration and [32] for more information.

6.4. Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example

Resting-state functional Magnetic Resonance Imaging (R-fMRI) holds the promise to reveal functional biomarkers of neuropsychiatric disorders. However, extracting such biomarkers is challenging for complex multi-faceted neuropatholo-gies, such as autism spectrum disorders. Large multi-site datasets increase sample sizes to compensate for this complexity, at the cost of uncontrolled heterogeneity. This heterogeneity raises new challenges, akin to those face in realistic diagnostic applications. Here, we demonstrate the feasibility of inter-site classification of neuropsychiatric status, with an application to the Autism Brain Imaging Data Exchange (ABIDE) database, a large (N=871) multi-site autism dataset. For this purpose, we investigate pipelines that extract the most predictive biomarkers from the data. These R-fMRI pipelines build participant-specific connectomes from functionally-defined brain areas. Connectomes are then compared across participants to learn patterns of connectivity that differentiate typical controls from individuals with autism. We predict this neuropsychiatric status for participants from the same acquisition sites or different, unseen, ones. Good choices

of methods for the various steps of the pipeline lead to 67% prediction accuracy on the full ABIDE data, which is significantly better than previously reported results. We perform extensive validation on multiple subsets of the data defined by different inclusion criteria. These enables detailed analysis of the factors contributing to successful connectome-based prediction. First, prediction accuracy improves as we include more subjects, up to the maximum amount of subjects available. Second, the definition of functional brain areas is of paramount importance for biomarker discovery: brain areas extracted from large R-fMRI datasets outperform reference atlases in the classification tasks.
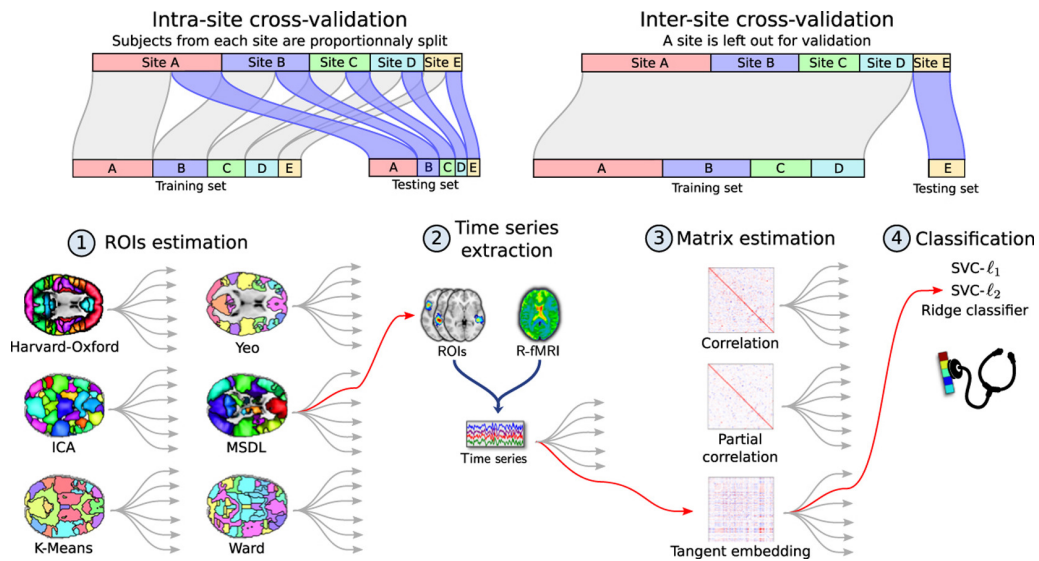
Figure 6. Validation of an fMRI-based pipeline for autism prediction. Several variants are considered for each pipeline step. See [1] for more information.

See Fig. 6 for an illustration and [1] for more information.

6.5. Seeing it all: Convolutional network layers map the function of the human visual system

Convolutional networks used for computer vision represent candidate models for the computations performed in mammalian visual systems. We use them as a detailed model of human brain activity during the viewing of natural images by constructing predictive models based on their different layers and BOLD fMRI activations. Analyzing the predictive performance across layers yields characteristic fingerprints for each visual brain region: early visual areas are better described by lower level convolutional net layers and later visual areas by higher level net layers, exhibiting a progression across ventral and dorsal streams. Our predictive model generalizes beyond brain responses to natural images. We illustrate this on two experiments, namely retinotopy and face-place oppositions, by synthesizing brain activity and performing classical brain mapping upon it. The synthesis recovers the activations observed in the corresponding fMRI studies, showing that this deep encoding model captures representations of brain function that are universal across experimental paradigms.

See Fig. 7 for an illustration and [10] for more information.

6.6. Formal Models of the Network Co-occurrence Underlying Mental Operations
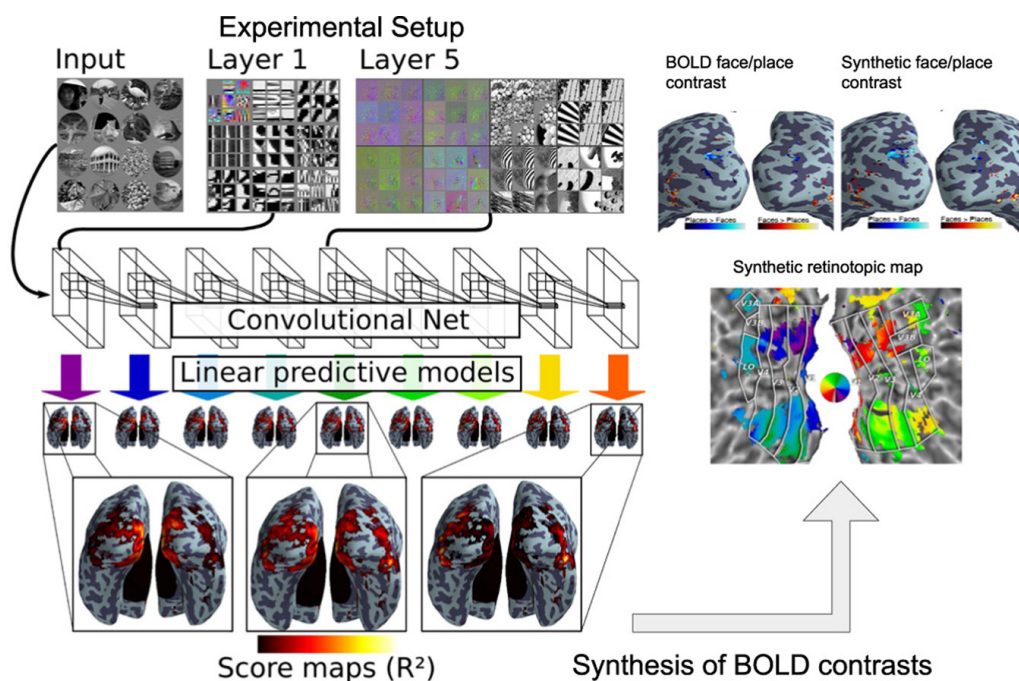
Figure 7. Overview of the vision mapping experiment: Convolutional network image representations of different layer depth explain brain activity throughout the full ventral visual stream. This mapping follows the known hierarchical organisation. Results from both static images and video stimuli. A model of brain activity for the full brain, based on the convolutional network, can synthesize brain maps for other visual experiments. Only deep models can reproduce observed BOLD activity. See [10] for more information.

Systems neuroscience has identified a set of canonical large-scale networks in humans. These have predominantly been characterized by resting-state analyses of the task-uncon-strained, mind-wandering brain. Their explicit relationship to defined task performance is largely unknown and remains challenging. The present work contributes a multivariate statistical learning approach that can extract the major brain networks and quantify their configuration during various psychological tasks. The method is validated in two extensive datasets (n = 500 and n = 81) by model-based generation of synthetic activity maps from recombination of shared network topographies. To study a use case, we formally revisited the poorly understood difference between neural activity underlying idling versus goal-directed behavior. We demonstrate that task-specific neural activity patterns can be explained by plausible combinations of resting-state networks. The possibility of decomposing a mental task into the relative contributions of major brain networks, the "network co-occurrence architecture" of a given task, opens an alternative access to the neural substrates of human cognition.

See Fig. 8 for an illustration and [6] for more information.

6.7. Transmodal Learning of Functional Networks for Alzheimer's Disease Prediction

Functional connectivity describes neural activity from resting-state functional magnetic resonance imaging (rs-fMRI). This noninvasive modality is a promising imaging biomarker of neurodegenerative diseases, such as Alzheimer's disease (AD), where the connectome can be an indicator to assess and to understand the pathology. However, it only provides noisy measurements of brain activity. As a consequence, it has shown fairly limited discrimination power on clinical groups. So far, the reference functional marker of AD is the fluorodeoxyglucose positron emission tomography (FDG-PET). It gives a reliable quantification of metabolic activity, but it is costly and invasive. Here, our goal is to analyze AD populations solely based on rs-fMRI, as functional connectivity is correlated to metabolism. We introduce transmodal learning: leveraging a prior from one modality to improve results of another modality on different subjects. A metabolic prior is learned from an independent FDG-PET dataset to improve functional connectivity-based prediction of AD. The prior acts as a regularization of connectivity learning and improves the estimation of discriminative patterns from distinct rs-fMRI datasets. Our approach is a two-stage classification strategy that combines several seed-based connectivity maps to cover a large number of functional networks that identify AD physiopathology. Experimental results show that our transmodal approach increases classification accuracy compared to pure rs-fMRI approaches, without resorting to additional invasive acquisitions. The method successfully recovers brain regions known to be impacted by the disease.

6.8. Assessing and tuning brain decoders: cross-validation, caveats, and guidelines

Decoding, ie prediction from brain images or signals, calls for empirical evaluation of its predictive power. Such evaluation is achieved via cross-validation, a method also used to tune decoders' hyper-parameters. This paper is a review on cross-validation procedures for decoding in neuroimaging. It includes a didactic overview of the relevant theoretical considerations. Practical aspects are highlighted with an extensive empirical study of the common decoders in within-and across-subject predictions, on multiple datasets –anatomical and functional MRI and MEG– and simulations. Theory and experiments outline that the popular " leave-one-out " strategy leads to unstable and biased estimates, and a repeated random splits method should be preferred. Experiments outline the large error bars of cross-validation in neuroimaging settings: typical confidence intervals of 10%. Nested cross-validation can tune decoders' parameters while avoiding circularity bias. However we find that it can be more favorable to use sane defaults, in particular for non-sparse decoders.

See Fig. 9 for an illustration and [16] for more information.

6.9. A projection algorithm for gradient waveforms design in Magnetic Resonance Imaging
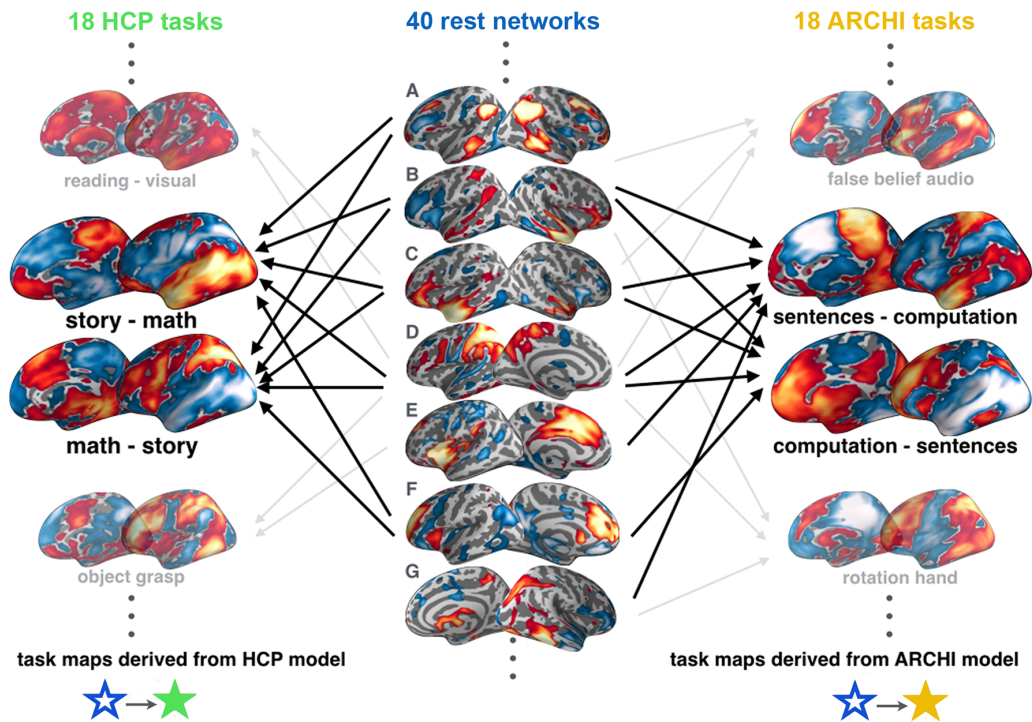
Figure 8. Task-rest correspondence: Reconstructing two similar tasks from two different datasets based on the same resting networks. 40 sparse PCA networks were discovered from the same rest data and used for feature engineering as a basis for classificationof 18 psychological tasks from HCP (left) and from ARCHI (right). Middle column: Examples of resting-state networks derived from decomposing rest data using sparse PCA. Networks B and C might be related to semantics processing in the anterior temporal lobe, network D covers extended parts of the parietal cortex, while networks E and F appear to be variants of the so-called "salience" network. Left/Right column: Examples of task-specific neural activity generated from network co-occurrence models of the HCP/ARCHI task batteries. Arrows: A diagnostic subanalysis indicated what rest networks were automatically ranked top-five in distinguishing a given task from the respective 17 other tasks. Although the experimental tasks in the HCP and ARCHI repositories, "story versus math" and "sentences versus computation" were the most similar cognitive contrasts in both datasets. For these four experimental conditions the model-derived task maps are highly similar. Consequently, two independent classification problems in two independent datasets with a six-fold difference in sample size resulted in two independent explicit models that, nevertheless, generated comparable task-specific maps. This indicated that network co-occurrence modeling indeed captures genuine aspects of neurobiology rather than arbitrary discriminatory aspects of the data. See [6] for more information.
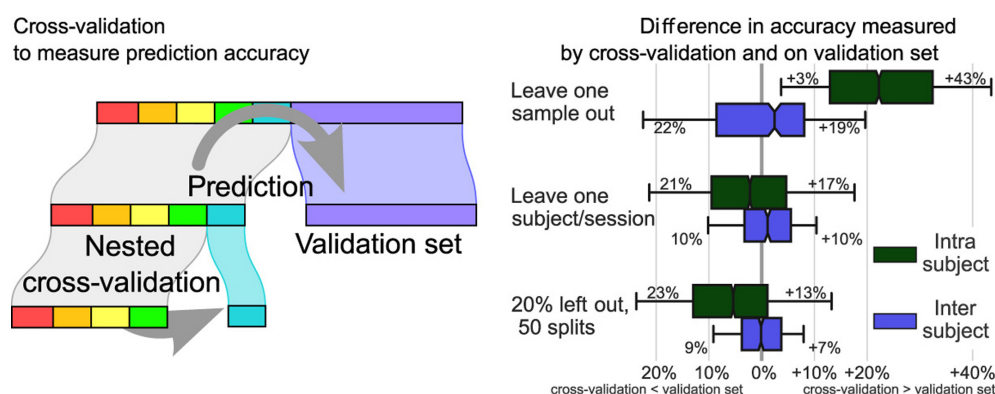
Figure 9. (Left) Illustration of the nested cross-validation principle. (Right) Typical cross-validated accuracy result: leave-one-out cross validation, when applied to imaging data, yields to optimistic bias (top) when used on dependent data, and in other cases leads to estimated with inflated variance. See [16] for more information.

Collecting the maximal amount of information in a given scanning time is a major concern in Magnetic Resonance Imaging (MRI) to speed up image acquisition. The hardware constraints (gradient magnitude, slew rate, ...), physical distortions (e.g., off-resonance effects) and sampling theorems (Shannon, compressed sensing) must be taken into account simultaneously, which makes this problem extremely challenging. To date, the main approach to design gradient waveform has consisted of selecting an initial shape (e.g. spiral, radial lines, ...) and then traversing it as fast as possible using optimal control. In this paper, we propose an alternative solution which first consists of defining a desired parameterization of the trajectory and then of optimizing for minimal deviation of the sampling points within gradient constraints. This method has various advantages. First, it better preserves the density of the input curve which is critical in sampling theory. Second, it allows to smooth high curvature areas making the acquisition time shorter in some cases. Third, it can be used both in the Shannon and CS sampling theories. Last, the optimized trajectory is computed as the solution of an efficient iterative algorithm based on convex programming. For piecewise linear trajectories, as compared to optimal control reparameterization, our approach generates a gain in scanning time of 10% in echo planar imaging while improving image quality in terms of signal-to-noise ratio (SNR) by more than 6 dB. We also investigate original trajectories relying on traveling salesman problem solutions. In this context, the sampling patterns obtained using the proposed projection algorithm are shown to provide significantly better reconstructions (more than 6 dB) while lasting the same scanning time.

See Fig. 10 for an illustration and [9] for more information.

6.10. Impact of perceptual learning on resting-state fMRI connectivity: A supervised classification study

Perceptual learning sculpts ongoing brain activity. This finding has been observed by statistically comparing the functional connectivity (FC) patterns computed from resting-state functional MRI (rs-fMRI) data recorded before and after intensive training to a visual attention task. Hence, functional connectivity serves a dynamic role in brain function, supporting the consolidation of previous experience. Following this line of research, we trained three groups of individuals to a visual discrimination task during a magneto-encephalography (MEG) experiment. The same individuals were then scanned in rs-fMRI. Here, in a supervised classification framework, we demonstrate that FC metrics computed on rs-fMRI data are able to predict the type of training the participants received. On top of that, we show that the prediction accuracies based on tangent embedding
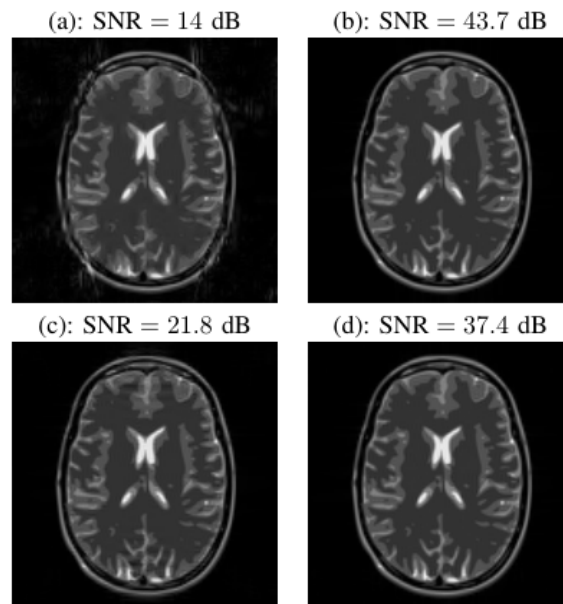
Figure 10. Reconstructed images from data collected along EPI-like trajectories. (a)-(b): Reconstruction results from the optimally reparameterized EPI readout. (c)-(d): Reconstructed results from data collected using the projected EPI trajectories. See [9] for more information.

FC measure outperform those based on our recently developed multivariate wavelet-based Hurst exponent estimator, which captures low frequency fluctuations in ongoing brain activity too.
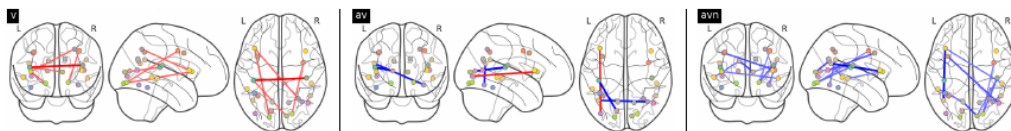
Figure 11. Statistical significant functional interactions (positive and negative values are color coded in red and blue, respectively) within each group of individuals (V: purely visual traing, AV: audio-visual training and AVn: unmatched audio-visual), Bonferroni-corrected for multiple comparisons at $\alpha = 0.05$. See [24] for more information.

See Fig. 11 for an illustration and [24] for more information.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Grants with Industry

7.1.1. The Wendelin FUI project

The Wendelin project has been granted on December 3rd, 2014. It has been selected at the *Programme d'Investissements d'Avenir (PIA)* that supports "cloud computing et Big Data". It gives visibility and fosters the French technological big data sector, and in particular the scikit-learn library, the NoSQL "NEO" et the decentralized "SlapOS" cloud, three open-source software supported by the Systematic *pôle de compétitivité*.

Scikit-learn is a worldwide reference library for machine learning. Gaël Varoquaux, Olivier Grisel and Alexandre Gramfort have been major players in the design of the library and Scikit-learn has then been supported by the growing scientific Python community. It is currently used by major internet companies as well as dynamic start-ups, including Google, Airbnb, Spotify, Evernote, AWeber, TinyClues; it wins more than half of the data science "Kaggle" competitions. Scikit-learn makes it possible to predict future outcomes given a training data, and thus to optimize company decisions. Almost 1 million euros will be invested to improve the algorithmic core of scikit-learn through the Wendelin project thanks to the Inria, ENS and Institut Mines Télécom teams. In particular, scikit-learn will be extended in order to ease online prediction and to include recent stochastic gradient algorithms.

NEO is the native NoSQL base of the Python language. It was initially designed by Nexedi and is currently used and embedded in the main software of company information systems. More than one million euros will be invested into NEO, so that scikit-learn can process within 10 years (out-of-core) data of 1 exabyte size.

Paris13 university and the Mines Télécom institute will extend the SlapOS distributed mesh cloud to deploy Wendelin in *Big Data as a Service* (BDaaS) mode, to achieve the interoperability between the Grid5000 and Teralab infrastructures and to extend the cloud toward smart sensor systems.

The combination of scikit-learn, NEO and SlapOS will improve the predictive maintenance of industrial plants with two major use cases: connected windmills (GDF SUEZ, Woelfel) and customer satisfaction in car sale systems (MMC Rus). In both cases it is about non-personal, yet profitable big data. The Wendelin project actually demonstrates that Big data can improve infrastructure and everyday-life equipment without intrusive data collection. For more information, please see http://www.wendelin.io.

The project partners are:
- Nexedi (leader)
- GDF SUEZ
- Abilian
- 2ndQuadrant
- Institut Mines Télécom
- Inria
- Université Paris 13

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. *CoSmic project*

Participants: Philippe Ciuciu [Correspondant], Carole Lazarus, Loubna El Gueddari.

This is a collaborative project with Jean-Luc Stark, (CEA) funded by the CEA program drf-impulsion.

Compressed Sensing is a recent theory in maths that allows the perfect recovery of signals or images from compressive acquisition scenarios. This approach has been popularized in MRI over the last decade as well as in astrophysics (noticeably in radio-astronomy). So far, both of these fields have developed skills in CS separately. The aim of the COSMIC project is to foster collaborations between CEA experts in MRI (Inria-CEA Parietal team within NeuroSpin) and in astrophysics (CosmoStat lab within the Astrophysics Department). These interactions will allow us to share different expertise in order to improve image quality, either in MRI or in radio-astronomy (thanks to the interferometry principle). In this field, given the data delivered by radio-telescopet he goal will consist of extracting high temporal resolution information in order to study fast transient events.

8.1.2. *BrainAMP project*

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Andre Monteiro Manoel.

This is a collaborative project with Lenka Zdeborová, Theoretical Physics Institute (CEA) funded by the CEA program drf-impulsion.

In many scientific fields, the data acquisition devices have benefited of hardware improvement to increase the resolution of the observed phenomena, leading to ever larger datasets. While the dimensionality has increased, the number of samples available is often limited, due to physical or financial limits. This is a problem when these data are processed with estimators that have a large sample complexity, such as multivariate statistical models. In that case it is very useful to rely on structured priors, so that the results reflect the state of knowledge on the phenomena of interest. The study of the human brain activity through high-field MRI belongs among these problems, with up to 10^6 features, yet a set of observations limited by cost and participant comfort.

We are missing fast estimators for multivariate models with structured priors, that furthermore provide statistical control on the solution. Approximate message passing methods are designed to work optimally with low-sample-complexity, they accommodate rather generic class of priors and come with an estimation of statistical significance. They are therefore well suited for our purposes.

We want to join forces to design a new generation of inverse problem solvers that can take into account the complex structure of brain images and provide guarantees in the low-sample-complexity regime. To this end, we will first adapt AMP to the brain mapping setting, using first standard sparsity priors (e.g. Gauss-Bernoulli) on the model. We will then consider more complex structured priors that control the variation of the learned image patterns in space. Crucial gains are expected from the use of the EM algorithm for parameter setting, that comes naturally with AMP. We will also examine the estimators provided by AMP for statistical significance. BrainAMP will design a reference inference toolbox released as a generic open source library. We expect a 3- to 10-fold improvement in CPU time, that will benefit to large-scale brain mapping investigations.

8.1.3. *iConnectom project*

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Elvis Dohmatob.

This is a Digiteo project (2014-2017).

Mapping brain functional connectivity from functional Magnetic Resonance Imaging (MRI) data has become a very active field of research. However, analysis tools are limited and many important tasks, such as the empirical definition of brain networks, remain difficult due to the lack of a good framework for the statistical modeling of these networks. We propose to develop population models of anatomical and functional connectivity data to improve the alignment of subjects brain structures of interest while inferring an average template of these structures. Based on this essential contribution, we will design new statistical inference procedures to compare the functional connections between conditions or populations and improve the sensitivity of connectivity analysis performed on noisy data. Finally, we will test and validate the methods on multiple datasets and distribute them to the brain imaging community.

8.1.4. *MetaCog project*

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Jérome Dockès.

This is a Digicosme project (2016-2019) and a collaboration with Fabian Suchanek (Telecom Paritech).

Understanding how cognition emerges from the billions of neurons that constitute the human brain is a major open problem in science that could bridge natural science –biology– to humanities –psychology. Psychology studies performed on humans with functional Magnetic Resonance Imaging (fMRI) can be used to probe the full repertoire of high-level cognitive functions. While analyzing the resulting image data for a given experiment is a relatively well-mastered process, the challenges in comparing data across multiple datasets poses serious limitation to the field. Indeed, such comparisons require to pool together brain images acquired under different settings and assess the effect of different *experimental conditions* that correspond to psychological effects studied by neuroscientists.

Such meta-analyses are now becoming possible thanks to the development of public data resources –OpenfMRI http://openfmri.org and NeuroVault http://neurovault.org. As many others, researchers of the Parietal team understand these data sources well and contribute to them. However, in such open-ended context, the description of experiments in terms of cognitive concepts is very difficult: there is no universal definition of cognitive terms that could be employed consistently by neuroscientists. Hence meta-analytic studies loose power and specificity. On the other hand, http://brainspell.org provide a set of curated annotation, albeit on much less data, that can serve as a seed or a ground truth to define a consensual ontology of cognitive concepts. Relating these terms to brain activity poses another challenge, of statistical nature, as brain patterns form high-dimensional data in perspective with the scarcity and the noise of the data.

The purpose of this project is to learn a semantic structure in cognitive terms from their occurrence in brain activations. This structure will simplify massive multi-label statistical-learning problems that arise in brain mapping by providing compact representations of cognitive concepts while capturing the imprecision on the definition these concepts.

8.1.5. *CDS2*

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Guillaume Lemaitre.

CDS2 is an "Strategic research initiatice" of the Paris Saclay University Idex http://datascience-paris-saclay.fr. Although it groups together many partners of the Paris Saclay ecosystem, Parietal has been deeply involved in the project. It currently funds a post-doc for Guillume Lemaitre.

8.2. National Initiatives

8.2.1. *ANR*

8.2.1.1. *MultiFracs project*

Participants: Philippe Ciuciu [Correspondant], Daria La Rocca.

The scale-free concept formalizes the intuition that, in many systems, the analysis of temporal dynamics cannot be grounded on specific and characteristic time scales. The scale-free paradigm has permitted the relevant analysis of numerous applications, very different in nature, ranging from natural phenomena (hydrodynamic turbulence, geophysics, body rhythms, brain activity,...) to human activities (Internet traffic, population, finance, art,...).

Yet, most successes of scale-free analysis were obtained in contexts where data are univariate, homogeneous along time (a single stationary time series), and well-characterized by simple-shape local singularities. For such situations, scale-free dynamics translate into global or local power laws, which significantly eases practical analyses. Numerous recent real-world applications (macroscopic spontaneous brain dynamics, the central application in this project, being one paradigm example), however, naturally entail large multivariate data (many signals), whose properties vary along time (non-stationarity) and across components (non-homogeneity), with potentially complex temporal dynamics, thus intricate local singular behaviors.

These three issues call into question the intuitive and founding identification of scale-free to power laws, and thus make uneasy multivariate scale-free and multifractal analyses, precluding the use of univariate methodologies. This explains why the concept of scale-free dynamics is barely used and with limited successes in such settings and highlights the overriding need for a systematic methodological study of multivariate scale-free and multifractal dynamics. The Core Theme of MULTIFRACS consists in laying the theoretical foundations of a practical robust statistical signal processing framework for multivariate non homogeneous scale-free and multifractal analyses, suited to varied types of rich singularities, as well as in performing accurate analyses of scale-free dynamics in spontaneous and task-related macroscopic brain activity, to assess their natures, functional roles and relevance, and their relations to behavioral performance in a timing estimation task using multimodal functional imaging techniques.

This overarching objective is organized into 4 Challenges:

1. Multivariate scale-free and multifractal analysis,
2. Second generation of local singularity indices,
3. Scale-free dynamics, non-stationarity and non-homogeneity,
4. Multivariate scale-free temporal dynamics analysis in macroscopic brain activity.

8.2.1.2. BrainPedia project

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux.

BrainPedia is an ANR JCJC (2011-2015) which addresses the following question: Neuroimaging produces huge amounts of complex data that are used to better understand the relations between brain structure and function. While the acquisition and analysis of this data is getting standardized in some aspects, the neuroimaging community is still largely missing appropriate tools to store and organize the knowledge related to the data. Taking advantage of common coordinate systems to represent the results of group studies, coordinate-based meta-analysis approaches associated with repositories of neuroimaging publications provide a crude solution to this problem, that does not yield reliable outputs and looses most of the data-related information. In this project, we propose to tackle the problem in a statistically rigorous framework, thus providing usable information to drive neuroscientific knowledge and questions.

8.2.1.3. Niconnect project

Participants: Bertrand Thirion, Gaël Varoquaux [Correspondant], Alexandre Abraham, Kamalaker Reddy Dadi, Darya Chyzhyk, Mehdi Rahim.

- **Context:** The NiConnect project (2012-2016) arises from an increasing need of medical imaging tools to diagnose efficiently brain pathologies, such as neuro-degenerative and psychiatric diseases or lesions related to stroke. Brain imaging provides a non-invasive and widespread probe of various features of brain organization, that are then used to make an accurate diagnosis, assess brain rehabilitation, or make a prognostic on the chance of recovery of a patient. Among different measures extracted from brain imaging, functional connectivity is particularly attractive, as it readily probes the integrity of brain networks, considered as providing the most complete view on brain functional organization.

- **Challenges:** To turn methods research into popular tool widely usable by non specialists, the NiConnect project puts specific emphasis on producing high-quality open-source software. NiConnect addresses the many data analysis tasks that extract relevant information from resting-state fMRI datasets. Specifically, the scientific difficulties are *i)* conducting proper validation of the models and tools, and *ii)* providing statistically controlled information to neuroscientists or medical doctors. More importantly, these procedures should be robust enough to perform analysis on limited quality data, as acquiring data on diseased populations is challenging and artifacts can hardly be controlled in clinical settings.

- **Outcome of the project:** In the scope of computer science and statistics, NiConnect pushes forward algorithms and statistical models for brain functional connectivity. In particular, we are investigating structured and multi-task graphical models to learn high-dimensional multi-subject brain connectivity models, as well as spatially-informed sparse decompositions for segmenting structures from brain imaging. With regards to neuroimaging methods development, NiConnect provides systematic comparisons and evaluations of connectivity biomarkers and a software library embedding best-performing state-of-the-art approaches. Finally, with regards to medical applications, the NiConnect project also plays a support role in on going medical studies and clinical trials on neurodegenerative diseases.

- **Consortium**
  - Parietal Inria research team: applied mathematics and computer science to model the brain from MRI

    – LIF INSERM research team: medical image data analysis and modeling for clinical applications

    – CATI center: medical image processing center for large scale brain imaging studies

    – Henri-Mondor hospital neurosurgery and neuroradiology: clinical teams conducting research on treatments for neurodegenerative diseases, in particular Huntington and Parkinson diseases

    – Logilab: consulting in scientific computing

8.3. European Initiatives

8.3.1. FP7 & H2020 Projects

8.3.1.1. HBP

Title: The Human Brain Project

Programm: FP7

Duration: October 2013 - September 2016

Coordinator: EPFL

Partners: 100 across Europe

Inria contact: Olivier Faugeras

Understanding the human brain is one of the greatest challenges facing 21st century science. If we can rise to the challenge, we can gain profound insights into what makes us human, develop new treatments for brain diseases and build revolutionary new computing technologies. Today, for the first time, modern ICT has brought these goals within sight. The goal of the Human Brain Project, part of the FET Flagship Programme, is to translate this vision into reality, using ICT as a catalyst for a global collaborative effort to understand the human brain and its diseases and ultimately to emulate its computational capabilities. The Human Brain Project will last ten years and will consist of a ramp-up phase (from month 1 to month 36) and subsequent operational phases. This Grant Agreement covers the ramp-up phase. During this phase the strategic goals of the project will be to design, develop and deploy the first versions of six ICT platforms dedicated to Neuroinformatics, Brain Simulation, High Performance Computing, Medical Informatics, Neuromorphic Computing and Neurorobotics, and create a user community of research groups from within and outside the HBP, set up a European Institute for Theoretical Neuroscience, complete a set of pilot projects providing a first demonstration of the scientific value of the platforms and the Institute, develop the scientific and technological capabilities required by future versions of the platforms, implement a policy of Responsible Innovation, and a programme of transdisciplinary education, and develop a framework for collaboration that links the partners under strong scientific leadership and professional project management, providing a coherent European approach and ensuring effective alignment of regional, national and European research and programmes. The project work plan is organized in the form of thirteen subprojects, each dedicated to a specific area of activity. A significant part of the budget will be used for competitive calls to complement the collective skills of the Consortium with additional expertise.

8.3.2. Collaborations in European Programs, Except FP7 & H2020

Program: Marie Curie

Project acronym: Neuroimaging Power

Project title: Effect size and power for neuroimaging.

Duration: mois année début - mois année fin

Coordinator: Inria

Other partners: Univ. Stanford, USA

Abstract: There is an increasing concern about statistical power in neuroscience research. Critically, an underpowered study has poor predictive power. Findings from a low-power study are unlikely to be reproducible, and thus a power analysis is a critical component of any paper. This project aims to promote and facilitate the use of power analyses.A key component of a power analysis is the specification of an effect size. However, in neuroimaging, there is no standardised way to communicate effect sizes, which makes the choice of an appropriate effect size a formidable task. The best way today to perform a power analysis is by collecting a pilot data set, a very expensive practice. To eliminate the need for pilot data, we will develop a standardised measure of effect size taking into account the spatial variance and the uncertainty of the measurements. Communicating effect sizes in new publications will facilitate the use of power analyses.To further alleviate the need for pilot data, we will provide a library of effect sizes for different tasks and contrasts, using open data projects in neuroimaging. We will integrate our effect size estimator in open repositories NeuroVault and OpenfMRI. Consequently, these effect sizes can then serve as a proxy for a pilot study, and as such, a huge cost in the design of an experiment is eliminated.A new experiment will not be identical to the open data and as such the hypothesised parameters might not be fully accurate. To address this issue, we present a flexible framework to analyse data mid-way without harming the control of the type I error rate. Such a procedure will allow re-evaluating halfway an experiment whether it is useful to continue a study, and how many more subjects are needed for statistically sound inferences.To make our methods maximally available, we will write a software suite including all these methods in different programming platforms and we will provide a GUI to further increase the use of power analyses.

8.4. International Initiatives

8.4.1. MetaMRI

Title: Machine learning for meta-analysis of functional neuroimaging data

International Partner (Institution - Laboratory - Researcher):

Stanford (United States) - Department of Psychology - Russ Poldrack

Start year: 2015

See also: https://team.inria.fr/metamri

Neuroimaging produces huge amounts of complex data that are used to better understand the relations between brain structure and function. Observing that the neuroimaging community is still largely missing appropriate tools to store and organize the knowledge related to the data, Parietal team and Poldrack's lab, have decided to join forces to set up a framework for functional brain image meta-analysis, i.e. a framework in which several datasets can be jointly analyzed in order to accumulate information on the functional specialization of brain regions. MetaMRI will build upon Poldrack's lab expertise in handling, sharing and analyzing multi-protocol data and Parietal's recent developments of machine learning libraries to develop a new generation of meta-analytic tools.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the Organizing Committees

- **Bertrand Thirion**: Organization for Human Brain Mapping.

9.1.1.2. Reviewer

- **Philippe Ciuciu**: IEEE ISBI (15 papers), IEEE ICASSP (10 papers), IEEE ICIP (5 papers), NIPS (4 papers), EUSIPCO (5 papers).
- **Bertrand Thirion**: IPMI, MICCAI, NIPS, ISBI, PRNI, AISTATS
- **Gaël Varoquaux**: IEEE ICASSP, MICCAI, NIPS, IPMI, ICML

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

- **Bertrand Thirion**: Medical Image Analysis, Frontiers in brain imaging
- **Gaël Varoquaux**: Frontiers in NeuroInformatics, Frontiers in brain imaging methods, NeuroImage

9.1.2.2. Reviewer - Reviewing Activities

- **Philippe Ciuciu**: Reviewer for Neuroimage, IEEE Signal Processing Letters, Signal Processing, IEEE Trans. Medical Imaging, Plos One, Plos Comput. Biology, Frontiers in Neuroscience.
- **Bertrand Thirion**: Human Brain Mapping, IEEE TMI, MedIA, NeuroImage, PNAS
- **Gaël Varoquaux**: NeuroImage, JSTSP, PNAS, HBM, PLOS Comp Bio, Gigascience
- **Olivier Grisel**: Journal of Machine Learning Research (software track).

9.1.3. Invited Talks

9.1.3.1. Bertrand Thirion

- February: invited talk at the *Imagerie du Vivant* National congress, entitled *Large-scale analyses in functional brain Imaging*.
- February: presentation at the Pasadena working group of the Digicosme Labex.
- April: invited presentation at European Neuroscience institute, Paris, entitled *Seeing it all: Convolutional network layers map the function of the human visual system*.
- April: presentation Functional connectomicts, at DTU Copenhagen, entitled *from large-scale estimators to empirical validation*.
- May: Talk at Atlas workshop, Grenoble, entitled *Learning representations from functional brain imaging*.
- June: organizer of a table ronde at the *Futur en Seine* event entitled *Computational methods for neurosciences & medical imaging*.
- October: talk at MPI Psychiatry, Munich, entitled *Machine learning for neuroimaging: current challenges and solutions*.
- June: Talk at Neurostic workshop, Grenoble, entitled *Learning representations from functional brain imaging*.
- October: Invited talk by the ITMO Neuroscience, Bordeaux, entitled *Working with large data samples: the case of human brain imaging*.

9.1.3.2. Philippe Ciuciu

- 12/16: IEEE Lecture at University of British Columbia (Vancouver, Canada): *Sparkling: Novel non-Cartesian sampling schemes for accelerated 2D anatomical imaging at 7 Tesla*.
- 12/16: Pacific Parkinson's research center (Vancouver, Canada): *Impact of perceptual learning on resting-state brain dynamics in fMRI: A supervised classification study*.
- 09/16: GdR d'Analyse Multifractale (Avignon, France): *Convergence of neural activity to multifractal attractors in MEG predicts learning*.
- 08/16: invitation to the Special session entitled "Unraveling brain networks from functional neuroimaging data" at EUSIPCO'16 (Budapest, Hungary): *Impact of perceptual learning on resting-state fMRI connectivity: A supervised classification study*.
- 06/16: Journées scientifiques d'Inria (Rennes, France): *Compressive Sampling in MRI*.

- 06/16: Inria Sophia-Antipolis, équipe Athena. *New physically plausible compressive sampling schemes for MRI: First results at 7 Tesla*

- 05/16: University of Geneva (Campus BioTech, Geneva, Switzlerand): *Convergence to asymptotic Multifractal dynamics in the brain predicts learning.*

- 02/16: Grenoble Institut of Neurosciences (Grenoble, France): *Physically plausible trajectories for Compressed Sensing in MRI.*

- 02/16: Workshop on 7 Tesla scanner at NeuroSpin (Gif-sur-Yvette, France) *Compressed sensing for high resolution MRI at 7 Tesla.*

- 01/16: Cosmostat lab, IRFU/CEA. *On the generation of compressed sampling schemes in MRI.*

9.1.3.3. Loïc Estève

- EuroScipy 2016: scikit-learn tutorial

- Budapest BI 2016 : scikit-learn tutorial and talk "Recent developments in scikit-learn and joblib"

9.1.3.4. Olivier Grisel

- PyData Berlin and PyData Paris 2016: *"Predictive modeling with Python, trends and tools*

- invited talk on *Some recent developments in Deep Learning researc* at Strata London 2016.

9.1.3.5. Gaël Varoquaux

- Paris Open Source summit 2016: scikit-learn, the vision and the community

- EuroScipy 2016 (Erlangen): keynote: "On writing code the science"

- Open Data Science Conference 2017 (London): keynote: "The code of data science"

- EuroPython 2016 (Bilbao): keynote "Scientists meet web dev: how Python became the language of data"

- PiterPy 2016 (St Petersbourg): keynote: "Python for data"

- Facebook AI Research: some statistical learning problems in brain imaging

- GDR ISIS Imagerie medicale: prediction de pathologies psychiatriques à partir d'imagerie fonction-nelle de repos

- Brain network analysis workshop, MICCAI 2016 (Athenes): keynote

- Journée Graphes et neuroscience à Marseilles: Machine learning on brain graphes

- Séminaire débat sur le Big data en Neuroscience, Lyon

- Seminar Max Planck Institute Leipzig: data mining for neuroimaging

- Seminar Telecom ParisTech: randomized methods for high-dimensional statistical learning

- Séminaire d'équipe Asclepios: Quelques problèmes d'apprentissage sur des images cérébrales

9.1.4. Leadership within the Scientific Community

- Gaël Varoquaux: Chair of the steering committee, IEEE PRNI

- Bertrand Thirion: member of the *Committee on Best Practices in Data Analysis and Sharing* for the OHBM community.

9.1.5. Scientific Expertise

- Philippe Ciuciu: ANR JC, NSERC au Canada, FWO

- Bertrand Thirion: ANR, NWO, NSF

- Gaël Varoquaux: Membre de la Commission Expertises Scientifiques, (CE23) ANR

- Olivier Grisel did 3 days of consulting with the CTO of the Therapixel startup to share expertise on the use of Deep Learning for the predictive analysis of 3D imaging data.

9.1.6. Research Administration

9.1.6.1. Philippe Ciuciu

- 03/16: Involvement in the CEA visiting committee on High Performance Computing.
- 05/16: Member of a Comité de sélection for hiring an Assistant Professor in Paris-Saclay University (Section 61 of CNU).
- 06/16: Member of the Inria scientific commission in charge of ranking PhD and post-doctoral applicants as well as delegations of Assistant Professors to Inria.

9.1.6.2. Bertrand Thirion

- Leader of the Datasense axis of the Digicosme Labex
- Member of the STIC department committee Paris-Saclay University and of the bureau thereof.
- DSA Saclay.

9.1.6.3. Gaël Varoquaux

- Member of "Comité de suivi doctoral", Inria Saclay
- Member of "Comité cluster", Inria Saclay
- Member of "Commission de Développement Technologique", Inria Saclay
- Member of the directorate of the Paris-Saclay CDS (Center for Data Science)

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

9.2.1.1. Bertrand Thirion

Master : Brain Computer interface and Functional Neuroimaging, 12 heures équivalent TD, niveau M2, ENS Cachan

9.2.1.2. Philippe Ciuciu

Master 2 : "Functional MRI: From data acquisition to analysis", 3h, Univ. Paris V René Descartes & Télécom-Paristech, Master of Biomedical Engineering

Master 2 : "FMRI data analysis", 3h, Univ. Paris-Saclay, Master of medical Physics

9.2.1.3. Gaël Varoquaux

Master 2 : "Brain functional connectivity analysis", 7h, Univ. Paris V René Descartes & Télécom-Paristech, Master of Biomedical Engineering

Master 2 : "Machine learning with scikit-learn", 2h, ENSAE

Master 2 : "Advanced Machine learning with scikit-learn", 3h, Centrale Paris, MSc in data sciences & business analytics

Ecole d'été multidisciplinaire analyse de données, Rennes, 1h

OHBM 2016: course on machine learning for cognitive neuroimaging 30mn

PRNI 2016: nilearn for machine learning on brain images, 8h

Max Planck Institute Leipzig: nilearn for machine learning on brain images, 8h

9.2.2. Supervision

9.2.2.1. Bertrand Thirion

PhD in progress: Elvis Dohmatob,

PhD in progress: Arthur Mensch,

PhD in progress: Andrés Hoyos Idrobo

9.2.2.2. Philippe Ciuciu

PhD defended: Aina Frau-Pascual, "Statistical models for the analysis of BOLD and ASL Magnetic Resonance modalities to study brain function and disease", University of Grenoble-Alpes (doctoral school: Mathématiques, Sciences et Technologies de l'Information, Informatique), defense: 19/12/2016, Advisors: Florence Forbes (Dir), Philippe Ciuciu (Co-Dir)

PhD in progress: Carole Lazarus, "Physically plausible compressed sensing for high resolution MRI at 7 Tesla in Humans" starting date: October 2015 (Univ. Paris-Saclay, doctoral school: EOBE). Advisors: Philippe Ciuciu (Dir), Alexandre Vignaud (Co-Dir)

PhD in progress: Loubna El Gueddari, "Parallel proximal algorithms for compressed sensing MRI reconstruction. Applications in ultra-high magnetic field imaging", starting date: October 2016 (Univ. Paris-Saclay, doctoral school: EOBE). Advisors: Philippe Ciuciu (Dir) and Jean-Christophe Pesquet (Co-Dir, Prof. at Centrale-Supélec)

9.2.2.3. Gael Varoquaux

PhD defended: Alexandre Abraham

PhD in progress: Elvis Dohmatob,

PhD in progress: Arthur Mensch,

PhD in progress: Andrés Hoyos Idrobo

9.2.3. Juries

9.2.3.1. Bertrand Thirion

- 04/29: Reviewer of Niklas Kasenburg PhD Thesis , Univ. Copenhagen, Denmark.
- 01/12: Examiner of Simona Schiavi PhD Thesis, Univ. Paris Saclay.
- 14/12: Reviewer of Olivier Marre habilitation, Paris.
- 15/12: Reviewer of Maite Termenon PhD thesis, Univ. Grenoble.

9.2.3.2. Philippe Ciuciu

- 04/16: Reviewer of Aiping Liu's PhD thesis (ECCS Dpt, Univ. British Columbia, Vancouvern Canada) entitled *"Brain Connectivity Network Modeling using fMRI signals"*
- 05/16: Reviewer of Andrea Laruelo-Fernandez's PhD thesis (INP Toulouse-IRIT- ENSEEIHT) entitled *"Integration of magnetic resonance spectroscopic imaging into the radiotherapy treatment planning"*
- 09/16: Examinor of Mohanad Albughdadi's PhD thesis (INP Toulouse-IRIT- ENSEEIHT) entitled *" Bayesian joint detection-estimation in functional MRI with automatic parcellation and functional constraints"*
- 10/16: Reviewer of Sébastien Combrexelle's PhD thesis (INP Toulouse-IRIT- ENSEEIHT) entitled *"Multifractal analysis for multivariate data with application to remote sensing"*.
- 10/16: Co-director of Aina Frau-Pascual's PhD thesis (see above).

9.2.3.3. Gaël Varoquaux

- 06/16: Examiner of Alberto García Durán, PhD Thesis, UTC Compiegne.

9.3. Popularization

9.3.1. Gaël Varoquaux

- Unithé ou Café, Inria Saclay Ile de France
- Atelier IHEST Les mots du numérique - 17 novembre

9.3.2. Loïc Estève

Software Carpentry workshops:

- git course at UNIC in Gif-sur-Yvette March 29-30
- helper at "Scientific Programming with Python and Software Engineering Best Practices" workshop, April 28-29 at Télécom Paris
- git course at Proto 204, May 24-25

Mentor at Startup Weekend Artificial Intelligence, November 4-6.

9.3.3. Olivier Grisel

"La tête au carré" radio show on France Inter in January 2016 to share his expertise and opinion on the use and impacts of Big Data and predictive algorithms [0].

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] A. ABRAHAM, M. MILHAM, A. DI MARTINO, R. C. CRADDOCK, D. SAMARAS, B. THIRION, G. VAROQUAUX.*Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example*, in "NeuroImage", November 2016 [*DOI :* 10.1016/J.NEUROIMAGE.2016.10.045], https://hal.inria.fr/hal-01398867.

[2] M. ALBUGHDADI, L. CHAARI, J.-Y. TOURNERET, F. FORBES, P. CIUCIU.*A Bayesian Non-Parametric Hidden Markov Random Model for Hemodynamic Brain Parcellation*, in "Signal Processing", 2017 [*DOI :* 10.1016/J.SIGPRO.2017.01.005], https://hal.archives-ouvertes.fr/hal-01426385.

[3] V. BORGHESANI, F. PEDREGOSA, M. BUIATTI, A. AMADON, E. EGER, M. PIAZZA.*Word meaning in the ventral visual path: a perceptual to conceptual gradient of semantic coding*, in "NeuroImage", September 2016 [*DOI :* 10.1016/J.NEUROIMAGE.2016.08.068], http://hal.upmc.fr/hal-01372551.

[4] C. BOYER, N. CHAUFFERT, P. CIUCIU, J. KAHN, P. WEISS.*On the generation of sampling schemes for Magnetic Resonance Imaging*, in "SIAM Journal on Imaging Sciences", December 2016, vol. 9, n⁰ 4, p. 2039-20972, https://hal.archives-ouvertes.fr/hal-01373758.

[5] D. BZDOK, G. HARTWIGSEN, A. REID, A. R. LAIRD, P. T. FOX, S. EICKHOFF.*Left inferior parietal lobe engagement in social cognition and language*, in "Neuroscience and Biobehavioral Reviews", April 2016 [*DOI :* 10.1016/J.NEUBIOREV.2016.02.024], https://hal.archives-ouvertes.fr/hal-01350512.

[6] D. BZDOK, G. VAROQUAUX, O. GRISEL, M. EICKENBERG, C. POUPON, B. THIRION.*Formal Models of the Network Co-occurrence Underlying Mental Operations*, in "PLoS Computational Biology", June 2016 [*DOI :* 10.1371/JOURNAL.PCBI.1004994], https://hal.archives-ouvertes.fr/hal-01338307.

[7] D. BZDOK, G. VAROQUAUX, B. THIRION.*Neuroimaging Research: From Null-Hypothesis Falsification to Out-of-sample Generalization*, in "Educational and Psychological Measurement", August 2016, https://hal.archives-ouvertes.fr/hal-01338313.

[8] N. CHAUFFERT, P. CIUCIU, J. KAHN, P. WEISS.*A Projection Method on Measures Sets*, in "Constructive Approximation", February 2017, vol. 45, n⁰ 1, p. 83 - 111 [*DOI :* 10.1007/S00365-016-9346-2], https://hal.inria.fr/hal-01432720.

[9] N. CHAUFFERT, P. WEISS, J. KAHN, P. CIUCIU.*A projection algorithm for gradient waveforms design in Magnetic Resonance Imaging*, in "IEEE Transactions on Medical Imaging", 2016 [*DOI :* 10.1109/TMI.2016.2544251], https://hal.inria.fr/hal-01317939.

---

[0]https://www.franceinter.fr/emissions/la-tete-au-carre/la-tete-au-carre-05-janvier-2016

[10] M. EICKENBERG, A. GRAMFORT, G. VAROQUAUX, B. THIRION.*Seeing it all: Convolutional network layers map the function of the human visual system*, in "NeuroImage", 2016 [*DOI :* 10.1016/J.NEUROIMAGE.2016.10.001], https://hal.inria.fr/hal-01389809.

[11] S. EICKHOFF, B. THIRION, G. VAROQUAUX, D. BZDOK.*Connectivity-Based Parcellation: Critique and Implications*, in "Human Brain Mapping", January 2016, 22 [*DOI :* 10.1002/HBM.22933], https://hal.archives-ouvertes.fr/hal-01184563.

[12] K. J. GORGOLEWSKI, T. AUER, V. D. CALHOUN, C. R. CRADDOCK, S. DAS, E. P. DUFF, G. FLANDIN, S. S. GHOSH, T. GLATARD, Y. O. HALCHENKO, D. A. HANDWERKER, M. HANKE, D. KEATOR, X. LI, Z. MICHAEL, C. MAUMET, N. B. NICHOLS, T. E. NICHOLS, J. PELLMAN, J.-B. POLINE, A. ROKEM, G. SCHAEFER, V. SOCHAT, W. TRIPLETT, J. A. TURNER, G. VAROQUAUX, R. A. POLDRACK.*The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments*, in "Scientific Data ", June 2016, vol. 3, 160044 [*DOI :* 10.1038/SDATA.2016.44], http://www.hal.inserm.fr/inserm-01345616.

[13] A. HOYOS-IDROBO, G. VAROQUAUX, J. KAHN, B. THIRION.*Recursive nearest agglomeration (ReNA): fast clustering for approximation of structured signals*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", August 2016, https://hal.inria.fr/hal-01366651.

[14] F. LIEM, G. VAROQUAUX, J. KYNAST, F. BEYER, S. K. MASOULEH, J. M. HUNTENBURG, L. LAMPE, M. RAHIM, A. ABRAHAM, R. C. CRADDOCK, S. RIEDEL-HELLER, T. LUCK, M. LOEFFLER, M. L. SCHROETER, A. V. WITTE, D. S. MARGULIES.*Predicting brain-age from multimodal imaging data captures cognitive impairment*, in "NeuroImage", November 2016 [*DOI :* 10.1016/J.NEUROIMAGE.2016.11.005], https://hal.inria.fr/hal-01403005.

[15] M. RAHIM, B. THIRION, C. COMTAT, G. VAROQUAUX.*Transmodal Learning of Functional Networks for Alzheimer's Disease Prediction*, in "IEEE Journal of Selected Topics in Signal Processing", October 2016, vol. 10, n° 7, p. 1204 - 1213 [*DOI :* 10.1109/JSTSP.2016.2600400], https://hal.inria.fr/hal-01353728.

[16] G. VAROQUAUX, P. A. REDDY RAAMANA, D. A. ENGEMANN, A. A. HOYOS-IDROBO, Y. A. SCHWARTZ, B. A. THIRION.*Assessing and tuning brain decoders: cross-validation, caveats, and guidelines*, in "NeuroImage", 2016 [*DOI :* 10.1016/J.NEUROIMAGE.2016.10.038], https://hal.archives-ouvertes.fr/hal-01332785.

Invited Conferences

[17] E. DOHMATOB.*A simple algorithm for computing Nash-equilibria in incomplete information games*, in "OPT2016 – NIPS workshop on optimization for machine learning", Barcelona, Spain, December 2016, https://hal.inria.fr/hal-01384460.

[18] E. DOHMATOB, M. EICKENBERG, B. THIRION, G. VAROQUAUX.*Local Q-Linear Convergence and Finite-time Active Set Identification of ADMM on a Class of Penalized Regression Problems*, in "ICASSP, International Conference on Acoustics, Speach, and Signal Processing", China, March 2016, http://icassp2016.org/default.asp , https://hal.archives-ouvertes.fr/hal-01265372.

[19] E. DOHMATOB, A. MENSCH, G. VAROQUAUX, B. THIRION.*Learning brain regions via large-scale online structured sparse dictionary-learning*, in "Neural Information Processing Systems (NIPS)", Barcelona, Spain, December 2016, https://hal.inria.fr/hal-01369134.

International Conferences with Proceedings

[20] M. ALBUGHDADI, L. CHAARI, F. FORBES, J.-Y. TOURNERET, P. CIUCIU.*Multi-subject joint parcellation detection estimation in functional MRI*, in "13th IEEE International Symposium on Biomedical Imaging", Prague, Czech Republic, April 2016, https://hal.inria.fr/hal-01261982.

[21] E. BELILOVSKY, G. VAROQUAUX, M. B. BLASCHKO.*Testing for Differences in Gaussian Graphical Models: Applications to Brain Connectivity*, in "Neural Information Processing Systems (NIPS) 2016", Barcelona, Spain, December 2016, https://hal.inria.fr/hal-01248844.

[22] A. MENSCH, J. MAIRAL, B. THIRION, G. VAROQUAUX.*Dictionary Learning for Massive Matrix Factorization*, in "International Conference on Machine Learning", New York, United States, Proceedings of the 33rd Internation Conference on Machine Learning, June 2016, vol. 48, p. 1737–1746, https://hal.archives-ouvertes.fr/hal-01308934.

[23] H. PELLÉ, P. CIUCIU, M. RAHIM, E. DOHMATOB, P. ABRY, V. VAN WASSENHOVE.*Multivariate Hurst Exponent Estimation in FMRI. Application to Brain Decoding of Perceptual Learning*, in "13th IEEE International Symposium on Biomedical Imaging", Prague, Czech Republic, April 2016, https://hal.inria.fr/hal-01261976.

[24] M. RAHIM, P. CIUCIU, S. BOUGACHA.*Impact of perceptual learning on resting-state fMRI connectivity: A supervised classification study*, in "Eusipco 2016", Budapest, Hungary, August 2016, https://hal.inria.fr/hal-01355478.

Conferences without Proceedings

[25] Y. BEKHTI, D. STROHMEIER, M. JAS, R. BADEAU, A. GRAMFORT.*M/EEG source localization with multi-scale time-frequency dictionaries*, in "6th International Workshop on Pattern Recognition in Neuroimaging (PRNI)", Trento, Italy, June 2016 [*DOI :* 10.1109/PRNI.2016.7552337], https://hal.archives-ouvertes.fr/hal-01313567.

[26] K. DADI, A. ABRAHAM, M. RAHIM, B. THIRION, G. VAROQUAUX.*Comparing functional connectivity based predictive models across datasets*, in "PRNI 2016: 6th International Workshop on Pattern Recognition in Neuroimaging", Trento, Italy, June 2016, https://hal.inria.fr/hal-01319131.

[27] A. HOYOS-IDROBO, G. VAROQUAUX, B. THIRION.*Fast brain decoding with random sampling and random projections*, in "PRNI 2016: the 6th International Workshop on Pattern Recognition in Neuroimaging", Trento, Italy, June 2016, https://hal.inria.fr/hal-01313814.

[28] C. LAROCHE, H. PAPADOPOULOS, M. KOWALSKI, G. RICHARD.*Genre specific dictionaries for harmonic/percussive source separation*, in "ISMIR The 17th International Society for Music Information Retrieval Conference", New York, United States, August 2016, https://hal.archives-ouvertes.fr/hal-01353252.

[29] C. LAROCHE, H. PAPADOPOULOS, M. KOWALSKI, G. RICHARD.*Drum extraction in single channel audio signals using multi-layer non negative matrix factor deconvolution*, in "ICASSP", Nouvelle Orleans, United States, March 2017, https://hal.archives-ouvertes.fr/hal-01438851.

[30] A. MENSCH, J. MAIRAL, G. VAROQUAUX, B. THIRION.*Subsampled online matrix factorization with convergence guarantees*, in "NIPS Workshop on Optimization for Machine Learning", Barcelona, Spain, December 2016, https://hal.archives-ouvertes.fr/hal-01405058.

[31] A. MENSCH, G. VAROQUAUX, B. THIRION.*Compressed Online Dictionary Learning for Fast Resting-State fMRI Decomposition*, in "International Symposium on Biomedical Imaging (ISBI 2016) "From Nano to Macro"", Prague, Czech Republic, 13th International Symposium on Biomedical Imaging (ISBI), IEEE, April 2016, p. 1282-1285 [*DOI :* 10.1109/ISBI.2016.7493501], https://hal.archives-ouvertes.fr/hal-01271033.

[32] G. VAROQUAUX, M. KOWALSKI, B. THIRION.*Social-sparsity brain decoders: faster spatial sparsity*, in "Pattern Recognition in NeuroImaging", Trento, Italy, June 2016, https://hal.archives-ouvertes.fr/hal-01334551.

Scientific Books (or Scientific Book chapters)

[33] B. THIRION.*Functional Neuroimaging Group Studies*, in "Handbook of Neuroimaging Data Analysis", H. OMBAO, M. LINDQUIST, W. THOMPSON, J. ASTON (editors), Handbook of Modern Statistical methods, Chapman & Hall / CRC, November 2016, https://hal.inria.fr/hal-01419347.

Other Publications

[34] M. ALBUGHDADI, L. CHAARI, J.-Y. TOURNERET, F. FORBES, P. CIUCIU.*Hemodynamic Brain Parcellation Using A Non-Parametric Bayesian Approach*, February 2016, working paper or preprint, https://hal.inria.fr/hal-01275622.

[35] E. BELILOVSKY, K. KASTNER, G. VAROQUAUX, M. BLASCHKO.*Learning to Discover Graphical Model Structures*, May 2016, working paper or preprint, https://hal.inria.fr/hal-01306491.

[36] D. BZDOK, B. T. YEO.*The Future of Data Analysis in the Neurosciences*, August 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01356923.

[37] L. CHAARI, S. BADILLO, T. VINCENT, G. DEHAENE-LAMBERTZ, F. FORBES, P. CIUCIU.*Subject-level Joint Parcellation-Detection-Estimation in fMRI*, January 2016, working paper or preprint, https://hal.inria.fr/hal-01255465.

[38] F. FENG, M. KOWALSKI.*A Unified Approach for Over and Under-Determined Blind Source Separation Based on Both Sparsity and Decorrelation*, April 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01297471.

[39] F. HADJ-SELEM, T. LÖFSTEDT, E. DOHMATOB, V. FROUIN, M. DUBOIS, V. GUILLEMOT, E. DUCHESNAY.*Supplementary materials: Iterative Smoothing Proximal Gradient for Regression with Structured Sparsity*, May 2016, working paper or preprint, https://hal-cea.archives-ouvertes.fr/cea-01324021.

[40] M. KOWALSKI, A. MEYNARD, H.-T. WU.*Convex Optimization approach to signals with fast varying instantaneous frequency*, April 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01199615.

[41] C. LAZARUS, N. CHAUFFERT, J. KAHN, P. WEISS, A. VIGNAUD, P. CIUCIU.*Physically plausible K-space trajectories for Compressed Sensing in MRI: From simulations to real acquisitions*, March 2016, CEA Visiting committee on High Performance Computing, Poster, https://hal.inria.fr/hal-01296496.

[42] A. MENSCH, J. MAIRAL, B. THIRION, G. VAROQUAUX.*Stochastic Subsampling for Factorizing Huge Matrices*, January 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01431618.

[43] M. RAHIM, P. CIUCIU, S. BOUGACHA.*Functional connectivity outperforms scale-free brain dynamics as fMRI predictive feature of perceptual learning underwent in MEG*, February 2016, working paper or preprint, https://hal.inria.fr/hal-01297845.

Project-Team PARSIFAL

Proof search and reasoning with logic specifications

IN COLLABORATION WITH: Laboratoire d'informatique de l'école polytechnique (LIX)

IN PARTNERSHIP WITH:
CNRS

Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Proofs and Verification

Table of contents

Project-Team PARSIFAL

Creation of the Project-Team: 2007 July 01

Keywords:

Computer Science and Digital Science:

2.1.1. - Semantics of programming languages
2.1.11. - Proof languages
2.4.2. - Model-checking
2.4.3. - Proofs
7.4. - Logic in Computer Science

Other Research Topics and Application Domains:

9.4.1. - Computer science
9.4.2. - Mathematics
9.6. - Reproducibility

1. Members

Research Scientists

Beniamino Accattoli [Inria, Researcher]
Kaustuv Chaudhuri [Inria, Researcher]
Stéphane Graham-Lengrand [CNRS, Researcher, HDR]
François Lamarche [Inria, Researcher]
Dale Miller [Team leader, Inria, Researcher]
Lutz Straßburger [Inria, Researcher, HDR]

Technical Staff

Danko Ilik [Inria, until Aug 2016, funded by FP7 ERC PROOFCERT]
Tomer Libal [Inria, funded by FP7 ERC PROOFCERT]
Marco Volpe [Inria, funded by FP7 ERC PROOFCERT]
Noam Zeilberger [Inria, until Oct 2016, funded by FP7 ERC PROOFCERT]

PhD Students

Roberto Blanco Martinez [Inria, funded by FP7 ERC PROOFCERT]
Ulysse Gerard [Inria, funded by FP7 ERC PROOFCERT]
Quentin Heath [Inria, funded by FP7 ERC PROOFCERT]
Sonia Marin [Inria, funded by FP7 ERC PROOFCERT]
François Thiré [ENS Cachan, from Apr 2016]

Post-Doctoral Fellows

Taus Brock-Nannestad [Inria, until Sep 2016, funded by FP7 ERC PROOFCERT]
Giselle Machado Nogueira Reis [Inria, until Jul 2016, funded by FP7 ERC PROOFCERT]
Matthias Puech [Inria, funded by FP7 ERC PROOFCERT]
Hugh Paul Steele [Inria, until Nov 2016, funded by FP7 ERC PROOFCERT]

Visiting Scientists

Jose Carlos Espirito Santo [University of Minho, Portugal, Nov 2016]
Chuck Liang [Hofstra University, NY, USA, Jun 2016]

Administrative Assistant

Christine Aklouche [Inria]

Others

Ameni Chtourou [Inria, intern, from May 2016 until Jul 2016]

Leonardo Augusto Lima Ferreira Da Silva [Inria, Intern, until Feb 2016]

2. Overall Objectives

2.1. Main themes

The aim of the Parsifal team is to develop and exploit *proof theory* and *type theory* in the specification, verification, and analysis of computational systems.

- *Expertise*: the team conducts basic research in proof theory and type theory. In particular, the team is developing results that help with automated deduction and with the manipulation and communication of formal proofs.
- *Design*: based on experience with computational systems and theoretical results, the team develops new logical principles, new proof systems, and new theorem proving environments.
- *Implementation*: the team builds prototype systems to help validate basic research results.
- *Examples*: the design and implementation efforts are guided by examples of specification and verification problems. These examples not only test the success of the tools but also drive investigations into new principles and new areas of proof theory and type theory.

The foundational work of the team focuses on *structural* and *analytic* proof theory, *i.e.*, the study of formal proofs as algebraic and combinatorial structures and the study of proof systems as deductive and computational formalisms. The main focus in recent years has been the study of the *sequent calculus* and of the *deep inference* formalisms.

An important research question is how to reason about computational specifications that are written in a *relational* style. To this end, the team has been developing new approaches to dealing with induction, co-induction, and generic quantification. A second important question is of *canonicity* in deductive systems, *i.e.*, when are two derivations "essentially the same"? This crucial question is important not only for proof search, because it gives an insight into the structure and an ability to manipulate the proof search space, but also for the communication of *proof objects* between different reasoning agents such as automated theorem provers and proof checkers.

Important application areas currently include:

- Meta-theoretic reasoning on functional programs, such as terms in the λ-calculus
- Reasoning about behaviors in systems with concurrency and communication, such as the π-calculus, game semantics, *etc.*
- Combining interactive and automated reasoning methods for induction and co-induction
- Verification of distributed, reactive, and real-time algorithms that are often specified using modal and temporal logics
- Representing proofs as documents that can be printed, communicated, and checked by a wide range of computational logic systems.
- Development of cost models for the evaluation of proofs and programs.

3. Research Program

3.1. General overview

There are two broad approaches for computational specifications. In the *computation as model* approach, computations are encoded as mathematical structures containing nodes, transitions, and state. Logic is used to *describe* these structures, that is, the computations are used as models for logical expressions. Intensional operators, such as the modals of temporal and dynamic logics or the triples of Hoare logic, are often employed to express propositions about the change in state.

The *computation as deduction* approach, in contrast, expresses computations logically, using formulas, terms, types, and proofs as computational elements. Unlike the model approach, general logical apparatus such as cut-elimination or automated deduction becomes directly applicable as tools for defining, analyzing, and animating computations. Indeed, we can identify two main aspects of logical specifications that have been very fruitful:

- *Proof normalization*, which treats the state of a computation as a proof term and computation as normalization of the proof terms. General reduction principles such as β-reduction or cut-elimination are merely particular forms of proof normalization. Functional programming is based on normalization [69], and normalization in different logics can justify the design of new and different functional programming languages [41].

- *Proof search*, which views the state of a computation as a a structured collection of formulas, known as a *sequent*, and proof search in a suitable sequent calculus as encoding the dynamics of the computation. Logic programming is based on proof search [75], and different proof search strategies can be used to justify the design of new and different logic programming languages [73].

While the distinction between these two aspects is somewhat informal, it helps to identify and classify different concerns that arise in computational semantics. For instance, confluence and termination of reductions are crucial considerations for normalization, while unification and strategies are important for search. A key challenge of computational logic is to find means of uniting or reorganizing these apparently disjoint concerns.

An important organizational principle is structural proof theory, that is, the study of proofs as syntactic, algebraic and combinatorial objects. Formal proofs often have equivalences in their syntactic representations, leading to an important research question about *canonicity* in proofs – when are two proofs "essentially the same?" The syntactic equivalences can be used to derive normal forms for proofs that illuminate not only the proofs of a given formula, but also its entire proof search space. The celebrated *focusing* theorem of Andreoli [43] identifies one such normal form for derivations in the sequent calculus that has many important consequences both for search and for computation. The combinatorial structure of proofs can be further explored with the use of *deep inference*; in particular, deep inference allows access to simple and manifestly correct cut-elimination procedures with precise complexity bounds.

Type theory is another important organizational principle, but most popular type systems are generally designed for either search or for normalization. To give some examples, the Coq system [85] that implements the Calculus of Inductive Constructions (CIC) is designed to facilitate the expression of computational features of proofs directly as executable functional programs, but general proof search techniques for Coq are rather primitive. In contrast, the Twelf system [80] that is based on the LF type theory (a subsystem of the CIC), is based on relational specifications in canonical form (*i.e.*, without redexes) for which there are sophisticated automated reasoning systems such as meta-theoretic analysis tools, logic programming engines, and inductive theorem provers. In recent years, there has been a push towards combining search and normalization in the same type-theoretic framework. The Beluga system [81], for example, is an extension of the LF type theory with a purely computational meta-framework where operations on inductively defined LF objects can be expressed as functional programs.

The Parsifal team investigates both the search and the normalization aspects of computational specifications using the concepts, results, and insights from proof theory and type theory.

3.2. Inductive and co-inductive reasoning

The team has spent a number of years in designing a strong new logic that can be used to reason (inductively and co-inductively) on syntactic expressions containing bindings. This work is based on earlier work by McDowell, Miller, and Tiu [71] [70] [76] [86], and on more recent work by Gacek, Miller, and Nadathur [4] [56]. The Parsifal team, along with our colleagues in Minneapolis, Canberra, Singapore, and Cachen, have been building two tools that exploit the novel features of this logic. These two systems are the following.

- Abella, which is an interactive theorem prover for the full logic.
- Bedwyr, which is a model checker for the "finite" part of the logic.

We have used these systems to provide formalize reasoning of a number of complex formal systems, ranging from programming languages to the λ-calculus and π-calculus.

Since 2014, the Abella system has been extended with a number of new features. A number of new significant examples have been implemented in Abella and an extensive tutorial for it has been written [1].

3.3. Developing a foundational approach to defining proof evidence

The team is developing a framework for defining the semantics of proof evidence. With this framework, implementers of theorem provers can output proof evidence in a format of their choice: they will only need to be able to formally define that evidence's semantics. With such semantics provided, proof checkers can then check alleged proofs for correctness. Thus, anyone who needs to trust proofs from various provers can put their energies into designing trustworthy checkers that can execute the semantic specification.

In order to provide our framework with the flexibility that this ambitious plan requires, we have based our design on the most recent advances within the theory of proofs. For a number of years, various team members have been contributing to the design and theory of *focused proof systems* [45] [48] [49] [50] [59] [67] [68] and we have adopted such proof systems as the corner stone for our framework.

We have also been working for a number of years on the implementation of computational logic systems, involving, for example, both unification and backtracking search. As a result, we are also building an early and reference implementation of our semantic definitions.

3.4. Deep inference

Deep inference [61], [63] is a novel methodology for presenting deductive systems. Unlike traditional formalisms like the sequent calculus, it allows rewriting of formulas deep inside arbitrary contexts. The new freedom for designing inference rules creates a richer proof theory. For example, for systems using deep inference, we have a greater variety of normal forms for proofs than in sequent calculus or natural deduction systems. Another advantage of deep inference systems is the close relationship to categorical proof theory. Due to the deep inference design one can directly read off the morphism from the derivations. There is no need for a counter-intuitive translation.

The following research problems are investigated by members of the Parsifal team:

- Find deep inference system for richer logics. This is necessary for making the proof theoretic results of deep inference accessible to applications as they are described in the previous sections of this report.
- Investigate the possibility of focusing proofs in deep inference. As described before, focusing is a way to reduce the non-determinism in proof search. However, it is well investigated only for the sequent calculus. In order to apply deep inference in proof search, we need to develop a theory of focusing for deep inference.

3.5. Proof nets and atomic flows

Proof nets and atomic flows are abstract (graph-like) presentations of proofs such that all "trivial rule permutations" are quotiented away. Ideally the notion of proof net should be independent from any syntactic formalism, but most notions of proof nets proposed in the past were formulated in terms of their relation to the sequent calculus. Consequently we could observe features like "boxes" and explicit "contraction links". The latter appeared not only in Girard's proof nets [58] for linear logic but also in Robinson's proof nets [83] for classical logic. In this kind of proof nets every link in the net corresponds to a rule application in the sequent calculus.

Only recently, due to the rise of deep inference, new kinds of proof nets have been introduced that take the formula trees of the conclusions and add additional "flow-graph" information (see e.g., [6], [5] and [62]. On one side, this gives new insights in the essence of proofs and their normalization. But on the other side, all the known correctness criteria are no longer available.

This directly leads to the following research questions investigated by members of the Parsifal team:

- Finding (for classical logic) a notion of proof nets that is deductive, i.e., can effectively be used for doing proof search. An important property of deductive proof nets must be that the correctness can be checked in linear time. For the classical logic proof nets by Lamarche and Straßburger [6] this takes exponential time (in the size of the net).

- Studying the normalization of proofs in classical logic using atomic flows. Although there is no correctness criterion they allow to simplify the normalization procedure for proofs in deep inference, and additionally allow to get new insights in the complexity of the normalization.

3.6. Cost Models and Abstract Machines for Functional Programs

In the *proof normalization* approach, computation is usually reformulated as the evaluation of functional programs, expressed as terms in a variation over the λ-calculus. Thanks to its higher-order nature, this approach provides very concise and abstract specifications. Its strength is however also its weakness: the abstraction from physical machines is pushed to a level where it is no longer clear how to measure the complexity of an algorithm.

Models like Turing machines or RAM rely on atomic computational steps and thus admit quite obvious cost models for time and space. The λ-calculus instead relies on a single non-atomic operation, β-reduction, for which costs in terms of time and space are far from evident.

Nonetheless, it turns out that the number of β-steps is a reasonable time cost model, i.e., it is polynomially related to those of Turing machines and RAM. For the special case of *weak evaluation* (i.e., reducing only β-steps that are not under abstractions)—which is used to model functional programming languages—this is a relatively old result due to Blelloch and Greiner [46] (1995). It is only very recently (2014) that the strong case—used in the implementation models of proof assistants—has been solved by Accattoli and Dal Lago [42].

With the recent recruitment of Accattoli, the team's research has expanded in this direction. The topics under investigations are:

1. *Complexity of Abstract Machines*. Bounding and comparing the overhead of different abstract machines for different evaluation schemas (weak/strong call-by-name/value/need λ-calculi) with respect to the cost model. The aim is the development of a complexity-aware theory of the implementation of functional programs.

2. *Reasonable Space Cost Models*. Essentially nothing is known about reasonable space cost models. It is known, however, that environment-based execution model—which are the mainstream technology for functional programs—do not provide an answer. We are exploring the use of the non-standard implementation models provided by Girard's Geometry of Interaction to address this question.

4. Application Domains

4.1. Integrating a model checker and a theorem prover

The goal of combining model checking with inductive and co-inductive theorem is appealing. The strengths of systems in these two different approaches are strikingly different. A model checker is capable of exploring a finite space automatically: such a tool can repeatedly explore all possible cases of a given computational space. On the other hand, a theorem prover might be able to prove abstract properties about a search space. For example, a model checker could attempt to discover whether or not there exists a winning strategy for, say, tic-tac-toe while an inductive theorem prover might be able to prove that if there is a winning strategy for one board then there is a winning strategy for any symmetric version of that board. Of course, the ability to combine proofs from these systems could drastically reduce the amount of state exploration and verification of proof certificates that are needed to prove the existence of winning strategies.

Our first step to providing an integration of model checking and (inductive) theorem proving was the development of a strong logic, that we call $\mathcal{G}$, which extends intuitionistic logic with notions of least and greatest fixed points. We had developed the proof theory of this logic in earlier papers [4] [56]. We have now recently converted the Bedwyr system so that it formally accepts almost all definitions and theorem statements that are accepted by the inductive theorem prover Abella. Thus, these two systems are proving theorems in the same logic and their results can now be shared.

Bedwyr's tabling mechanism has been extended so that its it can make use of previously proved lemmas. For instance, when trying to prove that some board position has a winning strategy, an available stored lemma can now be used to obtain the result if some symmetric board position is already in the table.

Heath and Miller have shown how model checking can be seen as constructing proof in (linear) logic [64]. For more about recent progress on providing checkable proof certificates for model checking, see the web site for Bedwyr http://slimmer.gforge.inria.fr/bedwyr/.

4.2. Implementing trusted proof checkers

Traditionally, theorem provers—whether interactive or automatic—are usually monolithic: if any part of a formal development was to be done in a particular theorem prover, then the whole of it would need to be done in that prover. Increasingly, however, formal systems are being developed to integrate the results returned from several, independent and high-performance, specialized provers: see, for example, the integration of Isabelle with an SMT solver [55] as well as the Why3 and ESC/Java systems.

Within the Parsifal team, we have been working on foundational aspects of this multi-prover integration problem. As we have described above, we have been developing a formal framework for defining the semantics of proof evidence. We have also been working on prototype checkers of proof evidence which are capable of executing such formal definitions. The proof definition language described in the papers [52], [51] is currently given an implementation in the λProlog programming language [74]. This initial implementation will be able to serve as a "reference" proof checker: others who are developing proof evidence definitions will be able to use this reference checker to make sure that they are getting their definitions to do what they expect.

Using λProlog as an implementation language has both good and bad points. The good points are that it is rather simple to confirm that the checker is, in fact, sound. The language also supports a rich set of abstracts which make it impossible to interfere with the code of the checker (no injection attacks are possible). On the negative side, the performance of our λProlog interpreters is lower than that of specially written checkers and kernels.

4.3. Trustworthy implementations of theorem proving techniques

Instead of integrating different provers by exchanging proof evidence and relying on a backend proof-checker, another approach to integration consists in re-implementing the theorem proving techniques as proof-search strategies, on an architecture that guarantees correctness. Focused systems can serve as the basis of such an architecture, identifying points for choice and backtracking, and providing primitives for the exploration of the search space. These form a trusted *Application Programming Interface* that can be used to program and experiment various proof-search heuristics without worrying about correctness. No proof-checking is needed if one trusts the implementation of the API.

This approach has led to the development of the Psyche engine.

Two major research directions are currently being explored, based on the above:

- The first one is about understanding how to deal with quantifiers in presence of one or more theories: On the one hand, traditional techniques for quantified problems, such as *unification* [40] or *quantifier elimination* are usually designed for either the empty theory or very specific theories. On the other hand, the industrial techniques for combining theories (Nelson-Oppen, Shostak, MCSAT [79], [84], [89], [65]) are designed for quantifier-free problems, and quantifiers there are dealt with incomplete *clause instantiation* methods or *trigger*-based techniques [54]. We are working on making the two approaches compatible.

- The above architecture's modular approach raises the question of how its different modules can safely cooperate (in terms of guaranteed correctness), while some of them are trusted and others are not. The issue is particularly acute if some of the techniques are run concurrently and exchange data at unpredictable times. For this we explore new solutions based on Milner's *LCF* [77]. In [60], we argued that our solutions in particular provide a way to fulfil the "Strategy Challenge for SMT-solving" set by De Moura and Passmore [90].

5. Highlights of the Year

5.1. Highlights of the Year

D. Miller gave invited talks at the following two regularly held international meetings.
- TYPES 2016: 22nd International Conference on Types for Proofs and Programs (Novi Sad, Serbia, 23-26 May 2016) and
- Linearity 2016: 4th International Workshop on Linearity (Porto, 25 June 2016).

D. Miller gave invited talks at the following research oriented meetings.
- Workshop on linear logic, mathematics and computer science as part of "LL2016-Linear Logic: interaction, proofs and computation", 7-10 November 2016, Lyon. France.
- Research seminar titled "Interactions between logic, computer science and linguistics: history and philosophy", Université de Lille 3, 15 June 2016.
- CIPPMI (Current issues in the philosophy of practice of mathematics and informatics) Workshop on Proofs, justifications and certificates. 3-4 June 2016, Toulouse, France.

A seminar in honor of the 60th birthday of Professor Miller was held on 15-16 December at Université Paris Diderot-Paris 7 in Paris, France. Several members of the team contributed talks and original research papers.
- Tomer Libal and Marco Volpe, *A general proof certification framework for modal logic*.
- Roberto Blanco and Zakaria Chihani, *An interactive assistant for the definition of proof certificates*. Preprint available as [36].
- Lutz Straßburger, *Combinatorial flows as proof certificates with built-in proof compression*.
- Taus Brock-Nannestad, *Substructural cut elimination*.

B. Accattoli gave an invited talk at the following regularly held international meeting.
- WPTE 2016: 3rd International Workshop on Rewriting Techniques for Program Transformations and Evaluation (Porto, 23 June 2016).

S. Graham-Lengrand gave an invited talk at the following international conference.
- CLAM 2016: 5th Latin American Congress of Mathematicians, thematic session on Logic and Computability (Barranquilla, Colombia, 15th July 2016).

6. New Software and Platforms

6.1. Abella

FUNCTIONAL DESCRIPTION

Abella is an interactive theorem prover for reasoning about computations given as relational specifications. Abella is particularly well suited for reasoning about binding constructs.
- Participants: Dale Miller, Olivier Savary-Bélanger, Mary Southern, Yuting Wang, Kaustuv Chaudhuri, Matteo Cimini and Gopalan Nadathur
- Partner: Department of Computer Science and Engineering, University of Minnesota
- Contact: Kaustuv Chaudhuri
- URL: http://abella-prover.org/

6.2. Bedwyr

Bedwyr - A proof search approach to model checking
FUNCTIONAL DESCRIPTION

Bedwyr is a generalization of logic programming that allows model checking directly on syntactic expression that possibly contain bindings. This system, written in OCaml, is a direct implementation of two recent advances in the theory of proof search.

It is possible to capture both finite success and finite failure in a sequent calculus. Proof search in such a proof system can capture both may and must behavior in operational semantics. Higher-order abstract syntax is directly supported using term-level lambda-binders, the nabla quantifier, higher-order pattern unification, and explicit substitutions. These features allow reasoning directly on expressions containing bound variables.

The distributed system comes with several example applications, including the finite pi-calculus (operational semantics, bisimulation, trace analyses, and modal logics), the spi-calculus (operational semantics), value-passing CCS, the lambda-calculus, winning strategies for games, and various other model checking problems.

- Participants: Quentin Heath, Roberto Blanco, and Dale Miller
- Contact: Quentin Heath
- URL: http://slimmer.gforge.inria.fr/bedwyr/

6.3. Checkers

Checkers - A proof verifier
KEYWORDS: Proof - Certification - Verification
FUNCTIONAL DESCRIPTION

Checkers is a tool in Lambda-prolog for the certification of proofs. Checkers consists of a kernel which is based on LKF and is based on the notion of ProofCert.

- Participants: Tomer Libal, Giselle Machado Nogueira Reis and Marco Volpe
- Contact: Tomer Libal
- URL: https://github.com/proofcert/checkers

6.4. Psyche

Proof-Search factorY for Collaborative HEuristics
FUNCTIONAL DESCRIPTION

Psyche is a modular platform for automated or interactive theorem proving, programmed in OCaml and built on an architecture (similar to LCF) where a trusted kernel interacts with plugins. The kernel offers an API of proof-search primitives, and plugins are programmed on top of the API to implement search strategies. This architecture is set up for pure logical reasoning as well as for theory-specific reasoning, for various theories. The major effort in 2016 was the release of version 2.1 that allows the combination of theories, integrating and subsuming both the Nelson-Oppen methodology [79] and the *model constructing satisfiability* (MCSAT) methodology recently proposed by De Moura and Jovanovic [89], [65].

- Participants: Assia Mahboubi, Jean-Marc Notin and Stéphane Graham-Lengrand
- Contact: Stéphane Graham-Lengrand
- URL: http://www.lix.polytechnique.fr/~lengrand/Psyche/

7. New Results

7.1. Linear rewriting systems for Boolean logic

Participant: Lutz Straßburger.

Last year's result on the nonexistence of a complete linear term rewriting system for propositional logic [53] has been generalized and some applications to proof theory have been invesigated. For example, we have found that the medial rule which plays a central role in deep inference systems is canonical in a strong sense: It is minimal, and every rule that reduce contraction to an atomic form is indeed derivable via medial. This is published in [15] (joint work with Anupam Das).

7.2. Non-crossing Tree Realizations of Ordered Degree Sequences

Participant: Lutz Straßburger.

We investigate the enumeration of non-crossing tree realizations of integer sequences, and we consider a special case in four parameters, that can be seen as a four-dimensional tetrahedron that generalizes Pascal's triangle and the Catalan numbers. This work is motivated by the study of ambiguities arising in the parsing of natural language sentences using categorial grammars. This is joint work with Laurent Méhats and published in [31].

7.3. Focusing for Nested Sequents

Participants: Kaustuv Chaudhuri, Sonia Marin, Lutz Straßburger.

Focusing is a general technique for transforming a sequent proof system into one with a syntactic separation of non-deterministic choices without sacrificing completeness. This not only improves proof search, but also has the representational benefit of distilling sequent proofs into synthetic normal forms. We have shown how to apply the focusing technique to nested sequent calculi, a generalization of ordinary sequent calculi to tree-like instead of list-like structures. We thus improve the reach of focusing to the most commonly studied modal logics, the logics of the modal S5 cube. Among our key contributions is a focused cut-elimination theorem for focused nested sequents. This is published in [25].

Then we further extend our results to intuitionistic nested sequents, which can capture all the logics of the intuitionistic S5 cube in a modular fashion. We obtained an internal cut-elimination procedure for the focused system which in turn is used to show its completeness. This is published in [26]

7.4. Combining inference systems: a generalization of Nelson-Oppen and MCSAT

Participant: Stéphane Graham-Lengrand.

Nelson-Oppen [79] and Model-Constructing Satisfiability (MCSAT) [89], [65] are two methodologies that allow the reasoning mechanisms of different theories to collaborate, in order to tackle hybrid problems. While these methodologies are often used and implemented for the practical applications of Automated Reasoning, their rather sophisticated foundations are traditionally explained in terms of model theory. SRI International pioneered some work providing such methodologies with new and more general foundations in terms of *inference systems* [57], closer to proof theory and to Parsifal's research. The more recent MCSAT methodology was not captured, more generally lacked any kind of theorem about the generic combination of arbitrary theories, and was also thought to be incompatible with the Nelson-Oppen approach, so that SMT-solvers are either working with one methodology or the other, unable to get the best of both worlds.

In 2016 we designed a combination methodology, based on *inference systems*, that supersedes both Nelson-Oppen and MCSAT [34]. We showed its soundness and completeness, and identified for this the properties that the theories to combine are required to satisfy. This generalized MCSAT with the generic combination mechanism that it lacked, and showed that it is perfectly compatible with the Nelson-Oppen methodology, which can now cohabit within the same solver.

7.5. Linear lambda terms as invariants of rooted trivalent maps

Participant: Noam Zeilberger.

Recent studies of the combinatorics of linear lambda calculus have uncovered some unexpected connections to the old and well-developed theory of graphs embedded on surfaces (also known as "maps") [47], [87], [88]. In [19], we aimed to give a simple and conceptual account for one of these connections, namely the correspondence (originally described by Bodini, Gardy, and Jacquot [47]) between α-equivalence classes of closed linear lambda terms and isomorphism classes of rooted trivalent maps on compact oriented surfaces without boundary. One immediate application of this new account was a characterization of trivalent maps which are *bridgeless* (in the graph-theoretic sense of having no disconnecting edge) as linear lambda terms with no closed proper subterms. In turn, this lead to a surprising but natural reformulation of the Four Color Theorem as a statement about typing in lambda calculus.

7.6. A bifibrational reconstruction of Lawvere's presheaf hyperdoctrine

Participant: Noam Zeilberger.

In joint work with Paul-André Melliès, we have been investigating the categorical semantics of type refinement systems, which are type systems built "on top of" a typed programming language to specify and verify more precise properties of programs. The fibrational view of type refinement we have been developing (cf. [72]) is closely related to the categorical perspective on first-order logic introduced by Lawvere [66], but with some important conceptual and technical differences that provide an opportunity for reflection. For example, Lawvere's axiomatization of first-order logic (his theory of so-called "hyperdoctrines") was based on the idea that existential and universal quantification can be described respectively as left and right adjoints to the operation of substitution, this giving rise to a family of *adjoint triples* $\Sigma_f \dashv \mathcal{P}_f \dashv \Pi_f$ (one such triple for every function $f : A \to B$). On the other hand, a bifibration only induces a family of *adjoint pairs* $\mathsf{push}_f \dashv \mathsf{pull}_f$ (again, one such pair for every $f : A \to B$). In [33], we resolved this and other apparent mismatches by applying ideas inspired by the semantics of linear logic and the shift from the cartesian closed category **Set** to the symmetric monoidal closed category **Rel**. Two other applications of our analysis include an axiomatic treatment of *directed* equality predicates (which can be modelled as "hom" presheaves, realizing an early vision of Lawvere), as well as a simple calculus of string diagrams that is highly reminiscent of C. S. Peirce's "existential graphs" for predicate logic.

7.7. Towards a link between CPS and focusing

Participant: Matthias Puech.

Continuation-passing style translations make a functional program more explicit by sequentializing its computations and reifying its control. They have been used as an intermediate language in many compilers. They are also understood as classical-to-intuitionistic proof embedding (so-called double negation translations). Matthias Puech studied a novel correspondence between CPS and focusing: to each CPS transform corresponds a focused proof system that is identifiable as a particular polarization of classical statements. Since, after Miller's and others work, we know the full design space of focused sequent calculi, we expect to understand the full design space of CPS translation.

The first step of this goal is to study the syntax and typing of variants of the CPS translation. Puech designed and implemented in OCaml a compacting, optimizing CPS translation, while using OCaml's type system to verify that it maps well-typed terms to well-typed terms in a tightly restricted syntactical form (the "typeful" approach to formalization) [82]. The resulting type system is in Curry-Howard isomorphism with a weakly focused proof system: LJQ.

7.8. Proof Checking and Logic Programming

Participants: Roberto Blanco, Tomer Libal, Dale Miller, Marco Volpe.

In a world where trusting software systems is increasingly important, formal methods and formal proofs can help provide some basis for trust. Proof checking can help to reduce the size of the *trusted base* since we do not need to trust an entire theorem prover: instead, we only need to trust a (smaller and simpler) proof checker. Many approaches to building proof checkers require embedding within them a full programming language. In most modern proof checkers and theorem provers, that programming language is a functional programming language, often a variant of ML. In fact, aspects of ML (e.g., strong typing, abstract data types, and higher-order programming) were designed to make ML a trustworthy "meta-language" for checking proofs. While there is considerable overlap between logic programming and proof checking (e.g., both benefit from unification, backtracking search, efficient term structures, etc), the discipline of logic programming has, in fact, played a minor role in the history of proof checking. Miller has been pushing the argument that logic programming can have a major role in the future of this important topic [18]. Many aspects of the ProofCert project are based on this perspective that logic programming techniques and methods can have significant utility within proof checking. This perspective stands in contrast to the work on the Dedukti proof checking framework [44] where functional programming principles are employed for proof checking.

7.9. Proof Certificates for First-Order Equational Logic

Participants: Dale Miller, Zakaria Chihani.

The kinds of inference rules and decision procedures that one writes for proofs involving equality and rewriting are rather different from proofs that one might write in first-order logic using, say, sequent calculus or natural deduction. For example, equational logic proofs are often chains of replacements or applications of oriented rewriting and normal forms. In contrast, proofs involving logical connectives are trees of introduction and elimination rules. Chihani and Miller have shown [13] how it is possible to check various equality-based proof systems with a programmable proof checker (the *kernel* checker) for first-order logic. That proof checker's design is based on the implementation of *focused proof search* and on making calls to (user-supplied) *clerks and experts* predicates that are tied to the two phases found in focused proofs. This particular design is based on the work of Chihani, Miller, and Renaud [14].

The specification of these clerks and experts provide a formal definition of the structure of proof evidence and they work just as well in the equational setting as in the logic setting where this scheme for proof checking was originally developed. Additionally, executing such a formal definition on top of a kernel provides an actual proof checker that can also do a degree of proof reconstruction. A number of rewriting based proofs have been defined and checked in this manner.

7.10. Extended Pattern Unification

Participants: Tomer Libal, Dale Miller.

Unification is a central operation in the construction of a range of computational logic systems based on first-order and higher-order logics. First-order unification has a number of properties that dominates the way it is incorporated within such systems. In particular, first-order unification is decidable, unary, and can be performed on untyped term structures. None of these three properties hold for full higher-order unification: unification is undecidable, unifiers can be incomparable, and term-level typing can dominate the search for unifiers. The so-called *pattern* subset of higher-order unification was designed to be a small extension to first-order unification that respected the basic laws governing λ-binding (the equalities of α, β, and η-conversion) but which also satisfied those three properties. While the pattern fragment of higher-order unification has been popular in various implemented systems and in various theoretical considerations, it is too weak for a number of applications. Libal and Miller [28] have defined an extension of pattern unification that is motivated by some existing applications and which satisfies these three properties. The main idea behind their extension is that the arguments to a higher-order, free variable can be more than just distinct bound variables: they can also be terms constructed from (sufficient numbers of) such variables using term constructors and where no

argument is a subterm of any other argument. This extension to pattern unification satisfies the three properties mentioned above. R. Blanco is currently adding this extended unification to the Abella theorem prover.

7.11. Focused proofs for modal logics

Participants: Tomer Libal, Sonia Marin, Dale Miller, Marco Volpe.

Several deductive formalisms (e.g., sequent, nested sequent, labeled sequent, hypersequent calculi) have been used in the literature for the treatment of modal logics, and some connections between these formalisms are already known. Marin, Miller, and Volpe [30] have propose a general framework, which is based on a focused version of the labeled sequent calculus by Negri [78], augmented with some parametric devices allowing to restrict the set of proofs. By properly defining such restrictions and by choosing an appropriate polarization of formulas, one can obtain different, concrete proof systems for the modal logic K and for its extensions by means of geometric axioms. The expressiveness of the labeled approach and the control mechanisms of focusing allow a clean emulation of a range of existing formalisms and proof systems for modal logic. These results make it possible to write Foundational Proof Certificate definitions of common modal logic proof systems.

7.12. Preserving differential privacy under finite-precision semantics

Participant: Dale Miller.

(Joint work with Ivan Gazeau and Catuscia Palamidessi). 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. Gazeau, Miller, and Palamidessi [16] have applied differential privacy—an approach to privacy that emerged from the area of statistical databases—to this problem. In their approach, privacy is protected by the addition of noise to a true (private) value. To date, this approach to privacy has been proved correct only in the ideal case in which computations are made using an idealized, infinite-precision semantics. An analysis of implementation levels, 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. In general there are violations of the differential privacy property but a limited (but, arguably, totally acceptable) variant of the property can be used insteand, under only a minor degradation of the privacy level. In fact, two cases of noise-generating distributions can be employed: 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.

7.13. Certification of Prefixed Tableau Proofs for Modal Logic

Participants: Tomer Libal, Marco Volpe.

This work [29] describes the theory and implementation of a proof checker for tableau theorem provers for modal logics. The tool supports proofs in both the traditional tableau format as well as the free variable variant. The implentation can be found at https://github.com/proofcert/checkers under the gandalf2016 branch.

7.14. Towards a Substitution Tree Based Index for Higher-order Resolution Theorem Provers

Participant: Tomer Libal.

First-order resolution theorem provers depend on efficient data structures for redundancy elimination. These data structures do not exist for higher-order resolution theorem provers. In [32] we discuss a new approach to this problem. (Joint work with Alexander Steen).

7.15. Open Call-by-Value

Participant: Beniamino Accattoli.

Functional programming languages are often based on the call-by-value λ-calculus, whose elegant theory relies on weak evaluation and closed terms, that are natural hypotheses in the study of programming languages. To model proof assistants, however, strong evaluation and open terms are required, and it is well known that the operational semantics of call-by-value becomes problematic in this case. In this joint work with Giulio Guerrieri we studied the intermediate setting—that we call Open Call-by-Value—of weak evaluation with open terms, on top of which Gregoire and Leroy designed the abstract machine of Coq. Various calculi for Open Call-by-Value already exist, each one with its pros and cons. We did a detailed comparative study of the operational semantics of four of them, coming from different areas such as the study of abstract machines, denotational semantics, linear logic proof nets, and sequent calculus. We showed that these calculi are all equivalent from a termination point of view, justifying the slogan Open Call-by-Value. The work has been published in the proceedings of the international conference APLAS 2016 [22].

7.16. A Reasonable Abstract Machine for the Strong λ-Calculus

Participant: Beniamino Accattoli.

We provided a new proof that the strong λ-calculus is a reasonable computational model. The original proof is by B. Accattoli and H. Dal Lago uses a calculus with explicit substitutions while the new one relies on a new sophisticated abstract machine, the Useful MAM. The work has been published in the proceeding of the international conference WoLLIC 2016 [21].

7.17. Space-efficient Acyclicity Constraints

Participant: Taus Brock-Nannestad.

Acyclicity constraints can be used to encode a large variety of useful constraints on graphs. The basic constraint itself can be encoded in terms of simpler constraints (e.g. integer linear constraints) in a straightforward and intuitive way, associating to each vertex of the (fixed) input graph a variable with domain linear in the size of the graph. For large graphs, this quickly becomes inefficient.

In [24], we show that in the case of planar graphs, a more efficient encoding (using a two-valued variable per vertex) is possible.

7.18. Exp-log normal form of types and the axioms for η-equality of the λ-calculus with sums

Participant: Danko Ilik.

In the presence of sum types, the λ-calculus has but one implemented (and incomplete) heuristic for deciding $\beta\eta$-equality of terms, in spite of a dozen of meta-theoretic works showing that the equality is decidable.

In the work discussed here, we first used the exp-log decomposition of the arrow type—inspired from the analytic transformation $a^b = \exp(b \times \log a)$—to obtain a type normal form for the type languages $\{\to, \times, +\}$. We then made a quotient of the $\beta\eta$-equality of terms modulo the terms coerced into their representation at the exp-log normal form of their type. This allows to obtain a *simplification* of the so far standard axioms for $\beta\eta$-equality.

Moreover, we provided a Coq implementation of a heuristic decision procedure for this equality. Although a heuristic, this implementation manages to tackle examples of equal terms that need a complex program analysis in the only previously implemented heuristic of Vincent Balat.

This work is described in a paper accepted for presentation at POPL 2017, [27].

7.19. Invertible-rule-free sequent calculi and an intuitionistic arithmetical hierarchy

Participants: Taus Brock-Nannestad, Danko Ilik.

In sequent calculi, proof rules can be divided into two groups: invertible (asynchronous) proof rules and non-invertible (synchronous) proof rules. Even in focusing sequent calculi the two groups of rules are present, albeit grouped together in synthetic rules (we speak of the synchronous and asynchronous phase).

In this work, we used the exp-log decomposition (described above) in the context of logic in order to obtain a version of sequent caclulus which contains synchronous rules only, a first such formalism for intuitionistic logic.

We extended the picture from the setting of propositional to the one of first-order intuitionistic logic, where the exp-log decomposition provided us with an intuitionistic hierarchy of formulas analogous to the classical arithmetical hierarchy; although the classical arithmetical hierarchy exists since the 1920s, a correspondingly versatile notion for intuitionistic logic has been elusive up to this day.

This work is described in the manuscript [37], submitted to an academic journal.

8. Partnerships and Cooperations

8.1. European Initiatives

8.1.1. FP7 & H2020 Projects

8.1.1.1. Proofcert

> Title: ProofCert: Broad Spectrum Proof Certificates
>
> Programm: FP7
>
> Type: ERC
>
> Duration: January 2012 - December 2016
>
> Coordinator: Inria
>
> Inria contact: Dale Miller
>
> There is little hope that the world will know secure software if we cannot make greater strides in the practice of formal methods: hardware and software devices with errors are routinely turned against their users. The ProofCert proposal aims at building a foundation that will allow a broad spectrum of formal methods—ranging from automatic model checkers to interactive theorem provers—to work together to establish formal properties of computer systems. This project starts with a wonderful gift to us from decades of work by logicians and proof theorist: their efforts on logic and proof has given us a universally accepted means of communicating proofs between people and computer systems. Logic can be used to state desirable security and correctness properties of software and hardware systems and proofs are uncontroversial evidence that statements are, in fact, true. The current state-of-the-art of formal methods used in academics and industry shows, however, that the notion of logic and proof is severely fractured: there is little or no communication between any two such systems. Thus any efforts on computer system correctness is needlessly repeated many time in the many different systems: sometimes this work is even redone when a given prover is upgraded. In ProofCert, we will build on the bedrock of decades of research into logic and proof theory the notion of proof certificates. Such certificates will allow for a complete reshaping of the way that formal methods are employed. Given the infrastructure and tools envisioned in this proposal, the world of formal methods will become as dynamic and responsive as the world of computer viruses and hackers has become.

8.1.2. Collaborations in European Programs, Except FP7 & H2020

8.1.2.1. FISP: ANR blanc International

Participants: Kaustuv Chaudhuri, François Lamarche, Sonia Marin, Dale Miller, Lutz Straßburger.

Title: The Fine Structure of Formal Proof Systems and their Computational Interpretations

Duration: 01/01/2016 – 31/12/2018

Partners:

University Paris VII, PPS (PI: Michel Parigot)

Inria Saclay–IdF, EPI Parsifal (PI: Lutz Straßburger)

University of Innsbruck, Computational Logic Group (PI: Georg Moser)

Vienna University of Technology, Theory and Logic Group (PI: Matthias Baaz)

Total funding by the ANR: 316 805 EUR

The FISP project is part of a long-term, ambitious project whose objective is to apply the powerful and promising techniques from structural proof theory to central problems in computer science for which they have not been used before, especially the understanding of the computational content of proofs, the extraction of programs from proofs and the logical control of refined computational operations. So far, the work done in the area of computational interpretations of logical systems is mainly based on the seminal work of Gentzen, who in the mid-thirties introduced the sequent calculus and natural deduction, along with the cut-elimination procedure. But that approach shows its limits when it comes to computational interpretations of classical logic or the modelling of parallel computing. The aim of our project, based on the complementary skills of the teams, is to overcome these limits. For instance, deep inference provides new properties, namely full symmetry and atomicity, which were not available until recently and opened new possibilities at the computing level, in the era of parallel and distributed computing.

8.1.2.2. COCA HOLA: ANR JCJC Project

Participant: Beniamino Accattoli.

Title: COst model for Complexity Analyses of Higher-Order programming LAnguages.

Collaborators: Ugo Dal Lago (University of Bologna & Inria), Delia Kesner (Paris Diderot University), Damiano Mazza (CNRS & Paris 13 University), Claudio Sacerdoti Coen (University of Bologna).

Duration: 01/10/2016 – 31/09/2019

Total funding by the ANR: 155 280 EUR

The COCA HOLA project aims at developing complexity analyses of higher-order computations, i.e. that approach to computation where the inputs and outputs of a program are not simply numbers, strings, or compound data-types, but programs themselves. The focus is not on analysing fixed programs, but whole programming languages. The aim is the identification of adequate units of measurement for time and space, i.e. what are called reasonable cost models. The problem is non-trivial because the evaluation of higher-order languages is defined abstractly, via high-level operations, leaving the implementation unspecified. Concretely, the project will analyse different implementation schemes, measuring precisely their computational complexity with respect to the number of high-level operations, and eventually develop more efficient new ones. The goal is to obtain a complexity-aware theory of implementations of higher-order languages with both theoretical and practical downfalls.

The projects stems from recent advances on the theory of time cost models for the lambda-calculus, the computational model behind the higher-order approach, obtained by the principal investigator and his collaborators (who are included in the project).

COCA HOLA will span over three years and is organised around three work packages, essentially:

1. extending the current results to encompass realistic languages;
2. explore the gap between positive and negative results in the literature;
3. use ideas from linear logic to explore space cost models, about which almost nothing is known.

8.2. International Initiatives

8.2.1. Participation in Other International Programs

8.2.1.1. PHC Amadeus: Analytic Calculi for Modal Logics
Participants: Kaustuv Chaudhuri, Sonia Marin, Giselle Reis, Lutz Straßburger.

Title: Analytic Calculi for Modal Logics

Duration: 01/01/2016 – 31/12/2017

Austrian Partner: TU Wien, Institute for Computer Science (Department III)

Modal logics are obtained from propositional logics by adding modalities $\Box$ and $\diamond$, meaning necessity and possibility. Originally studied by philosophers in order to reason about knowledge and belief, modal logics have nowadays many applications in computer science. Well known examples are epistemic logics, which allow to formally reason about the knowledge of independently acting and interacting agents, temporal logics, which allow to reason about temporal properties of processes, and authentication logics, which are used to formally reason about authentication protocols.

The purpose of this project is to develop a proof theory for variants of modal logic that have applications in modern computer science but that have been neglected by traditional proof theory so far.

8.3. International Research Visitors

8.3.1. Visits of International Scientists

Professor Chuck Liang (from Hofstra University, NY, USA) visited the team from 5 June to 25 June 2016 in order to continue his collaborations with team members on basic questions of proof theory. In particular, he worked with Miller on identifying possible means to allow classical and intuitionistic logic to be mixed in a common proof system. Miller is exploring how the resulting ideas might be able to reorganize the notion of kernel logic used within the ProofCert project.

8.3.1.1. Internships

Ameni Chtourou was an intern funded by ProofCert during May, June, and July 2016. She was advised by Accattoli and worked with using the Abella theorem prover to formalize connections various connections between λ-term evaluation and abstract machine models.

8.3.2. Visits to International Teams

8.3.2.1. Research Stays Abroad

Stéphane Graham-Lengrand spent 8 months, from January 2016 to August 2016, at SRI International, Computer Science Lab. This visit developed a collaboration with N. Shankar, MP Bonacina, D. Jovanovic, and Martin Schaeff on new algorithms and new architectures for automated and interactive theorem proving, as well as on new programme verification techniques.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the Organizing Committees

D. Miller was on the Steering Committee for the FSCD series of International Conference on Formal Structures for Computation and Deduction.

D. Miller was a member of the jury for selecting the 2016 Ackermann Award (the EACSL award for outstanding doctoral dissertation in the field of Logic in Computer Science).

D. Miller was an ex officio member of the Executive Committee of the ACM Special Interest Group on Logic and Compuation (SIGLOG), from April 2014 to June 2016. He was also a member of the SIGLOG advisory board, starting November 2015.

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

D. Miller was on the Program Committee of the following meetings.

FSCD'16: First International Conference on Formal Structures for Computation and Deduction, Porto, Portugal, 22-26 June.

IJCAR 2016: International Joint Conference on Automated Reasoning, Coimbra, Portugal, 27 June - 2 July.

CPP 2016, Fifth International Conference on Certified Programs and Proofs, 18-19 January, Saint Petersburg, Florida.

B. Accattoli was one of the two Program Committee chairs of the 5th International Workshop on Confluence (IWC 2016).

N. Zeilberger served on the program committee for workshops Computational Logic and Applications (CLA 2016) and Off the Beaten Track (OBT 2016).

N. Zeilberger served on external review committee for POPL 2017

L. Straßburger was on the Program Committee for LICS 2016.

9.1.2.2. Reviewer

D. Miller was a reviewer for CONCUR 2016: the International Conference on Concurrency Theory.

B. Accattoli was a reviewer for the international conferences ICTAC 2016, FSCD 2016, LICS 2016 (twice), FOSSACS 2017.

S. Graham-Lengrand was a reviewer for the international conferences FSCD 2016 (twice), LICS 2016 (three times), Concur 2016, VSTTE 2017, HATT 2017, FSTTCS 2017.

L. Straßburger was a reviewer for the international conferences LICS 2016 (8 times), FLOPS 2016.

M. Volpe was a reviewer for the international conferences IJCAR 2016 and CSL 2016.

H. Steele was a reviewer for the internal conference LICS 2016.

9.1.3. Journal

9.1.3.1. Member of the editorial boards

D. Miller is on the editorial board of the following journals: *ACM Transactions on Computational Logic*, *Journal of Automated Reasoning* (Springer), *Theory and Practice of Logic Programming* (Cambridge University Press), and *Journal of Applied Logic* (Elsevier).

9.1.3.2. Reviewer - Reviewing Activities

S. Graham-Lengrand has been a reviewer for the journals *Fundamenta Informaticae*, *Transactions on Computational Logic*, *Journal of Logic and Computation*, *Logical Methods in Computer Science*, *Journal of Automated Reasoning*.

Danko Ilik was a reviewer for Mathematical Reviews and Zentralblatt MATH.

F. Lamarche was a reviewer for *Mathematical Structures in Computer Science*.

Lutz Straßburger was a reviewer for the journals *Theoretical Computer Science* and *Logical Methods in Computer Science*.

Marco Volpe was a reviewer for the journal *Annals of Mathematics and Artificial Intelligence*.

Beniamino Accattoli was a reviewer for the journals *Theoretical Computer Science* and *Logical Methods in Computer Science*.

9.1.4. Invited Talks

D. Miller was an invited speaker at the following conferences and workshops.

Workshop on linear logic, mathematics and computer science as part of "LL2016-Linear Logic: interaction, proofs and computation", 7-10 November 2016, Lyon. France.

Linearity 2016. Porto, 25 June 2016.

CIPPMI (Current issues in the philosophy of practice of mathematics and informatics) Workshop on Proofs, justifications and certificates. 3-4 June 2016, Toulouse, France.

TYPES 2016: 22nd International Conference on Types for Proofs and Programs. Novi Sad, Serbia, 23-26 May 2016.

D. Miller was an invited speaker at the research seminar titled "Interactions between logic, computer science and linguistics: history and philosophy", Université de Lille 3, 15 June 2016.

D. Miller was an invited speaker at the ACADIA research centre, Ca' Foscari University, Venice, 27 April 2016.

B. Accattoli was invited speaker at WPTE 2016: 3rd International Workshop on Rewriting Techniques for Program Transformations and Evaluation (Porto, 23 June 2016).

S. Graham-Lengrand gave an invited talk at CLAM 2016: 5th Latin American Congress of Mathematicians, thematic session on Logic and Computability (Barranquilla, Colombia, 15th July 2016).

N. Zeilberger was an invited lecturer at OPLSS 2016: Oregon Programming Languages Summer School on Types, Logic, Semantics, and Verification.

9.1.5. Leadership within the Scientific Community

D. Miller was a member of the ACM SIGLOG Advisory Board, the LICS Organizing Board, the CPP Steering Committee, and the ACM SIGLOG Executive Committee Nominating Committee.

S. Graham-Lengrand is the head of the National Workgroup on "Logic, Algebra, and Computation", within the Informatique Mathématique section of CNRS.

9.1.6. Research Administration

L. Straßburger serves on the "commission développement technologique (CDT)" for Inria Saclay–Île-de-France since June 2012

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master: D. Miller, "*MPRI 2-1: Logique linéaire et paradigmes logiques du calcul*", 12 hours, M2, Master Parisien de Recherche en Informatique, France.

Licence: S. Graham-Lengrand, "*INF412: Fondements de l'Informatique: Logique, Modèles, Calcul*", 32 hours eq. TD, L3, École Polytechnique, France.

Master: S. Graham-Lengrand, "*INF551: Computational Logic*", 45 hours eq. TD, M1, École Polytechnique, France.

Master: S. Graham-Lengrand, "*MPRI 2-1: Logique linéaire et paradigmes logiques du calcul*", 6 hours, M2, Master Parisien de Recherche en Informatique, France.

Undergraduate: K. Chaudhuri, R. Blanco, M. Volpe, G. Reis, T. Libal all taught or tutored exercises for first and second year undergrad courses, mostly at École Polytechnique.

9.2.2. Supervision

PhD in progress: Sonia Marin, 1 Nov 2014, supervised by L. Straßburger and D. Miller

PhD in progress: Roberto Blanco, Ulysse Gérard, and Quentin Heath, supervised by D. Miller

PhD in progress: François Thiré (since 1st October 2016), supervised by S. Graham-Lengrand (joint with G. Dowek)

9.2.3. Juries

Miller was a reporter for the PhD juries of Raphaël Cauderlier (CNAM, 10 October 2016) and Gabriel Scherer (Université Paris-Diderot, 30 March 2016).

Graham-Lengrand was a reporter for the PhD juries of Pierre Halmagrand (CNAM, 10 December 2016).

10. Bibliography

Major publications by the team in recent years

[1] D. BAELDE, K. CHAUDHURI, A. GACEK, D. MILLER, G. NADATHUR, A. TIU, Y. WANG.*Abella: A System for Reasoning about Relational Specifications*, in "Journal of Formalized Reasoning", 2014, vol. 7, n° 2, p. 1-89 [*DOI :* 10.6092/ISSN.1972-5787/4650], https://hal.inria.fr/hal-01102709.

[2] K. CHAUDHURI, N. GUENOT, L. STRASSBURGER.*The Focused Calculus of Structures*, in "Computer Science Logic: 20th Annual Conference of the EACSL", Leibniz International Proceedings in Informatics (LIPIcs), Schloss Dagstuhl–Leibniz-Zentrum für Informatik, September 2011, p. 159–173, http://drops.dagstuhl.de/opus/frontdoor.php?source_opus=3229.

[3] M. FAROOQUE, S. GRAHAM-LENGRAND, A. MAHBOUBI.*A bisimulation between DPLL(T) and a proof-search strategy for the focused sequent calculus*, in "Proceedings of the 2013 International Workshop on Logical Frameworks and Meta-Languages: Theory and Practice (LFMTP 2013)", A. MOMIGLIANO, B. PIENTKA, R. POLLACK (editors), ACM Press, September 2013 [*DOI :* 10.1145/2503887.2503892].

[4] A. GACEK, D. MILLER, G. NADATHUR.*Nominal abstraction*, in "Information and Computation", 2011, vol. 209, n° 1, p. 48–73, http://arxiv.org/abs/0908.1390.

[5] A. GUGLIELMI, T. GUNDERSEN, L. STRASSBURGER.*Breaking Paths in Atomic Flows for Classical Logic*, in "Proceedings of the 25th Annual IEEE Symposium on Logic in Computer Science (LICS 2010)", Edinburgh, United Kingdom, July 2010, p. 284–293 [*DOI :* 10.1109/LICS.2010.12], http://www.lix.polytechnique.fr/~lutz/papers/AFII.pdf.

[6] F. LAMARCHE, L. STRASSBURGER.*Naming Proofs in Classical Propositional Logic*, in "Typed Lambda Calculi and Applications, TLCA 2005", P. URZYCZYN (editor), LNCS, Springer-Verlag, 2005, vol. 3461, p. 246–261.

[7] C. LIANG, D. MILLER.*Focusing and Polarization in Linear, Intuitionistic, and Classical Logics*, in "Theoretical Computer Science", 2009, vol. 410, n° 46, p. 4747–4768.

[8] C. LIANG, D. MILLER.*A Focused Approach to Combining Logics*, in "Annals of Pure and Appl. Logic", 2011, vol. 162, n° 9, p. 679–697 [*DOI :* 10.1016/J.APAL.2011.01.012], http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/lku.pdf.

[9] D. MILLER.*A proposal for broad spectrum proof certificates*, in "CPP: First International Conference on Certified Programs and Proofs", J.-P. JOUANNAUD, Z. SHAO (editors), LNCS, 2011, vol. 7086, p. 54–69, http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/cpp11.pdf.

[10] L. STRASSBURGER.*On the Axiomatisation of Boolean Categories with and without Medial*, in "Theory and Applications of Categories", 2007, vol. 18, nº 18, p. 536–601, http://arxiv.org/abs/cs.LO/0512086.

[11] A. TIU, D. MILLER.*Proof Search Specifications of Bisimulation and Modal Logics for the π-calculus*, in "ACM Trans. on Computational Logic", 2010, vol. 11, nº 2, http://arxiv.org/abs/0805.2785.

Publications of the year

Articles in International Peer-Reviewed Journal

[12] B. ACCATTOLI, U. DAL LAGO.*(Leftmost-outermost) beta reduction is invariant, indeed*, in "Logical Methods in Computer Science", 2016 [*DOI :* 10.2168/LMCS-12(1:4)2016], https://hal.inria.fr/hal-01337712.

[13] Z. CHIHANI, D. MILLER.*Proof Certificates for Equality Reasoning*, in "Electronic Notes in Theoretical Computer Science", 2016, vol. 323, p. 93 - 108 [*DOI :* 10.1016/J.ENTCS.2016.06.007], https://hal.inria.fr/hal-01390919.

[14] Z. CHIHANI, D. MILLER, F. RENAUD.*A Semantic Framework for Proof Evidence*, in "Journal of Automated Reasoning", 2016 [*DOI :* 10.1007/S10817-016-9380-6], https://hal.inria.fr/hal-01390912.

[15] A. DAS, L. STRASSBURGER.*On linear rewriting systems for Boolean logic and some applications to proof theory*, in "Logical Methods in Computer Science", December 2016, https://hal.inria.fr/hal-01422611.

[16] I. GAZEAU, D. MILLER, C. PALAMIDESSI.*Preserving differential privacy under finite-precision semantics*, in "Journal of Theoretical Computer Science (TCS)", 2016, https://hal.inria.fr/hal-01390927.

[17] W. HEIJLTJES, L. STRASSBURGER.*Proof nets and semi-star-autonomous categories*, in "Mathematical Structures in Computer Science", June 2016, vol. 26, nº 5, p. 789-828 [*DOI :* 10.1017/S0960129514000395], https://hal.inria.fr/hal-01417643.

[18] D. MILLER.*Proof Checking and Logic Programming*, in "Formal Aspects of Computing", 2016 [*DOI :* 10.1007/S00165-016-0393-z], https://hal.inria.fr/hal-01390901.

[19] N. ZEILBERGER.*Linear lambda terms as invariants of rooted trivalent maps*, in "Journal of Functional Programming", November 2016, vol. 26 [*DOI :* 10.1017/S095679681600023X], https://hal.archives-ouvertes.fr/hal-01247757.

Invited Conferences

[20] B. ACCATTOLI.*The Complexity of Abstract Machines*, in "Third International Workshop on Rewriting Techniques for Program Transformations and Evaluation (WPTE 2016)", Porto, Portugal, June 2016, vol. 235, p. 1 - 15 [*DOI :* 10.4204/EPTCS.235.1], https://hal.inria.fr/hal-01425560.

International Conferences with Proceedings

[21] B. ACCATTOLI.*The Useful MAM, a Reasonable Implementation of the Strong λ-Calculus*, in "23rd International Workshop on Logic, Language, Information, and Computation (WoLLIC 2016)", Puebla, Mexico, August 2016, p. 1 - 21 [*DOI :* 10.1007/978-3-662-52921-8_1], https://hal.inria.fr/hal-01425534.

[22] B. ACCATTOLI, G. GUERRIERI.*Open Call-by-Value*, in "14th Asian Symposium on Programming Languages and Systems (APLAS)", Hanoi, Vietnam, November 2016, p. 206 - 226 [*DOI :* 10.1007/978-3-319-47958-3_12], https://hal.inria.fr/hal-01425465.

[23] S. AZAIEZ, D. DOLIGEZ, M. LEMERRE, T. LIBAL, S. MERZ.*Proving Determinacy of the PharOS Real-Time Operating System*, in "Abstract State Machines, Alloy, B, TLA, VDM, and Z - 5th International Conference, ABZ 2016", Linz, Austria, M. J. BUTLER, K.-D. SCHEWE, A. MASHKOOR, M. BIRÓ (editors), LNCS - Lecture Notes in Computer Science, Springer, May 2016, vol. 9675, p. 70-85 [*DOI :* 10.1007/978-3-319-33600-8_4], https://hal.inria.fr/hal-01322335.

[24] T. BROCK-NANNESTAD.*Space-efficient Planar Acyclicity Constraints - A Declarative Pearl*, in "FLOPS 2016 - 13th International Symposium on Functional and Logic Programming", Kochi, Japan, March 2016, https://hal.inria.fr/hal-01426753.

[25] K. CHAUDHURI, S. MARIN, L. STRASSBURGER.*Focused and Synthetic Nested Sequents*, in "FOSSACS 2016 - 19th International Conference Foundations of Software Science and Computation Structures", Eindhoven, Netherlands, B. JACOBS, C. LÖDING (editors), LNCS - Lecture Notes in Computer Science, Springer, April 2016, vol. 9634, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2016 [*DOI :* 10.1007/978-3-662-49630-5_23], https://hal.inria.fr/hal-01417618.

[26] K. CHAUDHURI, S. MARIN, L. STRASSBURGER.*Modular Focused Proof Systems for Intuitionistic Modal Logics*, in "FSCD 2016 - 1st International Conference on Formal Structures for Computation and Deduction", Porto, Portugal, June 2016 [*DOI :* 10.4230/LIPIcs.FSCD.2016.16], https://hal.inria.fr/hal-01417603.

[27] D. ILIK.*On the exp-log normal form of types: Decomposing extensional equality and representing terms compactly*, in "Proceedings of the 44th ACM SIGPLAN Symposium on Principles of Programming Languages", Paris, France, January 2017, p. 387-399, https://hal.inria.fr/hal-01167162.

[28] T. LIBAL, D. MILLER.*Functions-as-constructors Higher-order Unification*, in "1st International Conference on Formal Structures for Computation and Deduction (FSCD 2016)", Porto, Portugal, Delia Kesner and Brigitte Pientka, June 2016, p. 1 - 17 [*DOI :* 10.4230/LIPIcs.FSCD.2016.26], https://hal.inria.fr/hal-01379683.

[29] T. LIBAL, M. VOLPE.*Certification of Prefixed Tableau Proofs for Modal Logic*, in "the Seventh International Symposium on Games, Automata, Logics and Formal Verification (GandALF 2016)", Catania, Italy, September 2016, vol. 226, p. 257 - 271 [*DOI :* 10.4204/EPTCS.226.18], https://hal.archives-ouvertes.fr/hal-01379625.

[30] S. MARIN, D. MILLER, M. VOLPE.*A focused framework for emulating modal proof systems*, in "11th conference on "Advances in Modal Logic"", Budapest, Hungary, L. BEKLEMISHEV, S. DEMRI, A. MÁTÉ (editors), Advances in Modal Logic, College Publications, August 2016, n⁰ 11, p. 469-488, https://hal.archives-ouvertes.fr/hal-01379624.

[31] L. MEHATS, L. STRASSBURGER.*Non-crossing Tree Realizations of Ordered Degree Sequences*, in "LACL 2016 - 9th International Conference Logical Aspects of Computational Linguistics - Celebrating 20 Years of LACL (1996–2016)", Nancy, France, M. AMBLARD, P. DE GROOTE, S. POGODALLA, C. RETORÉ (editors), LNCS - Lecture Notes in Computer Science, Springer, December 2016, vol. 10054 [*DOI :* 10.1007/978-3-662-53826-5_13], https://hal.inria.fr/hal-01417567.

Conferences without Proceedings

[32] T. LIBAL, A. STEEN.*Towards a Substitution Tree Based Index for Higher-order Resolution Theorem Provers*, in "5th Workshop on Practical Aspects of Automated Reasoning", Coimbra, Portugal, Proceedings of the 5th Workshop on Practical Aspects of Automated Reasoning co-located with International Joint Conference on Automated Reasoning (IJCAR 2016), Coimbra, Portugal, July 2nd, 2016, July 2016, https://hal.archives-ouvertes.fr/hal-01424749.

[33] P.-A. MELLIÈS, N. ZEILBERGER.*A bifibrational reconstruction of Lawvere's presheaf hyperdoctrine*, in "IEEE/ACM Logic in Computer Science (LICS) 2016", New York, United States, Proceedings of LICS'16, June 2016, https://hal.archives-ouvertes.fr/hal-01261955.

Research Reports

[34] M. P. BONACINA, S. GRAHAM-LENGRAND, N. SHANKAR.*A model-constructing framework for theory combination*, Universita degli Studi di Verona, September 2016, n$^{\text{o}}$ RR-99/2016, https://hal.inria.fr/hal-01425305.

[35] K. CHAUDHURI, S. MARIN, L. STRASSBURGER.*Focused and Synthetic Nested Sequents (Extended Technical Report)*, Inria, April 2016, https://hal.inria.fr/hal-01251722.

Other Publications

[36] R. BLANCO, Z. CHIHANI.*An interactive assistant for the definition of proof certificates*, December 2016, working paper or preprint, https://hal.inria.fr/hal-01422829.

[37] T. BROCK-NANNESTAD, D. ILIK.*An Intuitionistic Formula Hierarchy Based on High-School Identities*, August 2016, working paper or preprint, https://hal.inria.fr/hal-01354181.

[38] Z. CHIHANI, D. ILIK, D. MILLER.*Classical polarizations yield double-negation translations*, August 2016, working paper or preprint, https://hal.inria.fr/hal-01354298.

[39] D. ILIK.*Perspectives for proof unwinding by programming languages techniques*, August 2016, Book chapter to appear, https://hal.inria.fr/hal-01354180.

References in notes

[40] J. A. ROBINSON, A. VORONKOV (editors). *Handbook of Automated Reasoning*, Elsevier and MIT press, 2001.

[41] S. ABRAMSKY.*Computational Interpretations of Linear Logic*, in "Theoretical Computer Science", 1993, vol. 111, p. 3–57.

[42] B. ACCATTOLI, U. DAL LAGO.*Beta reduction is invariant, indeed*, in "Joint Meeting of the Twenty-Third EACSL Annual Conference on Computer Science Logic (CSL) and the Twenty-Ninth Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), CSL-LICS '14, Vienna, Austria, July 14 - 18, 2014", 2014, p. 8:1–8:10, http://doi.acm.org/10.1145/2603088.2603105.

[43] J.-M. ANDREOLI.*Logic Programming with Focusing Proofs in Linear Logic*, in "Journal of Logic and Computation", 1992, vol. 2, n⁰ 3, p. 297–347.

[44] A. ASSAF, G. BUREL, R. CAUDERLIER, D. DELAHAYE, G. DOWEK, C. DUBOIS, F. GILBERT, P. HALMAGRAND, O. HERMANT, R. SAILLARD.*Dedukti: a Logical Framework based on the λΠ-Calculus Modulo Theory*, 2016, Unpublished, http://www.lsv.ens-cachan.fr/~dowek/Publi/expressing.pdf.

[45] D. BAELDE, D. MILLER, Z. SNOW.*Focused Inductive Theorem Proving*, in "Fifth International Joint Conference on Automated Reasoning (IJCAR 2010)", J. GIESL, R. HÄHNLE (editors), LNCS, 2010, n⁰ 6173, p. 278–292 [*DOI :* 10.1007/978-3-642-14203-1], http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/ijcar10.pdf.

[46] G. E. BLELLOCH, J. GREINER.*Parallelism in Sequential Functional Languages*, in "Proceedings of the seventh international conference on Functional programming languages and computer architecture, FPCA 1995, La Jolla, California, USA, June 25-28, 1995", 1995, p. 226–237, http://doi.acm.org/10.1145/224164.224210.

[47] O. BODINI, D. GARDY, A. JACQUOT.*Asymptotics and random sampling for BCI and BCK lambda terms*, in "Theoretical Computer Science", 2013, vol. 502, p. 227–238.

[48] K. CHAUDHURI.*The Focused Inverse Method for Linear Logic*, Carnegie Mellon University, December 2006, Technical report CMU-CS-06-162, http://reports-archive.adm.cs.cmu.edu/anon/2006/CMU-CS-06-162.pdf.

[49] K. CHAUDHURI, N. GUENOT, L. STRASSBURGER.*The Focused Calculus of Structures*, in "Computer Science Logic: 20th Annual Conference of the EACSL", Leibniz International Proceedings in Informatics (LIPIcs), Schloss Dagstuhl–Leibniz-Zentrum für Informatik, September 2011, p. 159–173 [*DOI :* 10.4230/LIPIcs.CSL.2011.159], http://drops.dagstuhl.de/opus/volltexte/2011/3229/pdf/16.pdf.

[50] K. CHAUDHURI, S. HETZL, D. MILLER.*A Multi-Focused Proof System Isomorphic to Expansion Proofs*, in "Journal of Logic and Computation", June 2014 [*DOI :* 10.1093/LOGCOM/EXU030], http://hal.inria.fr/hal-00937056.

[51] Z. CHIHANI, D. MILLER, F. RENAUD.*Checking Foundational Proof Certificates for First-Order Logic (extended abstract)*, in "Third International Workshop on Proof Exchange for Theorem Proving (PxTP 2013)", J. C. BLANCHETTE, J. URBAN (editors), EPiC Series, EasyChair, 2013, vol. 14, p. 58–66.

[52] Z. CHIHANI, D. MILLER, F. RENAUD.*Foundational proof certificates in first-order logic*, in "CADE 24: Conference on Automated Deduction 2013", M. P. BONACINA (editor), Lecture Notes in Artificial Intelligence, 2013, n⁰ 7898, p. 162–177.

[53] A. DAS, L. STRASSBURGER.*No complete linear term rewriting system for propositional logic*, in "26th International Conference on Rewriting Techniques and Applications (RTA 2015)", Warsaw, Poland, 26th International Conference on Rewriting Techniques and Applications (RTA 2015), June 2015 [*DOI :* 10.4230/LIPIcs.RTA.2015.127], https://hal.inria.fr/hal-01236948.

[54] C. DROSS, S. CONCHON, J. KANIG, A. PASKEVICH.*Reasoning with Triggers*, in "10th Intern. Worksh. on Satisfiability Modulo Theories, SMT 2012", P. FONTAINE, A. GOEL (editors), EPiC Series, EasyChair, June 2012, vol. 20, p. 22–31, http://www.easychair.org/publications/?page=2135488790.

[55] P. FONTAINE, J.-Y. MARION, S. MERZ, L. P. NIETO, A. TIU.*Expressiveness + Automation + Soundness: Towards Combining SMT Solvers and Interactive Proof Assistants*, in "TACAS: Tools and Algorithms for the Construction and Analysis of Systems, 12th International Conference", H. HERMANNS, J. PALSBERG (editors), LNCS, Springer, 2006, vol. 3920, p. 167–181 [*DOI : 10.1007/11691372_11*].

[56] A. GACEK, D. MILLER, G. NADATHUR.*Combining generic judgments with recursive definitions*, in "23th Symp. on Logic in Computer Science", F. PFENNING (editor), IEEE Computer Society Press, 2008, p. 33–44, http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/lics08a.pdf.

[57] H. GANZINGER, H. RUEB, N. SHANKAR.*Modularity and Refinement in Inference Systems*, SRI, 2004, n^o SRI-CSL-04-02.

[58] J.-Y. GIRARD.*Linear Logic*, in "Theoretical Computer Science", 1987, vol. 50, p. 1–102.

[59] S. GRAHAM-LENGRAND, R. DYCKHOFF, J. MCKINNA.*A Focused Sequent Calculus Framework for Proof Search in Pure Type Systems*, in "Logical Methods in Computer Science", 2011, vol. 7, n° 1, http://www.lix.polytechnique.fr/~lengrand/Work/Reports/TTSC09.pdf.

[60] S. GRAHAM-LENGRAND.*Slot Machines: an approach to the Strategy Challenge in SMT solving (presentation only)*, in "13th International Workshop on Satisfiability Modulo Theories", San Francisco, United States, July 2015, https://hal.inria.fr/hal-01211209.

[61] A. GUGLIELMI.*A System of Interaction and Structure*, in "ACM Trans. on Computational Logic", 2007, vol. 8, n° 1.

[62] A. GUGLIELMI, T. GUNDERSEN.*Normalisation Control in Deep Inference Via Atomic Flows*, in "Logical Methods in Computer Science", 2008, vol. 4, n° 1:9, p. 1–36, http://arxiv.org/abs/0709.1205.

[63] A. GUGLIELMI, L. STRASSBURGER.*Non-commutativity and MELL in the Calculus of Structures*, in "Computer Science Logic, CSL 2001", L. FRIBOURG (editor), LNCS, Springer-Verlag, 2001, vol. 2142, p. 54–68.

[64] Q. HEATH, D. MILLER.*A framework for proof certificates in finite state exploration*, in "Proceedings of the Fourth Workshop on Proof eXchange for Theorem Proving", C. KALISZYK, A. PASKEVICH (editors), Electronic Proceedings in Theoretical Computer Science, Open Publishing Association, August 2015, n° 186, p. 11–26 [*DOI : 10.4204/EPTCS.186.4*], http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/pxtp2015.pdf.

[65] D. JOVANOVIĆ, C. BARRETT, L. DE MOURA.*The Design and Implementation of the Model Constructing Satisfiability Calculus*, in "Proc. of the 13th Int. Conf. on Formal Methods In Computer-Aided Design (FMCAD '13)", FMCAD Inc., 2013, Portland, Oregon, http://www.cs.nyu.edu/~barrett/pubs/JBdM13.pdf.

[66] F. W. LAWVERE.*Adjointness in Foundations*, in "Dialectica", 1969, vol. 23, p. 281–296.

[67] C. LIANG, D. MILLER.*Focusing and Polarization in Linear, Intuitionistic, and Classical Logics*, in "Theoretical Computer Science", 2009, vol. 410, n° 46, p. 4747–4768 [*DOI : 10.1016/J.TCS.2009.07.041*], http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/tcs09.pdf.

[68] C. LIANG, D. MILLER.*A Focused Approach to Combining Logics*, in "Annals of Pure and Applied Logic", 2011, vol. 162, n° 9, p. 679–697 [*DOI : 10.1016/J.APAL.2011.01.012*].

[69] P. MARTIN-LÖF.*Constructive Mathematics and Computer Programming*, in "Sixth International Congress for Logic, Methodology, and Philosophy of Science", Amsterdam, North-Holland, 1982, p. 153–175.

[70] R. MCDOWELL, D. MILLER.*Reasoning with Higher-Order Abstract Syntax in a Logical Framework*, in "ACM Trans. on Computational Logic", 2002, vol. 3, n° 1, p. 80–136, http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/mcdowell01.pdf.

[71] R. MCDOWELL, D. MILLER.*A Logic for Reasoning with Higher-Order Abstract Syntax*, in "Proceedings, Twelfth Annual IEEE Symposium on Logic in Computer Science", Warsaw, Poland, G. WINSKEL (editor), IEEE Computer Society Press, July 1997, p. 434–445.

[72] P.-A. MELLIÈS, N. ZEILBERGER.*Functors are Type Refinement Systems*, in "42nd ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL 2015)", Mumbai, India, January 2015 [*DOI : 10.1145/2676726.2676970*], https://hal.inria.fr/hal-01096910.

[73] D. MILLER.*Forum: A Multiple-Conclusion Specification Logic*, in "Theoretical Computer Science", September 1996, vol. 165, n° 1, p. 201–232.

[74] D. MILLER, G. NADATHUR.*Programming with Higher-Order Logic*, Cambridge University Press, June 2012 [*DOI : 10.1017/CBO9781139021326*].

[75] D. MILLER, G. NADATHUR, F. PFENNING, A. SCEDROV.*Uniform Proofs as a Foundation for Logic Programming*, in "Annals of Pure and Applied Logic", 1991, vol. 51, p. 125–157.

[76] D. MILLER, A. TIU.*A Proof Theory for Generic Judgments: An extended abstract*, in "Proc. 18th IEEE Symposium on Logic in Computer Science (LICS 2003)", IEEE, June 2003, p. 118–127, http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/lics03.pdf.

[77] R. MILNER.*LCF: A Way of Doing Proofs with a Machine*, in "Proc. of the 8th Intern. Symp. on the Mathematical Foundations of Computer Science", J. BECVÁR (editor), LNCS, Springer, 1979, vol. 74, p. 146-159.

[78] S. NEGRI.*Proof Analysis in Modal Logic*, in "Journal of Philosophical Logic", 2005, vol. 34, n° 5-6, p. 507–544 [*DOI : 10.1007/s10992-005-2267-3*].

[79] G. NELSON, D. C. OPPEN.*Simplification by Cooperating Decision Procedures*, in "ACM Press Trans. on Program. Lang. and Syst.", October 1979, vol. 1, n° 2, p. 245–257 [*DOI : 10.1145/357073.357079*].

[80] F. PFENNING, C. SCHÜRMANN.*System Description: Twelf — A Meta-Logical Framework for Deductive Systems*, in "16th Conference on Automated Deduction", Trento, H. GANZINGER (editor), LNAI, Springer, 1999, n° 1632, p. 202–206.

[81] B. PIENTKA, J. DUNFIELD.*Beluga: A Framework for Programming and Reasoning with Deductive Systems (System Description)*, in "Fifth International Joint Conference on Automated Reasoning", J. GIESL, R. HÄHNLE (editors), LNCS, 2010, n° 6173, p. 15–21.

[82] M. PUECH.*Typeful Continuations*, in "Journées Francophones des Langages Applicatifs (JFLA) 2017", Gourette, France, January 2017, https://hal.archives-ouvertes.fr/hal-01419473.

[83] E. P. ROBINSON.*Proof Nets for Classical Logic*, in "Journal of Logic and Computation", 2003, vol. 13, p. 777–797.

[84] R. E. SHOSTAK.*Deciding Combinations of Theories*, in "J. ACM", 1984, vol. 31, n° 1, p. 1–12 [*DOI :* 10.1145/2422.322411].

[85] THE COQ DEVELOPMENT TEAM.*The Coq Proof Assistant Version 8.3 Reference Manual*, Inria, October 2010.

[86] A. TIU, D. MILLER.*Proof Search Specifications of Bisimulation and Modal Logics for the π-calculus*, in "ACM Trans. on Computational Logic", 2010, vol. 11, n° 2, http://arxiv.org/abs/0805.2785.

[87] N. ZEILBERGER, A. GIORGETTI.*A correspondence between rooted planar maps and normal planar lambda terms*, in "Logical Methods in Computer Science", September 2015, vol. 11, n° 3:22, p. 1-39 [*DOI :* 10.2168/LMCS-11(3:22)2015], https://hal.inria.fr/hal-01057269.

[88] N. ZEILBERGER.*Counting isomorphism classes of β-normal linear lambda terms*, September 2015, 5 pages, https://hal.archives-ouvertes.fr/hal-01214121.

[89] L. M. DE MOURA, D. JOVANOVIC.*A Model-Constructing Satisfiability Calculus*, in "Proc. of the 14th Int. Conf. on Verification, Model Checking, and Abstract Interpretation (VMCAI'13)", R. GIACOBAZZI, J. BERDINE, I. MASTROENI (editors), LNCS, Springer-Verlag, 2013, vol. 7737, p. 1–12 [*DOI :* 10.1007/978-3-642-35873-9_1].

[90] L. M. DE MOURA, G. O. PASSMORE.*The Strategy Challenge in SMT Solving*, in "Automated Reasoning and Mathematics - Essays in Memory of William W. McCune", M. P. BONACINA, M. E. STICKEL (editors), LNCS, Springer, 2013, vol. 7788, p. 15–44 [*DOI :* 10.1007/978-3-642-36675-8_2], http://dx.doi.org/10.1007/978-3-642-36675-8.

Project-Team POEMS

Wave propagation: mathematical analysis and simulation

Table of contents

Project-Team POEMS

Creation of the Project-Team: 2005 January 01

Keywords:

Computer Science and Digital Science:

6. - Modeling, simulation and control
6.1. - Mathematical Modeling
6.1.1. - Continuous Modeling (PDE, ODE)
6.1.4. - Multiscale modeling
6.1.5. - Multiphysics modeling
6.1.6. - Fractal Modeling
6.2. - Scientific Computing, Numerical Analysis & Optimization
6.2.1. - Numerical analysis of PDE and ODE
6.2.7. - High performance computing
6.3.1. - Inverse problems

Other Research Topics and Application Domains:

4.1. - Fossile energy production (oil, gas)
5.3. - Nanotechnology
5.4. - Microelectronics
6. - IT and telecom
6.2. - Network technologies
6.2.1. - Wired technologies
6.2.2. - Radio technology
6.2.3. - Satellite technology
6.2.4. - Optic technology
9.4.2. - Mathematics
9.4.3. - Physics
9.9.1. - Environmental risks

1. Members

Research Scientists

Eliane Bécache [Inria, Researcher, HDR]
Anne-Sophie Bonnet-Ben Dhia [Team leader, CNRS, Senior Researcher, HDR]
Marc Bonnet [CNRS, Researcher, HDR]
Stéphanie Chaillat [CNRS, Researcher]
Gary Cohen [Inria, Senior Researcher, until Mar 2016, HDR]
Christophe Hazard [CNRS, Researcher, HDR]
Patrick Joly [Inria, Senior Researcher, HDR]
Jean-François Mercier [CNRS, Researcher, HDR]
Axel Modave [CNRS, Researcher, from Oct 2016]

Faculty Members

Laurent Bourgeois [ENSTA, Associate Professor, HDR]
Patrick Ciarlet [ENSTA, Professor, HDR]

Bérangère Delourme [Univ. Paris XIII, Associate Professor, until Nov 2016]
Sonia Fliss [ENSTA, Associate Professor]
Marc Lenoir [ENSTA, Professor]
Eric Lunéville [CNRS, Associate Professor]
Konstantin Pankrashkin [Univ. Paris XI, Associate Professor, from Sep 2016]

Technical Staff

Colin Chambeyron [CNRS]
Nicolas Kielbasiewicz [CNRS]
Christophe Mathulik [ENSTA]

PhD Students

Zouhair Adnani
Marc Bakry [until Oct 2016]
Geoffrey Beck [until Mar 2016]
Antoine Bensalah
Antoine Bera [from Oct 2016]
Rémi Cornaggia [until Sep 2016]
Luca Desiderio
Léandre Giret
Sandrine Paolantoni [from Sep 2016]
Laure Pesudo
Arnaud Recoquillay
Yohanes Tjandrawidjaja [from Oct 2016]
Elizaveta Vasilevskaya [until Jul 2016]
Valentin Vinoles [until Aug 2016]
Emmanuel Zerbib

Post-Doctoral Fellows

Faisal Amlani [CNRS, from Nov 2016]
Samuel Groth [CNRS, until Oct 2016]
Maryna Kachanovska [ENSTA]
Dmitry Ponomarev [ENSTA, from Oct 2016]
Nicolas Salles [ENSTA]

Administrative Assistants

Corinne Chen [ENSTA]
Katia Evrat [Inria]
Emmanuelle Perrot [Inria]

Others

Shashank Anand [Inria, Internship, from May 2016 until Jul 2016]
Omar Darwiche Domingues [Inria, Internship, from Dec 2016]

2. Overall Objectives

2.1. The topic of waves

The propagation of waves is one of the most common physical phenomena one can meet in nature. From the human scale (sounds, vibrations, water waves, telecommunications, radar) to the scales of the universe (electromagnetic waves, gravity waves) and of the atoms (spontaneous or stimulated emission, interferences between particles), the emission and the reception of waves are our privileged way to understand the world that surrounds us.

The study and the simulation of wave propagation phenomena constitute a very broad and active field of research in various domains of physics and engineering sciences.

The variety and the complexity of the underlying problems, their scientific and industrial interest, the existence of a common mathematical structure to these problems from different areas justify together a research project in Scientific Computing entirely devoted to this theme.

2.2. POEMS activities

The project POEMS is an UMR (Unité Mixte de Recherche) between CNRS, ENSTA ParisTech and Inria (UMR 7231). The general activity of the project is oriented toward the design, the analysis, the numerical approximation and the control of mathematical models for the description of wave propagation in mechanics, physics and engineering sciences.

Beyond the general objective of contributing to the progress of the scientific knowledge, four goals can be ascribed to the project:

- the development of expertise relative to various types of waves (acoustic, elastic, electromagnetic, gravity waves, ...), their modelling and numerical simulation,
- the treatment of complex problems whose simulation is close enough to real life situations and industrial applications,
- the development of original mathematical and numerical techniques,
- the development of computational codes, in particular in collaboration with external partners (scientists from other disciplines, industry, state companies...)

3. Research Program

3.1. General description

Our activity relies on the existence of boundary value problems established by physicists to model the propagation of waves in various situations. The basic ingredient is a partial differential equation of the hyperbolic type, whose prototype is the wave equation (or the Helmholtz equation if time-periodic solutions are considered). Nowadays, the numerical techniques for solving the basic academic problems are well mastered. However, the solution of complex wave propagation problems close to real applications still raises (essentially open) problems which constitute a real challenge for applied mathematicians. In particular, several difficulties arise when extending the results and the methods from the scalar wave equation to vectorial problems modeling wave propagation in electromagnetism or elastodynamics.

A large part of research in mathematics, when applied to wave propagation problems, is oriented towards the following goals:

- The design of new numerical methods, increasingly accurate and efficient.
- The development of artificial transparent boundary conditions for handling unbounded propagation domains.
- The treatment of more and more complex configurations (non local models, non linear models, coupled systems, periodic media).
- The study of specific phenomena such as guided waves and resonances, which raise mathematical questions of spectral theory.
- The development of approximate models via asymptotic analysis with multiple scales (thin layers, boundary layers effects, small heterogeneities, homogenization, ...).
- The development and the analysis of algorithms for inverse problems (in particular for inverse scattering problems) and imaging techniques, using data from wave phenomena.

3.2. New schemes for time-domain simulations

Problems of wave propagation naturally arise as problems of evolution and it is necessary to have efficient methods for the calculation of their solution, directly in the time domain. The development and analysis of such methods has been in the past an important part of POEMS activity. Nowadays, there exists a large variety of higher order numerical methods that allow us to solve with good accuracy and in short computational time most classical wave propagation problems. However, when on wishes to deal with real life applications, one has to tackle problems which are complex in many ways: they involve multi-physics, non standard (possibly nonlinear) constitutive laws, highly heterogeneous media with high contrasts of coefficients, complex geometries... In many cases, such problems escape to the direct application of the above mentioned methods and ad hoc dedicated methods have to be designed. Such methods are most often of hybrid nature, which includes domain decomposition methods and subgridding, mixing of integral equations and PDEs, and artificial boundary conditions. In time domain, a particularly challenging issue is the time stability, in particular concerning the coupling of algorithms. To cope with this major difficulty, a key issue (and a kind of graal for numerical analysts) is the development of energy preserving methods which is one of the specificity of the research developed at POEMS in this field.

3.3. Integral equations

Our activity in this field aims at developing accurate and fast methods for 3D problems.

On one hand, we developed a systematic approach to the analytical evaluation of singular integrals, which arise in the computation of the matrices of integral equations when two elements of the mesh are either touching each other or geometrically close.

On the other hand, POEMS is developing Fast Boundary Element Methods for 3D acoustics or elastodynamics, with applications to soil-structure interaction, seismology or seismic imaging.

Finally, a posteriori error analysis methodologies and adaptivity for boundary integral equation formulations of acoustic, electromagnetic and elastic wave propagation is investigated in the framework of the ANR project RAFFINE.

3.4. Domain decomposition methods

This is a come back to a topic in which POEMS contributed in the 1990's. It is motivated by our collaborations with the CEA-CESTA and the CEA-LIST, for the solution of large problems in time-harmonic electromagnetism and elastodynamics.

We combine in an original manner classical ideas of Domain Decomposition Methods with the specific formulations that we use for wave problems in unbounded domains, taking benefit of the available analytical representations of the solution (integral representation, modal expansion etc...).

One ANR project (NonLocalDD) supports this research.

3.5. Wave propagation in complex media

Our objective is first to develop efficient numerical approaches for the propagation of waves in heterogeneous media, taking into account their complex microstructure.

We aim on one hand to improve homogenized modeling of periodic media, by deriving enriched boundary conditions (or transmission conditions if the periodic structure is embedded in a homogeneous matrix) which take into account the boundary layer phenomena. On the other hand, we like to develop multi-scale numerical methods when the assumption of periodicity on the spatial distribution of the heterogeneities is relaxed, or even completely lost. The general idea consists in a coupling between a macroscopic solver, based on a coarse mesh, with some microscopic representation of the field. This latter can be obtained by a numerical microscopic solver or by an analytical asymptotic expansion. This leads to two very different approaches which may be relevant for very different applications.

Extraordinary phenomena regarding the propagation of electromagnetic or acoustic waves appear in materials which have non classical properties: materials with a complex periodic microstructure that behave as materials with negative physical parameters, metals with a negative dielectric permittivity at optical frequencies, magnetized plasmas endowed with a strongly anisotropic and sign-indefinite permittivity tensor. These non classical materials raise original questions from theoretical and numerical points of view.

The objective is to study the well-posedness in this unusual context where physical parameters are sign-changing. New functional frameworks must be introduced, due, for instance, to hypersingularities of the electromagnetic field which appear at corners of metamaterials. This has of course numerical counterparts. In particular, classical Perfectly Matched Layers are unstable in these dispersive media, and new approaches must be developed.

Two ANR projects (METAMATH and CHROME) are related to this activity.

3.6. Spectral theory and modal approaches

The study of waveguides is a longstanding and major topic of the team. Concerning the selfadjoint spectral theory for open waveguides, we turned recently to the very important case of periodic media. One objective is to design periodic structures with localized perturbations to create gaps in the spectrum, containing isolating eigenvalues.

Then, we would like to go further in proving the absence of localized modes in non uniform open waveguides. An original approach has been successfully applied to the scalar problem of a waveguides junctions or bent waveguides. The challenge now is to extend these ideas to vectorial problems (for applications to electromagnetism or elastodynamics) and to junctions of periodic waveguides.

Besides, we will continue our activity on modal methods for closed waveguides. In particular, we aim at extending the enriched modal method to take into account curvature and rough boundaries.

Finally, we are developing asymptotic models for networks of thin waveguides which arise in several applications (electric networks, simulation of lung, nanophotonics...).

The study of waveguides is a longstanding and major topic of the team.

On this topic, a workshop entitled « New trends in theoretical and numerical analysis of waveguides » was co-organized by Anne-Sophie Bonnet-Ben Dhia (and Philippe Briet and Eric Soccrosi from CPT, Marseille and Michel Cristofol from I2M, Marseille) at IGESA (Porquerolles) from May 16th to 19th. This workshop is part of series of workshops organised from 2011 (in 2011 at Irmar, Rennes, in 2012 at Marseille, in 2013 at POems, Palaiseau, in 2015 at Napoli). The aim of these workshops is to bring together researchers from Mathematics, mathematical physics, theoretical physics and numerical analysis in order to encourage and stimulate the interactions between the different communities on problems associated to waveguides.

3.7. Inverse problems

Building on the strong expertise of POEMS in the mathematical modeling of waves, most of our contributions aim at improving inverse scattering methodologies.

We acquired some expertise on the so called Linear Sampling Method, from both the theoretical and the practical points of view. Besides, we are working on topological derivative methods, which exploit small-defect asymptotics of misfit functionals and can thus be viewed as an alternative sampling approach, which can take benefit of our expertise on asymptotic methods.

An originality of our activity is to consider inverse scattering in waveguides (the inverse scattering community generally considers only free-space configurations). This is motivated at the same time by specific issues concerning the ill-posedness of the identification process and by applications to non-destructive techniques, for waveguide configurations (cables, pipes, plates etc...).

Lastly, we continue our work on the so-called exterior approach for solving inverse obstacle problems, which associates quasi-reversibility and level set methods. The objective is now to extend it to evolution problems.

4. Application Domains

4.1. Acoustics

Two particular subjects have retained our attention recently.

1- Aeroacoustics, or more precisely, acoustic propagation in a moving compressible fluid, has been for our team a very challenging topic, which gave rise to a lot of open questions, from the modeling until the numerical approximation of existing models. Our works in this area are partially supported by EADS and Airbus. The final objective is to reduce the noise radiated by Airbus planes.

2- Musical acoustics constitute a particularly attractive application. We are concerned by the simulation of musical instruments whose objectives are both a better understanding of the behavior of existing instruments and an aid for the manufacturing of new instruments. We have successively considered the timpani, the guitar and the piano. This activity is continuing in the framework of the European Project BATWOMAN.

4.2. Electromagnetism

Applied mathematics for electromagnetism during the last ten years have mainly concerned stealth technology and electromagnetic compatibility. These areas are still motivating research in computational sciences (large scale computation) and mathematical modeling (derivation of simplified models for multiscale problems). These topics are developed in collaboration with CEA, DGA and ONERA.

Electromagnetic propagation in non classical media opens a wide and unexplored field of research in applied mathematics. This is the case of wave propagation in photonic crystals, metamaterials or magnetized plasmas. Two ANR projects (METAMATH and CHROME) support this research.

Finally, the simulation electromagnetic (possibly complex, even fractal) networks is motivated by destructive testing applications. This topic is developed in partnership with CEA-LIST.

4.3. Elastodynamics

Wave propagation in solids is with no doubt, among the three fundamental domains that are acoustics, electromagnetism and elastodynamics, the one that poses the most significant difficulties from mathematical and numerical points of view. A major application topic has emerged during the past years : the non destructive testing by ultra-sounds which is the main topic of our collaboration with CEA-LIST. On the other hand, we are developing efficient integral equation modelling for geophysical applications (soil-structure interaction for civil engineering, seismology).

5. Highlights of the Year

5.1. Highlights of the Year

Workshop METAMATH

This event marked the end of the project METAMATH,

funded by the French National Research Agency (ANR). The METAMATH project, led by Sonia Fliss, involved:

- from POEMS, Sonia Fliss, Anne-Sophie Bonnet Ben Dhia, Patrick Ciarlet, Patrick Joly, Camille Carvalho and Valentin Vinoles;

- from DEFI, Lucas Chesnel, Houssem Haddar, Mathieu Chamaillard and Thi Phong Nguyen;

- from Laboratoire Jacques Louis Lions, Xavier Claeys ;

- from IMATH, Université de Toulon, Guy Bouchitté and Christophe Bourel.

The motivation of this project was to contribute to the development of mathematical models for the study of periodic media and metamaterials, which are both physically relevant and available for numerical computations.

The aim of the workshop was to bring together physicists and mathematicians to make an overview of the recent researches and the new perspectives on the field.

The colloquium has taken place at Institut d'études scientifiques de Cargese, near Ajaccio, at Corsica from November 23rd until November 25th. There were about 40 participants.

Workshop on Mathematical and Numerical Modeling in Optics

This workshop, co-organized by Anne-Sophie Bonnet-Ben Dhia, was a part of the yearlong IMA (Institute of Mathematics and Applications) program in Mathematics and Optics, which brings together applied mathematicians, physical scientists and engineers to confront challenging problems arising in optics. It has taken place in Minneapolis from December 12th to December 16th.

It concerned more specifically researchers interested in the mathematical and numerical modeling of optical phenomena, especially spectral problems arising in photonics involving dispersion relations and band structures, eigenfunctions, and scattering resonances. Specific areas of focus included: (i) efficient computational methods for scattering and spectral problems and (ii) properties and optimal design of extreme materials and photonic devices. These problems arise in the study of photonic crystals and periodic media, diffraction gratings, metamaterials, graphene and related materials with Dirac points, and cloaking devices.

There were about 70 participants.

6. New Software and Platforms

6.1. COFFEE

FUNCTIONAL DESCRIPTION

COFFEE is a 3D solver for linear elastodynamics based on fast BEMs (full implementation in Fortran 90). The 3-D elastodynamic equations are solved with the boundary element method accelerated by the multi-level fast multipole method or H-matrix based solvers. The fundamental solutions for the infinite or half-space are used. A boundary element-boundary element coupling strategy is also implemented so multi-region problems (strata inside a valley for example) can be solved.

- Contact: Stéphanie Chaillat
- URL: http://perso.ensta-paristech.fr/~chaillat/index.php?page=softwares

6.2. XLiFE++

FUNCTIONAL DESCRIPTION

XLiFE++ is a Finite Element library written in C++ based on a variational approach and standard finite element methods, boundary element methods, spectral approximations. It allows to mix these different methods in a easy way to deal with complex models. A new version (v1.5) has been released in November 2016. This year, many bugs have been fixed and many improvements have been done. The main new features are: introduction of hierarchical matrix, new development for BEM computation (2D Helmholtz is now available), re-engineering of the automatic installation and compilation and connection to a Jenkins test platform. Two training days gathering thirty people have been organized in June 2016.

- Contact: Eric Lunéville and Nicolas Kielbasiewicz
- URL: http://uma.ensta-paristech.fr/soft/XLiFE++/

7. New Results

7.1. New schemes for time-domain simulations

7.1.1. *Solving the Homogeneous Isotropic Linear Elastodynamics Equations Using Potentials*
Participant: Patrick Joly.

This work is done in collaboration with Sébastien Impériale (EPI M3DISIM) and Jorge Albella from the University of Santiago de Compostela. We consider the numerical solution of 2D elastodynamic equations using the decomposition of the displacement fields into potentials. This appears as a challenge for finite element methods. We address here the particular question of free boundary conditions. A stable (mixed) variational formulation of the evolution problem is proposed based on a clever choice of Lagrange multipliers. This is expected to be efficient when the velocity of shear waves is much smaller than the velocity of pressure waves, since one can adapt the discretization to each type of waves.

7.1.2. *Discontinuous Galerkin method with high-order absorbing boundary conditions*
Participant: Axel Modave.

This work is done in collaboration with Andreas Atle from TOTAL, Jesse Chan from Rice University and Tim Warburton from Virginia Tech.

Discontinuous Galerkin finite element schemes exhibit attractive features for large-scale time-domain wave-propagation simulations on modern parallel architectures (e.g. GPU clusters). For many applications, these schemes must be coupled with non-reflective boundary treatments to limit the size of the computational domain without losing accuracy or computational efficiency, which remains a challenging task.

We propose a combination of a nodal discontinuous Galerkin method with high-order absorbing boundary conditions (HABCs) for cuboidal computational domains. Compatibility conditions are derived for HABCs intersecting at the edges and the corners of a cuboidal domain. We propose a GPU implementation of the computational procedure, which results in a multidimensional solver with equations to be solved on 0D, 1D, 2D and 3D spatial regions. Numerical results demonstrate both the accuracy and the computational efficiency of our approach.

7.2. Integral equations

7.2.1. *Mesh adaptation for the fast multipole method in acoustics*
Participants: Faisal Amlani, Stéphanie Chaillat, Samuel Groth.

This work is done in collaboration with Adrien Loseille (EPI Gamma3). We introduce a metric-based anisotropic mesh adaptation strategy for the fast multipole accelerated boundary element method (FM-BEM) applied to exterior boundary value problems of the three-dimensional Helmholtz equation. The present methodology is independent of discretization technique and iteratively constructs meshes refined in size, shape and orientation according to an *optimal* metric reliant on a reconstructed Hessian of the boundary solution. The resulting adaptation is anisotropic in nature and numerical examples demonstrate optimal convergence rates for domains that include geometric singularities such as corners and ridges.

7.2.2. *Coupling integral equations and high-frequency methods*
Participants: Marc Bonnet, Marc Lenoir, Eric Lunéville, Laure Pesudo, Nicolas Salles.

This theme concerns wave propagation phenomena which involve two different space scales, namely, on the one hand, a medium scale associated with lengths of the same order of magnitude as the wavelength (medium-frequency regime) and on the other hand, a long scale related to lengths which are large compared to the wavelength (high-frequency regime). Integral equation methods are known to be well suited for the former, whereas high-frequency methods such as geometric optics are generally used for the latter. Because of the presence of both scales, both kinds of simulation methods are simultaneously needed but these techniques do not lend themselves easily to coupling.

A first situation, considered by Marc Lenoir, Eric Lunéville and Nicolas Salles, is the scattering of an acoustic wave by two sound-hard obstacles: a large obstacle subject to high-frequency regime relatively to the wavelength and a small one subject to medium-frequency regime. The technique proposed in this case consists in an iterative method which allows to decouple the two obstacles and to use Geometric Optics for the large obstacle and Boundary Element Method for the small obstacle. The method is implemented on the XLife++ library developed in the lab.

The second situation, undertaken in the context of the PhD thesis of Laure Pesudo, is the subject of a partnership with CEA LIST and a collaboration with Francis Collino. Modelling ultrasonic non destructive testing (NDT) experiments simultaneously involves the scattering of waves by defects of moderate size (for which discretization-based methods such as the BEM are appropriate) and propagation over large distances (requiring high-frequency approximations). A new hybrid strategy between the boundary element method (BEM) and ray tracing is proposed in order to allow the accurate and quick simulation of high frequency Non Destructive Testing (NDT) configurations involving diffraction phenomena. Results from its implementation to 2D acoustic NDT-like diffraction configurations have been obtained. The strategy proposed is however generic, and can be extended to three-dimensional configurations and elastodynamic wave propagation.

7.2.3. *Dynamic soil-structure interaction*

Participants: Marc Bonnet, Stéphanie Chaillat, Zouhair Adnani.

This work, undertaken in the context of the PhD thesis of Zouhair Adnani (CIFRE partnership with EDF), concerns the simulation of dynamic soil-structure interaction (SSI) in connection with seismic assessment of civil engineering structures. Because of the complementary specificities of the finite element method (FEM) and the boundary element method (BEM), it is natural to use the BEM to model the unbounded soil domain, while the FEM is applied for the bounded region comprising the structure undergoing assessment, and possibly its close-range soil environment.

The originality of this work is to formulate, implement, and evaluate on realistic test examples, a computational strategy that combines the fast multipole accelerated boundary element method (visco-elastodynamic COFFEE solver), and the EDF in-house FEM code Code_Aster. In a preliminary phase, the evaluation of transient elastodynamic responses via the Fourier synthesis of frequency domain solutions computed using COFFEE (see Section 5.1) has been studied on several test problems, achieving substantial improvements of computational efficiency for this component of SSI analysis.

The coupling between the two methods is then done in a black-box fashion with the substructuring method by computing the soil impedance (i.e. elastodynamic Poincaré-Steklov) operator relating forces to displacements on the FEM-BEM coupling interface. One of the main challenges is that this operator cannot be assembled due to the iterative nature of the FM-BEM and the potentially large number of degrees of freedom supported by the interface. To reduce the computational costs, we instead compute its projection on a reduced basis of interface modes, which requires to perform as many FM-BEM calculations as interface modes selected. This approach has so far been compared to reference solutions and validated for superficial and buried foundations on homogeneous or heterogeneous soil.

7.2.4. *Volume Integral Formulations*

Participant: Marc Bonnet.

Volume integral equations (VIEs), also known as Lippmann-Schwinger integral equations, arise naturally when considering the scattering of waves by penetrable, and possibly heterogeneous, inhomogeneities embedded in a homogeneous background medium (for which a fundamental solution is explicitly known). Their derivation and use in e.g. acoustics, elastodynamics or electromagnetism goes back several decades. Since their geometrical support is confined to the spatial region where material properties differ from the background, VIEs are in particular useful for the derivation and justification of homogenized or asymptotic models (the latter providing our main motivation for this study, in connection with [section gradient topologique]). By directly linking remote measurements to unknown inhomogeneities, VIEs also provide a convenient forward modeling approach for medium imaging inverse problems. However, whereas the theory of boundary integral

equations is extensively documented, the mathematical properties of VIEs have undergone a comparatively modest coverage, much of it pertaining to electromagnetic scattering problems.

In this work, we investigate the solvability of VIE formulations arising in elastodynamic scattering by penetrable obstacles. The elasticity tensor and mass density are allowed to be smoothly heterogeneous inside the obstacle and may be discontinuous across the background-obstacle interface, the background elastic material being homogeneous. Both materials may be anisotropic, within certain limitations for the background medium.

Towards this goal, we have introduced a modified version of the singular volume integral equation (SVIE) governing the corresponding elastostatic (i.e. zero frequency) problem, and shown it to be of second kind involving a contraction operator, i.e. solvable by Neumann series, for any background material and inhomogeneity material and geometry. Then, the solvability of VIEs for frequency-domain elastodynamic scattering problems follows by a compact perturbation argument, assuming uniqueness to be established. In particular, in an earlier work, we have established a uniqueness result for the anisotropic background case (where, to avoid difficulties associated with existing radiation conditions for anisotropic elastic media, we have proposed an alternative definition of the radiating character of solutions, which is equivalent to the classical Sommerfeld-Kupradze conditions for the isotropic background case). This investigation extends work by Potthast (1999) on 2D electromagnetic problems (transverse-electric polarization conditions) involving orthotropic inhomogeneities in a isotropic background, and contains recent results on the solvability of Eshelby's equivalent inclusion problem as special cases. The proposed modified SVIE is also useful for fixed-point iterative solution methods, as Neumannn series then converge (i) unconditionally for static problems and (ii) on some inhomogeneity configurations for which divergence occurs with the usual SVIE for wave scattering problems.

7.3. Domain decomposition methods

7.3.1. *Transparent boundary conditions with overlap in unbounded anisotropic media*
Participants: Anne-Sophie Bonnet-Ben Dhia, Sonia Fliss, Yohanes Tjandrawidjaja.

This work is done in the framework of the PhD of Yohanes Tjandrawidjaja, funded by CEA-LIST, in collaboration with Vahan Baronian form CEA. This follows the PhD of Antoine Tonnoir (now Assistant Professor at Insa of Rouen) who developed a new approach, the Half-Space Matching Method, to solve scattering problems in 2D unbounded anisotropic media. The objective is to extend the method to a 3D plate of finite width.

In 2D, our approach consists in coupling several plane-waves representations of the solution in half-spaces surrounding the defect with a FE computation of the solution around the defect. The difficulty is to ensure that all these representations match, in particular in the infinite intersections of the half-spaces. It leads to a Fredholm formulation which couples, via integral operators, the solution in a bounded domain including the defect and some traces of the solution on the edge of the half-planes.

The extension to 3D elastic plates requires some generalizations of the formulation which are not obvious. In particular, we have to use Neumann traces of the solution, which raises difficult theoretical questions.

As a first step, we have considered a scattering problem outside a convex polygonal scatterer for a general class of boundary conditions, using the Half-Space Matching Method. Using the Mellin Transform, we are able to show that this system is coercive + compact in presence of dissipation.

Let us mention that the Half-Space Matching Method has been extended successfully by Julian Ott (Karlsruhe Institut für Technologie) to the scattering by junctions of open waveguides in 2D.

7.3.2. *Domain Decomposition Methods for the neutron diffusion equation*
Participants: Patrick Ciarlet, Léandre Giret.

This work is done in collaboration with Erell Jamelot (CEA-DEN, Saclay) and Félix Kpadonou (LMV, UVSQ). Studying numerically the steady state of a nuclear core reactor is expensive, in terms of memory

storage and computational time. In its simplest form, one must solve a neutron diffusion equation with low-regularity solutions, discretized by mixed finite element techniques, within a loop. Iterating in this loop allows to compute the smallest eigenvalue of the system, which determines the critical, or non-critical, state of the core. This problem fits within the framework of high performance computing so, in order both to optimize the memory storage and to reduce the computational time, one can use a domain decomposition method, which is then implemented on a parallel computer: this is the strategy used for the APOLLO3 neutronics code. The development of non-conforming DD methods for the neutron diffusion equation with low-regularity solutions has recently been finalized, cf. [PC,EJ,FK'1x]. The theory for the eigenvalue problem is also understood. The current research now focuses on the numerical analysis of the full suite of algorithms to prove convergence for the complete multigroup SPN model (which involves coupled diffusion equations).

7.4. Wave propagation in complex media

7.4.1. *Perfectly Matched Layers in plasmas and metamaterials*
Participants: Eliane Bécache, Maryna Kachanovska.

In this work we consider the problem of the modelling of 2D anisotropic dispersive wave propagation in unbounded domains with the help of perfectly matched layers (PML). We study the Maxwell equations in passive media with the frequency-dependent diagonal tensor of dielectric permittivity and magnetic permeability. An application of the traditional PMLs to this kind of problems often results in instabilities, due to the presence of so-called backward propagating waves. In previous works, this instability was overcome with the help of the frequency-dependent correction of the PML, for isotropic dispersive models.

We show that this idea can be extended to a more general class of models (uniaxial cold plasma, some anisotropic metamaterials). Crucially, we base our considerations on the Laplace-domain techniques. This allows to avoid the analysis of the group and phase velocity (used before) but study (rather formally) coercivity properties of the sesquilinear form corresponding to the PML model in the Laplace domain. The advantage of this method is that it permits to treat problems with dissipation, and provides an intuition on how to obtain explicit energy estimates for the resulting PML models in the time domain. However, such analysis does not allow to obtain easily the necessary stability condition of the PML. We demonstrate that the necessary stability conditions of the PML can be rewritten for a class of models in a form that is easy to verify, and demonstrate that these conditions are sufficient for the stability of the new PMLs with the help of the Laplace-domain techniques. Thanks to the Laplace domain analysis, we are able to rewrite a PML system in the time domain in a form, for which the derivation of energy estimates is simplified (compared to other formulations).

7.4.2. *Transparent Boundary Conditions for the Wave Propagation in Fractal Trees*
Participants: Patrick Joly, Maryna Kachanovska.

This work, done in collaboration with Adrien Semin (Postdoctoral student at the Technische Universität of Berlin), is dedicated to an efficient resolution of the wave equation in self-similar trees (e.g. wave propagation in a human lung). In this case it is possible to avoid computing the solution at deeper levels of the tree by using the transparent boundary conditions. The corresponding DtN operator is defined by a functional equation in the frequency domain. In this work we propose and compare two approaches to the discretization of this operator in the time domain. The first one is based on the multistep convolution quadrature, while the second one stems from the rational approximations.

7.4.3. *High order transmission conditions between homogeneous and homogenized periodic half-spaces*
Participants: Sonia Fliss, Valentin Vinoles.

This work is a part of the PhD of Valentin Vinoles, and is done in collaboration with Xavier Claeys from Paris 6 University and EPI ALPINE. It is motivated by the fact that classical homogenization theory poorly takes into account interfaces, which is particularly unfortunate when considering negative materials, because important phenomena arise precisely at their surface (plasmonic waves for instance). To overcome this

limitation, we want to construct high order transmission conditions. For now, we have treated the case of a plane interface between a homogeneous and a periodic half spaces. Using matched asymptotic techniques, we have derived high order transmission conditions. We have then introduced an approximate model associated to this asymptotic expansions which consists in replacing the periodic media by an effective one but the transmission conditions are not classical. The obtained conditions involve Laplace- Beltrami operators at the interface and requires to solve cell problems in periodicity cell (as in classical homogenisation) and in infinite strips (to take into account the phenomena near the interface). We establish well posedness for the approximate and error estimate which justify that this new model is more accurate near the interface and in the bulk. From a numerical point of view, the only difficulty comes from the problems set in infinite strips (one half is homogeneous and the other is periodic). This is overcome using DtN operators corresponding to the homogeneous and the periodic media.The numerical results confirm the theoretical ones.

7.4.4. *Finite Element Heterogeneous Multiscale Method for Maxwell's Equations*
Participants: Patrick Ciarlet, Sonia Fliss.

This work is done in collaboration with Christian Stohrer (Karlsruhe Institute of Technology, Allemagne). In recent years, the Finite Element Heterogeneous Multiscale Method (FE-HMM) has been used to approximate the effective behavior of solutions to PDEs in highly oscillatory media. We started on the extension of the FE-HMM to the Helmholtz equation in such media, and recently we solved the time-harmonic Maxwell equations case. Using a combination of results regarding the FE-HMM and the notion of T-coercivity applied to Maxwell's equations, we derive an a priori error bound and the error. Moreover, numerical experiments corroborate the analytical findings, cf. [PC,SF,CS'1x].

7.5. Spectral theory and modal approaches for waveguides

7.5.1. *Plasmonic waveguides*
Participants: Anne-Sophie Bonnet-Ben Dhia, Patrick Ciarlet.

This work is done in collaboration with Camille Carvalho (UC Merced, California, USA) and Lucas Chesnel (EPI DEFI). A plasmonic waveguide is a cylindrical structure consisting of metal and dielectric parts. In a certain frequency range, the metal can be seen as a lossless material with a negative dielectric permittivity. The study of the modes of a plasmonic waveguide is then presented as a model eigenvalue problem with a sign-change of coefficients in the main part of the operator. Depending on the values of the contrast of permittivities at the metal/dielectric interface, different situations may occur. We focus on the situation where the interface between metal and dielectric presents corners. For a particular contrast range, the problem is neither self-adjoint nor with compact resolvent, this is the "critical" case. Whereas in the "nice" case, the problem is self-adjoint with compact resolvent and admits two sequences of eigenvalues tending to $-\infty$ and $+\infty$. In the "critical" case, Kondratiev's theory of singularities allows to build extensions of the operator, with compact resolvent. We show that the eigenvalues for one of these extensions can be computed by combining finite elements and Perfectly Matched Layers at the corners. The paradox is that a specific treatment has to be done to capture the corners singularities, even to compute regular eigenmodes. In the "nice" case, we propose and analyze numerical techniques based on the notion of T-coercive meshes that allow to solve the model problem.

7.5.2. *Modal analysis of electromagnetic dispersive media*
Participants: Christophe Hazard, Sandrine Paolantoni.

We investigate the spectral effects of an interface between a usual dielectric and a negative-index material (NIM), that is, a dispersive material whose electric permittivity and magnetic permeability become negative in some frequency range. We consider here an elementary situation, namely, 1) the simplest existing model of NIM : the Drude model (for which negativity occurs at low frequencies); 2) a two-dimensional scalar model derived from the complete Maxwell's equations; 3) the case of a simple bounded cavity: a camembert-like domain partially

lled with a portion of non dissipative Drude material. Because of the frequency dispersion (the permittivity and permeability depend on the frequency), the spectral analysis of such a cavity is unusual since it yields a nonlinear eigenvalue problem. Thanks to the use of an additional unknown, we show how to linearize the problem and we present a complete description of the spectrum.

7.5.3. *Formulation of invisibility in waveguides as an eigenvalue problem*
Participants: Antoine Bera, Anne-Sophie Bonnet-Ben Dhia.

This work is done in collaboration with Lucas Chesnel from EPI DEFI, Vincent Pagneux from Laboratoire d'Acoustique de l'Université du Maine and Sergei Nazarov from Russian Academy of Sciences. A scatterer placed in an infinite waveguide may be *invisible* at particular discrete frequencies. We consider two different definitions of invisibility: no reflection (but possible conversion or phase shift in transmission) or perfect invisibility (the scattered field is exponentially decaying at infinity). Our objective is to show that the invisibility frequencies can be characterized as eigenvalues of some spectral problems. Two different approaches will be used for the two different definitions of invisibility, leading to non-selfadjoint eigenvalue problems. Concerning the non-reflection case, our approach based on an original use of PMLs allows to extend to higher dimension and to complex eigenvalues the results obtained by Hernandez-Coronado, Krejcirik and Siegl on a 1D model problem.

7.5.4. *Transparent boundary conditions for general waveguide problems*
Participants: Anne-Sophie Bonnet-Ben Dhia, Sonia Fliss.

In this work, done in collaboration with Antoine Tonnoir from INSA of Rouen, we propose a construction of transparent boundary conditions which can be used for quite general waveguide problems. Classical Dirichlet-to-Neumann maps used for homogeneous acoustic waveguides can be constructed using separation of variables and the orthogonality of the modes on one transverse section. These properties are also important for the mathematical and numerical analysis of problems involving DtN maps. However this framework does not extend directly to strati

ed, anisotropic or periodic waveguides and for Maxwell's or elastic equations. The difficulties are that (1) the separation of variables is not always possible and (2) the modes of the waveguides are not necessarily orthogonal on the transverse section. We propose an alternative to the DtN maps which uses two artificial boundaries and is constructed using a more general orthogonality property.

7.6. Inverse problems

7.6.1. *Linear Sampling Method with realistic data in waveguides*
Participants: Laurent Bourgeois, Arnaud Recoquillay.

Our activities in the field of inverse scattering in waveguides with the help of sampling methods has now a quite long history. We now intend to apply these methods in the case of realistic data, that is surface data in the time domain. This is the subject of the PhD of Arnaud Recoquillay. It is motivated by Non Destructive Testing activities for tubular structures and is the object of a partnership with CEA List (Vahan Baronian).

Our strategy consists in transforming the time domain problem into a multi-frequency problem by the Fourier transform. This allows us to take full advantage of the established efficiency of modal frequency-domain sampling methods. We have already proved the feasibility of our approach in the 2D acoustic and 2D elastic case. In particular, we have shown how to optimize the number of sources/receivers and the distance between them in order to obtain the best possible identification result. Experiments are currently carried in CEA.

7.6.2. *The "exterior approach" to solve inverse obstacle problems*

Participants: Laurent Bourgeois, Arnaud Recoquillay.

We consider some inverse obstacle problems in acoustics by using a single incident wave, either in the frequency or in the time domain. When so few data are available, a Linear Sampling type method cannot be applied. In order to solve those kinds of problem, we propose an "exterior approach", coupling a mixed formulation of quasi-reversibility and a simple level set method. In such iterative approach, for a given defect D, we update the solution u with the help of a mixed formulation of quasi-reversibility while for a given solution u, we update the defect D with the help of a level set method based on a Poisson problem. The case of data in the frequency domain has been studied for the waveguide geometry. We currently investigate the case of data in a finite time domain.

7.6.3. *Topological derivatives of leading- and second-order homogenized coefficients.*

Participants: Marc Bonnet, Rémi Cornaggia.

This work is done in collaboration with Bojan Guzina from University of Minnesota. We derive the topological derivatives of the homogenized coefficients associated to a periodic material, with respect of the small size of a penetrable inhomogeneity introduced in the unit cell that defines such material. In the context of antiplane elasticity, this work extends existing results to (i) time-harmonic wave equation and (ii) second-order homogenized coefficients, whose contribution reflects the dispersive behavior of the material.

7.6.4. *A continuation method for building large invisible obstacles in waveguides*

Participants: Antoine Bera, Anne-Sophie Bonnet-Ben Dhia.

In collaboration with Lucas Chesnel (EPI DEFI) and Sergei Nazarov (Saint-Petersburg University), we consider time harmonic acoustic problems in waveguides. We are interested in finding localized perturbations of a straight waveguide which are not detectable in the far field, as they produce neither reflection nor conversion of propagative modes. In other words, such *invisible* perturbation produces a scattered field which is exponentially decaying at infinity in the two infinite outlets of the waveguide.

In our previous contributions, we found a way to build smooth and small perturbations of the boundary which were almost invisible, in the sense that they were producing no reflexions but maybe a phase shift in transmission.

The method is constructive and has been validated numerically. But the drawback is that it is limited to small perturbations. In the present work, we show that the previous idea can be combined with a continuation method, in order to get larger invisible perturbations.

7.7. Aeroacoustics

7.7.1. *Time-harmonic acoustic scattering in a vortical flow*

Participants: Antoine Bensalah, Patrick Joly, Jean-François Mercier.

This activity is done in the framework of the PhD of Antoine Bensalah, in partnership with Airbus Group. We study the time-harmonic acoustic radiation in a fluid in a general flow which is not curl free, but has restricted vortical areas. The objective is to take into account the complicated coupling between acoustics and hydrodynamics. The Galbrun approach developed previously in 2D is too expensive in terms of degrees of freedom for 3D simulations. As an alternative, we propose to consider instead the Goldstein equations, which are vectorial only in the vortical areas and remain scalar elsewhere.

To begin with, we aim at determining the acoustic field radiated in 2D by a time-harmonic source in a fluid in flow. Goldstein's equations are proved to be well-posed outside a spectrum of frequencies corresponding to resonant streamlines. This band spectrum is explicitly determined for two simple geometries (an annular domain and a rectangular one with periodic conditions). Then the full model is shown to be well-posed under a coercivity condition, implying a subsonic flow with a small enough vorticity.

7.7.2. *Propagation of solitons through Helmholtz resonators*
Participant: Jean-François Mercier.

With Bruno Lombard (Laboratoire de Mécanique et Acoustique of Marseille), we study the propagation of nonlinear solitary acoustic waves in a 1D waveguide connected to a lattice of Helmholtz resonators. We start from an homogenized model of the literature, consisting of two coupled equations evolution: a nonlinear PDE describing acoustic waves (similar to the Burgers equation), and a linear ODE describing oscillations in the Helmholtz resonators. We have already developed a numerical modeling of this model and we have compared simulations with experimental data.

The drawback of the homogenized model is that all the resonators must be the same. In particular the reflection of an incident wave by a defect cannot be considered. To remedy this limitation, we have proposed an extension of the model, predicting two-way propagation across variable resonators. Thanks to a new discrete description of the resonators, the improved model takes into account two important features: resonators of different strengths and back-scattering effects. Comparisons with experimental data show that a closer agreement is obtained.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Contract POEMS-DGA
Participants: Eric Lunéville, Marc Lenoir, Séphanie Chaillat, Nicolas Kielbasiewicz, Nicolas Salles.

Start : 2015, End : 2018. Administrator : ENSTA.
This contract is in partnership with François Alouges and Matthieu Aussal (CMAP, Ecole Polytechnique) and concerns the improvement of Boundary Element Methods for wave propagation problems.

Contract POEMS-CEA-LIST
Participants: Marc Bonnet, Laure Pesudo.

Start : 12/01/2014, End : 11/31/2017. Administrator : CNRS.
This contract is about the coupling between high frequency methods and integral equations.

Contract POEMS-SHELL
Participants: Stéphanie Chaillat, Patrick Ciarlet, Luca Desiderio.

Start : 10/01/2013, End : 09/31/2016. Administrator : CNRS.
This contract is about fast direct solvers to simulate seismic wave propagation in complex media.

Contract POEMS-EDF
Participants: Stéphanie Chaillat, Marc Bonnet, Zouhair Adnani.

Start : 12/01/2014, End : 11/31/2017. Administrator : CNRS.
This contract is about fast solvers to simulate soil-structure interactions.

9. Partnerships and Cooperations

9.1. Regional Initiatives

The post-doc of Maryna Kachanovska si funded by the Fondation Mathématique Jacques Hadamard (FMJH).

9.2. National Initiatives

9.2.1. ANR

- ANR project *METAMATH: modélisation mathématique et numérique pour la propagation des ondes en présence de métamatériaux.* Partners: EPI DEFI (Inria Saclay), IMATH-Université de Toulon, LJLL-Paris 6 University.
  Start : 12/01/2011, End : 11/30/2016. Administrator : Inria. Coordinator : Sonia Fliss.
- ANR project *CHROME: Chauffage , réflectométrie et Ondes pour les plasmas magnétiques*
  Partners: LJLL-Paris 6 University, Université de Lorraine
  Start : 10/01/2012, End : 19/09/2016 Administrator : Inria Coordinator for POEMS: Eliane Bécache
- ANR project *RAFFINE: Robustesse, Automatisation et Fiabilité des Formulations INtégrales en propagation d'ondes : Estimateurs a posteriori et adaptivité*
  Partners: EADS, IMACS, ONERA, Thales
  Start : January 2013. End : June 2017. Administrator : Inria. Coordinator: Marc Bonnet.
- ANR project *Non-Local Domain Decomposition Methods in Electromagnetism.*
  Partners: Inria Alpines, Inria POEMS, Inria Magique 3D.
  Start : 2015, End : 2019. Administrator : Inria. Coordinator: Xavier Claeys.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. BATWOMAN

Type: FP7 Marie Curie

Objectif: Basic Acoustics Training - & Workprogram On Methodologies for Acoustics - Network

Duration: September 2013 - August 2017

Coordinator: Martin Wifling, VIRTUAL VEHICLE (AT)

Inria contact: P. Joly

Abstract: The BATWOMAN ITN aims at structuring research training in basic and advanced acoustics and setting up a work program on methodologies for acoustics for skills development in a highly diverse research field offering multiple career options.

9.4. International Initiatives

9.4.1. Inria International Partners

9.4.1.1. Informal International Partners

Wilkins Aquino (Duke University)

Eric Chung (Chinese University of Hong Kong)

Bojan Guzina (University of Minnesota)

Sergei Nazarov (Saint-Petersburg University)

Jeronimo Rodriguez (University of Santiago de Compostela)

Adrien Semin (Technische Universitat Berlin)

Julian Ott (Karlsruhe Institut für Technologie)

9.5. International Research Visitors

9.5.1. Visits of International Scientists

Stefan Sauter, University of Zürich, Switzerland (2 months)

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Advisory and management activities

- P. Joly is a member of the scientific committee of CEA-DAM.
- E. Lunéville is the Head of UMA (Unité de Mathématiques Appliquées) at ENSTA ParisTech.

10.1.2. Scientific events organisation end selection

- E. Bécache, A. S. Bonnet-Ben Dhia, M. Bonnet, S. Fliss, C. Hazard, P. Joly and E. Lunéville are members of the scientific committee for the 13rd international conference on mathematical and numerical aspects of wave propagation (WAVES 2017), which will take place in Minneapolis in May 2017.
- A. S. Bonnet-Ben Dhia has been a member of the organizing committee of the workshop *NEW TRENDS IN THEORETICAL AND NUMERICAL ANALYSIS OF WAVEGUIDES* (40 participants) which has been held in Porquerolles in May 2016.
- S. Fliss has been a member of the organizing committee of the workshop *METAMATH* (40 participants) which has been held in Cargese in November 2016.
- A. S. Bonnet-Ben Dhia has been a member of the organizing committee of the workshop on *Mathematical and Numerical Modeling in Optics* (70 participants) which has been held in Minneapolis in December 2016.

10.1.3. Journal

- A. S. Bonnet-Ben Dhia is associate editor of SINUM (SIAM Journal of Numerical Analysis) and SIAP (SIAM Journal of Applied Mathematics).
- M. Bonnet is associate editor of Engineering Analysis with Boundary Elements
- M. Bonnet is in the editorial board of Inverse Problems.
- M. Bonnet is in the editorial board of Computational Mechanics.
- M. Bonnet is in the editorial board of Journal of Optimization Theory and Application.
- P. Ciarlet is an editor of CAMWA (Computers & Mathematics with Applications).
- P. Ciarlet is an editor of ESAIM:M2AN (Mathematical Modeling and Numerical Analysis).
- P. Joly is an editor of ESAIM:M2AN (Mathematical Modeling and Numerical Analysis).
- P. Joly is a member of the editorial board of AAMM (Advances in Applied Mathematics and Mechanics).
- P. Joly is a member of the Book Series Scientific Computing of Springer Verlag.
- The team members regularly review papers for many international journals.

10.2. Teaching - Supervision

10.2.1. Teaching

Eliane Bécache
- *Méthode des éléments finis*, ENSTA ParisTech (2nd year)
- *Compléments sur la méthode des éléments finis*, ENSTA ParisTech, (2nd year)
- *Fonctions d'une variable complexe*, ENSTA ParisTech (1st year)

Marc Bonnet
- *Problèmes inverses*, Master MS2SC (Centrale Paris and ENS Cachan)

- *Méthodes intégrales*, Master TACS (ENS Cachan)
- *Outils élémentaires d'analyse pour les équations aux dérivées partielles*, ENSTA Paris-Tech (1st year)

Anne-Sophie Bonnet-Ben Dhia

- *Fonctions d'une variable complexe*, ENSTA ParisTech (1st year)
- *Propagation dans les guides d'ondes*, ENSTA ParisTech (3rd year) and Master "Modeling and Simulation" (M2)
- *Théorie spectrale des opérateurs autoadjoints et applications aux guides optiques*, ENSTA ParisTech (2nd year)

Laurent Bourgeois

- *Outils élémentaires pour l'analyse des équations aux dérivées partielles*, ENSTA Paris-Tech (1st year)
- *Inverse problems: mathematical analysis and numerical algorithms*, (Master AN& EDP, Paris 6 and Ecole Polytechnique)

Stéphanie Chaillat

- *Introduction à la discrétisation des équations aux dérivées partielles*, ENSTA ParisTech (1st year)
- *Fonctions d'une variable complexe*, ENSTA ParisTech (1st year)
- *Equations intégrales et multipôles rapides*, Ecole doctorale MODES (Univ. Paris Est, Marne la Vallée)
- *Résolution des problèmes de diffraction par équations intégrales*, ENSTA ParisTech (3rd year) and Master "Modeling and Simulation" (M2)

Colin Chambeyron

- *Analyse réelle: optimisation libre et sous contraintes*, Dauphine University (1st year)
- *Outils mathématiques*, Dauphine University (1st year)
- *Algèbre linéaire*, Dauphine University (2nd year)

Patrick Ciarlet

- Advanced Finite Element Methods, ENSTA ParisTech (2nd year)
- Parallel Scientific Computing, ENSTA ParisTech (3rd year), and Master "Analysis, Modelling, Simulation" (M2)
- Mathematical Models and their Discretisation in Electromagnetism, ENSTA ParisTech (3rd year), and Master "Analysis, Modelling, Simulation" (M2)
- Deputy Head of the Master's Program Analysis, Modelling, Simulation, Paris-Saclay University

Sonia Fliss

- *Méthode des éléments finis*, ENSTA ParisTech (2nd year)
- *Introduction à la discrétisation des équations aux dérivées partielles*, ENSTA ParisTech (1st year).
- *Propagation des ondes dans les milieux périodiques*, ENSTA ParisTech (3rd year) and Master "Modeling and Simulation" (M2)
- *Homogeneisation*, Master ANEDP Paris 6 and Ecole Polytechnique (M2)

Christophe Hazard

- *Outils élémentaires d'analyse pour les équations aux dérivées partielles*, ENSTA Paris-Tech (1st year)

- *Théorie spectrale des opérateurs autoadjoints et applications aux guides optiques*, ENSTA ParisTech (2nd year)

Patrick Joly

- *Introduction à la discrétisation des équations aux dérivées partielles*, ENSTA ParisTech (1st year)
- *Propagation des ondes dans les milieux périodiques*, ENSTA ParisTech (3rd year) and Master "Modeling and Simulation" (M2)

Nicolas Kielbasiewicz

- *Programmation scientifique et simulation numérique*, ENSTA ParisTech (2nd year)
- *Parallélisme et calcul réparti*, ENSTA ParisTech (Master 2)

Marc Lenoir

- *Fonctions d'une variable complexe*, ENSTA ParisTech (2nd year)
- *Equations intégrales*, ENSTA ParisTech (3rd year) and Master "Modeling and Simulation" (M2)
- *Méthodes asymptotiques hautes fréquences pour les équations d'ondes - course notes*, ENSTA ParisTech (3rd year) and Master "Modeling and Simulation" (M2)

Eric Lunéville

- *Introduction au Calcul Scientifique*, ENSTA ParisTech (2nd year).
- *Programmation scientifique et simulation numérique*, ENSTA ParisTech (2nd year).
- *Propagation dans les guides d'ondes*, ENSTA ParisTech (3rd year) and Master "Modeling and Simulation" (M2)

Jean-François Mercier

- *Outils élémentaires d'analyse pour les équations aux dérivées partielles*, ENSTA ParisTech (1st year)
- *Fonctions d'une variable complexe*, ENSTA ParisTech, ENSTA ParisTech (2nd year)
- *Théorie spectrale des opérateurs autoadjoints et application aux guides optiques*, ENSTA ParisTech (2nd year)

Axel Modave

- *Scientific parallel computing*, ENSTA ParisTech (3rd year) and Master "Modeling and Simulation" (M2)

10.2.2. Supervision

PhD: Marc Bakry, "Estimateurs a posteriori pour la résolution des problèmes de diffraction par équations intégrales", October 2016, Patrick Ciarlet and Sébastien Pernet.

PhD: Geoffrey Beck, "Modélisation de la propagation d'ondes électromagnétiques dans des câbles co-axiaux", March 2016, Patrick Joly et Sébastien Impériale

PhD: Mathieu Chamaillard, "Conditions aux limites effectives pour des revêtements minces périodiques", October 2011, Patrick Joly and Houssem Haddar

PhD: Rémi Cornaggia, "Asymptotique petit-défaut de fonctions-coût et son application en identification: justifications théorique et expérimentale, extensions", September 2016, Marc Bonnet and Bojan Guzina

PhD: Elizaveta Vasilevskaia, "Modes localisés dans les guides d'onde quantiques", July 2016, Patrick Joly

PhD: Valentin Vinoles, "Analyse asymptotique des équations de Maxwell en présence de métamatériaux", September 2016, Sonia Fliss and Patrick Joly

PhD in progress : Zouhair Adnani , "Modélisation numérique tridimensionnelle des effets de site en interaction sol-structure par une méthode adaptée aux problèmes sismiques de très grande taille", October 2014, Marc Bonnet and Stéphanie Chaillat

PhD in progress : Antoine Bensalah, "Une approche nouvelle de la modélisation mathématique et numérique en aéroacoustique par les équations de Goldstein et applications en aéronautique", October 2014, Patrick Joly and Jean-François Mercier

PhD: Antoine Bera, "Conception de perturbations invisibles pour les ondes électromagnétiques ou acoustiques", October 2016, Anne-Sophie Bonnet-Ben Dhia and Lucas Chesnel

PhD in progress :Luca Desiderio, "Efficient visco-eleastic wave propagation in 3D for high contrast media", October 2013, Stéphanie Chaillat and Patrick Ciarlet

PhD in progress :Léandre Giret, "Development of a domain decomposition method on non-conforming meshes: application to the modeling of a Reactivity-Initiated Accident (RIA) in a Pressurized Water Reactor (PWR)", October 2014, Patrick Ciarlet

PhD in progress :Sandrine Paolantoni, "Analyse spectrale et simulation numérique de la diffraction électromagnétique par des métamatériaux", October 2016, Christophe Hazard and Boris Gralak

PhD in progress :Laure Pesudo , "Modélisation de la réponse ultrasonore de défauts de type fissure par méthode BEM et couplage à un modèle de propagation - Application à la simulation des contrôle non destructifs", October 2014, Marc Bonnet and Stéphanie Chaillat

PhD in progress : Arnaud Recoquillay, "Identification de défauts dans un guide d'ondes en régime temporel", October 2014, Laurent Bourgeois

PhD in progress : Yohanes Tjandrawidjaja, "Modélisation de la propagation d'ondes guidées et de leur interaction avec des défauts localisés dans une plaque élastique anisotrope pour des applications en SHM", October 2016, Anne-Sophie Bonnet-Ben Dhia and Sonia Fliss

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] M. BAKRY.*Reliability and optimization of integral formulation based computations for wave propagation*, Université Paris-Saclay, October 2016, https://pastel.archives-ouvertes.fr/tel-01422039.

[2] G. BECK.*Mathematical modeling of electrical networks*, Université Paris-Saclay, March 2016, https://pastel.archives-ouvertes.fr/tel-01421740.

[3] M. CHAMAILLARD.*Effective boundary conditions for thin periodic coatings*, Université Paris Saclay, January 2016, https://hal.archives-ouvertes.fr/tel-01421937.

[4] R. CORNAGGIA. *Development and use of higher-order asymptotics to solve inverse scattering problems*, Université Paris Saclay, September 2016, https://hal.archives-ouvertes.fr/tel-01395525.

[5] E. VASILEVSKAYA.*Open periodic waveguides. Theory and computation*, Université Paris 13, July 2016, https://tel.archives-ouvertes.fr/tel-01426166.

[6] V. VINOLES.*Interface problems with metamaterials : modelling, analysis and simulations*, Université Paris-Saclay, September 2016, https://pastel.archives-ouvertes.fr/tel-01426226.

Articles in International Peer-Reviewed Journal

[7] L. BARATCHART, L. BOURGEOIS, J. LEBLOND.*Uniqueness results for inverse Robin problems with bounded coefficient*, in "Journal of Functional Analysis", 2016 [*DOI :* 10.1016/J.JFA.2016.01.011], https://hal.inria.fr/hal-01084428.

[8] V. BARONIAN, A.-S. BONNET-BENDHIA, S. FLISS, A. TONNOIR.*Iterative methods for scattering problems in isotropic or anisotropic elastic waveguides*, in "Wave Motion", 2016, https://hal.inria.fr/hal-01164794.

[9] V. BARONIAN, L. BOURGEOIS, A. RECOQUILLAY.*Imaging an acoustic waveguide from surface data in the time domain*, in "Wave Motion", 2016, vol. 66, p. 68 - 87 [*DOI :* 10.1016/J.WAVEMOTI.2016.05.006], https://hal.inria.fr/hal-01379890.

[10] M. BONNET.*A modified volume integral equation for anisotropic elastic or conducting inhomogeneities. Unconditional solvability by Neumann series*, in "Journal of Integral Equations and Applications", 2016, forthcoming, https://hal.archives-ouvertes.fr/hal-01417944.

[11] M. BONNET.*Solvability of a Volume Integral Equation Formulation for Anisotropic Elastodynamic Scattering*, in "Journal of Integral Equations and Applications", 2016, vol. 28, p. 169-203, https://hal.archives-ouvertes.fr/hal-01246135.

[12] M. BONNET, A. BUREL, M. DURUFLÉ, P. JOLY.*Effective transmission conditions for thin-layer transmission problems in elastodynamics. The case of a planar layer model*, in "ESAIM: Mathematical Modelling and Numerical Analysis", 2016, vol. 50, p. 43-75 [*DOI :* 10.1051/M2AN/2015030], https://hal.archives-ouvertes.fr/hal-01144401.

[13] A.-S. BONNET-BENDHIA, C. CARVALHO, L. CHESNEL, P. CIARLET.*On the use of Perfectly Matched Layers at corners for scattering problems with sign-changing coefficients*, in "Journal of Computational Physics", October 2016, vol. 322, p. 224-247 [*DOI :* 10.1016/J.JCP.2016.06.037], https://hal.archives-ouvertes.fr/hal-01225309.

[14] C. CALDINI-QUEIROS, B. DESPRÉS, L.-M. IMBERT-GÉRARD, M. KACHANOVSKA.*A numerical study of the solution of x-mode equations around the hybrid resonance* , in "ESAIM: Proceedings and Surveys", 2016, vol. 53, p. 1-21, https://hal.archives-ouvertes.fr/hal-01114791.

[15] J. CHABASSIER, M. DURUFLÉ, P. JOLY.*Time Domain Simulation of a Piano. Part 2 : Numerical Aspects*, in "ESAIM: Mathematical Modelling and Numerical Analysis", January 2016, vol. 50, n$^{\rm o}$ 1, p. 93-133, L'article est en cours de revision, https://hal.archives-ouvertes.fr/hal-01085477.

[16] P. CIARLET.*On the approximation of electromagnetic fields by edge finite elements. Part 1: Sharp interpolation results for low-regularity fields*, in "Computers and Mathematics with Applications", 2016 [*DOI :* 10.1016/J.CAMWA.2015.10.020], https://hal.archives-ouvertes.fr/hal-01176476.

[17] B. DELOURME, P. JOLY, E. VASILEVSKAYA.*Existence of guided waves due to a lineic perturbation of a 3D periodic medium*, in "Applied Mathematics Letters", 2016 [*DOI :* 10.1016/J.AML.2016.11.017], https://hal.archives-ouvertes.fr/hal-01412396.

[18] A. HÜPPE, G. COHEN, S. IMPERIALE, M. KALTENBACHER.*Construction and analysis of an adapted spectral finite element method to convective acoustic equations*, in "Communications in Computational Physics", June 2016 [*DOI :* 10.4208/CICP.250515.161115A], https://hal.inria.fr/hal-01363149.

[19] P. JOLY, J. RODRÍGUEZ.*Mathematical Aspects of variational boundary integral equations for time dependent wave propagation*, in "Journal of Integral Equations and Applications", 2016, https://hal.archives-ouvertes.fr/hal-01398379.

[20] K. C. MEZA FAJARDO, J.-F. SEMBLAT, S. CHAILLAT, L. LENTI.*Seismic Wave Amplification in 3D Alluvial Basins: 3D/1D Amplification Ratios from Fast Multipole BEM Simulations*, in "Bulletin of the Seismological Society of America", June 2016, vol. 106, n° 3, p. 1267-1281 [*DOI :* 10.1785/0120150159], https://hal.archives-ouvertes.fr/hal-01358724.

[21] J. F. MERCIER, A. MAUREL.*Improved multimodal method for the acoustic propagation in waveguides with a wall impedance and a uniform flow*, in "Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences", 2016, vol. 472, n° 2190 [*DOI :* 10.1098/RSPA.2016.0094], https://hal.inria.fr/hal-01396081.

[22] A. MODAVE, J. LAMBRECHTS, C. GEUZAINE.*Perfectly matched layers for convex truncated domains with discontinuous Galerkin time domain simulations*, in "Computers & Mathematics with Applications", 2017 [*DOI :* 10.1016/J.CAMWA.2016.12.027], https://hal.archives-ouvertes.fr/hal-01378501.

[23] A. TONNOIR.*Dirichlet-to-Neumann operator for diffraction problems in stratified anisotropic acoustic waveguides*, in "Comptes Rendus Mathématique", February 2016 [*DOI :* 10.1016/J.CRMA.2015.12.018], https://hal.inria.fr/hal-01195869.

[24] A. VIGNERON, E. DEMALDENT, M. BONNET.*A multi-step solution algorithm for Maxwell boundary integral equations applied to low-frequency electromagnetic testing of conductive objects*, in "IEEE Transactions on Magnetics", 2016, vol. 52, 7005208 [*DOI :* 10.1109/TMAG.2016.2584018], https://hal.archives-ouvertes.fr/hal-01336835.

Scientific Books (or Scientific Book chapters)

[25] G. COHEN, S. PERNET.*Finite Element and Discontinuous Galerkin Methods for Transient Wave Equations*, Springer, 2017 [*DOI :* 10.1007/978-94-017-7761-2], https://hal.archives-ouvertes.fr/hal-01253166.

Books or Proceedings Editing

[26] G. COHEN, S. PERNET (editors). *Finite Element and Discontinuous Galerkin Methods for Transient Wave Equations*, Scientific Computation, Springer, 2017, 381 [*DOI :* 10.1007/978-94-017-7761-2], https://hal-ensta.archives-ouvertes.fr/hal-01398287.

Research Reports

[27] B. DELOURME, S. FLISS, P. JOLY, E. VASILEVSKAYA.*Trapped modes in thin and infinite ladder like domains: existence and asymptotic analysis*, (CNRS-ENSTA Paristech-Inria, Université Paris-Saclay) ; LAGA, Université Paris 13, March 2016, n^o RR-8882, https://hal.inria.fr/hal-01287023.

Other Publications

[28] L. BARATCHART, S. CHEVILLARD, J. LEBLOND, E. A. LIMA, D. PONOMAREV.*Moments estimation of magnetic source terms from partial data*, 2016, working paper or preprint, https://hal.inria.fr/hal-01421157.

[29] A.-S. BONNET-BENDHIA, C. CARVALHO, P. CIARLET.*Mesh requirements for the finite element approximation of problems with sign-changing coefficients*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01335153.

[30] A.-S. BONNET-BENDHIA, L. CHESNEL, S. NAZAROV.*Perfect transmission invisibility for waveguides with sound hard walls*, September 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01371163.

[31] L. BOURGEOIS, J. DARDÉ.*The "exterior approach" applied to the inverse obstacle problem for the heat equation*, January 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01366850.

[32] E. BÉCACHE, P. JOLY, M. KACHANOVSKA.*Stable Perfectly Matched Layers for a Cold Plasma in a Strong Background Magnetic Field* , November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01397581.

[33] E. BÉCACHE, P. JOLY, V. VINOLES.*On the analysis of perfectly matched layers for a class of dispersive media and application to negative index metamaterials*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01327315.

[34] E. BÉCACHE, M. KACHANOVSKA.*Stable perfectly matched layers for a class of anisotropic dispersive models. PART I: Necessary and sufficient conditions of stability*, December 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01356811.

[35] C. CARVALHO, L. CHESNEL, P. CIARLET.*Eigenvalue problems with sign-changing coefficients*, November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01394856.

[36] M. CASSIER, C. HAZARD, P. JOLY.*Spectral theory for Maxwell's equations at the interface of a metamaterial. Part I: Generalized Fourier transform*, October 2016, 34 pages, 8 figures, https://hal-ensta.archives-ouvertes.fr/hal-01379118.

[37] P. CIARLET, S. FLISS, C. STOHRER.*On the Approximation of Electromagnetic Fields by Edge Finite Elements. Part 2: A Heterogeneous Multiscale Method for Maxwell's equations*, October 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01364782.

[38] P. CIARLET, E. JAMELOT, F. D. KPADONOU.*Domain decomposition methods for the diffusion equation with low-regularity solution*, July 2016, working paper or preprint, https://hal.inria.fr/hal-01349385.

[39] B. DELOURME, S. FLISS, P. JOLY, E. VASILEVSKAYA.*Trapped modes in thin and infinite ladder like domains: existence and asymptotic analysis*, March 2016, working paper or preprint, https://hal.inria.fr/hal-01287127.

[40] M. KACHANOVSKA.*Stable perfectly matched layers for a class of anisotropic dispersive models. part II: stability in corners and energy estimates* , December 2016, This work was supported by a public grant as part of the Investissement d'avenir project, reference ANR-11-LABX-0056-LMH, LabEx LMH, as well as a co-financing program PRESTIG, https://hal.inria.fr/hal-01419682.

[41] J.-F. MERCIER, B. LOMBARD.*A two-way model for nonlinear acoustic waves in a non-uniform lattice of Helmholtz resonators*, September 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01312801.

[42] A. MODAVE, A. ATLE, J. CHAN, T. WARBURTON.*High-order absorbing boundary conditions with corner/edge compatibility for GPU-accelerated discontinuous Galerkin wave simulations*, October 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01383074.

Project-Team SELECT

Model selection in statistical learning

IN COLLABORATION WITH: Laboratoire de mathématiques d'Orsay de l'Université de Paris-Sud (LMO)

IN PARTNERSHIP WITH:
CNRS

Université Paris-Sud (Paris 11)

RESEARCH CENTER
Saclay - Île-de-France

THEME
Optimization, machine learning and statistical methods

Table of contents

<center>**Project-Team SELECT**</center>

Creation of the Project-Team: 2007 January 01

Keywords:

Computer Science and Digital Science:
- 3.1.1. - Modeling, representation
- 3.1.8. - Big data (production, storage, transfer)
- 3.2.2. - Knowledge extraction, cleaning
- 3.3.2. - Data mining
- 3.3.3. - Big data analysis
- 3.4.1. - Supervised learning
- 3.4.2. - Unsupervised learning
- 3.4.3. - Reinforcement learning
- 3.4.4. - Optimization and learning
- 3.4.5. - Bayesian methods
- 3.4.6. - Neural networks
- 3.4.7. - Kernel methods
- 3.4.8. - Deep learning
- 5.3.3. - Pattern recognition
- 6.2.4. - Statistical methods
- 6.2.6. - Optimization

Other Research Topics and Application Domains:
- 1.1.5. - Genetics
- 1.1.6. - Genomics
- 1.1.9. - Bioinformatics
- 1.1.10. - Mathematical biology
- 9.4.2. - Mathematics

1. Members

Research Scientists
Kevin Bleakley [Inria, Researcher]
Gilles Celeux [Inria, Senior Researcher]
Julie Josse [Inria, Researcher]
Matthieu Lerasle [CNRS, Researcher]

Faculty Members
Pascal Massart [Team leader, Univ. Paris XI, Professor]
Sylvain Arlot [Univ. Paris XI, Professor]
Christine Keribin [Univ. Paris XI, Associate Professor]
Claire Lacour [Univ. Paris XI, Associate Professor]
Patrick Pamphile [Univ. Paris XI, Associate Professor]
Jean-Michel Poggi [Univ. Paris V, Professor]

Technical Staff

Benjamin Auder [CNRS, Researcher]
Josselin Demont [Inria]
Jonas Renault [Inria]
Christian Poli [Inria]

PhD Students
Benjamin Goehry [Univ. Paris XI]
Neska Haouij [Univ. Paris XI and Univ. Tunis El Manar]
Valerie Robert [Univ. Paris XI]
Yann Vasseur [Univ. Paris XI]
Jeanne Nguyen [Univ. Paris XI]
Florence Ducros [Cifre: Univ. Paris XI and Nexter]
Claire Brécheteau [Univ. Paris XI]
Eddie Aamari [Univ. Paris XI]
Damien Garreau [Univ. Paris XI]
Guillaume Maillard [Univ. Paris XI]

Post-Doctoral Fellow
Kaniav Kamary [Inria]

Administrative Assistant
Olga Mwana Mobulakani [Inria]

2. Overall Objectives

2.1. Model selection in Statistics

The research domain for the SELECT project is statistics. Statistical methodology has made great progress over the past few decades, with a variety of statistical learning software packages that support many different methods and algorithms. Users now face the problem of choosing among them, to select the most appropriate method for their data sets and objectives. The problem of model selection is an important but difficult problem, both theoretically and practically. Classical model selection criteria, which use penalized minimum-contrast criteria with fixed penalties, are often based on unrealistic assumptions.

SELECT aims to provide efficient model selection criteria with data-driven penalty terms. In this context, SELECT aims to improve the toolkit of statistical model selection criteria from both theoretical and practical perspectives. Currently, SELECT is focusing its effort on variable selection in statistical learning, hidden-structure models and supervised classification. Its domains of application concern reliability, curve classification, phylogenetic analysis and classification in genetics. New developments in SELECT activities are concerned with applications in biostatistics (statistical analysis of medical images) and biology.

3. Research Program

3.1. General presentation

From applications we treat on a day-to-day basis, we have learned that some assumptions currently used in asymptotic theory for model selection are often irrelevant in practice. For instance, it is not realistic to assume that the target belongs to the family of models in competition. Moreover, in many situations, it is useful to make the size of the model depend on the sample size, which makes asymptotic analyses breakdown. An important aim of SELECT is to propose model selection criteria which take such practical constraints into account.

3.2. A nonasymptotic view of model selection

An important goal of SELECT is to build and analyze penalized log-likelihood model selection criteria that are efficient when the number of models in competition grows to infinity with the number of observations. Concentration inequalities are a key tool for this, and lead to data-driven penalty choice strategies. A major research direction for SELECT consists of deepening the analysis of data-driven penalties, both from the theoretical and practical points of view. There is no universal way of calibrating penalties, but there are several different general ideas that we aim to develop, including heuristics derived from Gaussian theory, special strategies for variable selection, and resampling methods.

3.3. Taking into account the modeling purpose in model selection

Choosing a model is not only difficult theoretically. From a practical point of view, it is important to design model selection criteria that accommodate situations in which the data probability distribution P is unknown, and which take the model user's purpose into account. Most standard model selection criteria assume that P belongs to one of a set of models, without considering the purpose of the model. By also considering the model user's purpose, we can avoid or overcome certain theoretical difficulties, and produce flexible model selection criteria with data-driven penalties. The latter is useful in supervised classification and hidden-structure models.

3.4. Bayesian model selection

The Bayesian approach to statistical problems is fundamentally probabilistic: a joint probability distribution is used to describe the relationships among all unknowns and the data. Inference is then based on the posterior distribution, i.e., the conditional probability distribution of the parameters given the observed data. Exploiting the internal consistency of the probability framework, the posterior distribution extracts relevant information in the data and provides a complete and coherent summary of post-data uncertainty. Using the posterior to solve specific inference and decision problems is then straightforward, at least in principle.

4. Application Domains

4.1. Introduction

A key goal of SELECT is to produce methodological contributions in statistics. For this reason, the SELECT team works with applications that serve as an important source of interesting practical problems and require innovative methodology to address them. Many of our applications involve contracts with industrial partners, e.g., in reliability, although we also have several academic collaborations, e.g., in genetics and image analysis.

4.2. Curve classification

The field of classification for complex data such as curves, functions, spectra and time series, is an important problem in current research. Standard data analysis questions are being looked into anew, in order to define novel strategies that take the functional nature of such data into account. Functional data analysis addresses a variety of applied problems, including longitudinal studies, analysis of fMRI data, and spectral calibration.

We are focused in particular on unsupervised classification. In addition to standard questions such as the choice of the number of clusters, the norm for measuring the distance between two observations, and vectors for representing clusters, we must also address a major computational problem: the functional nature of the data, which requires new approaches.

4.3. Computer experiments and reliability

For several years now, SELECT has collaborated with the EDF-DER *Maintenance des Risques Industriels* group. One important theme involves the resolution of inverse problems using simulation tools to analyze incertainty in highly complex physical systems.

The other major theme concerns reliability, through a research collaboration with Nexter involving a Cifre convention. This collaboration concerns a lifetime analysis of a vehicle fleet to assess aging.

Moreover, a collaboration has begun with Dassault Aviation on the modal analysis of mechanical structures, which aims to identify the vibration behavior of structures under dynamic excitation. From the algorithmic point of view, modal analysis amounts to estimation in parametric models on the basis of measured excitations and structural response data. In literature and existing implementations, the model selection problem associated with this estimation is currently treated by a rather weighty and heuristic procedure. In the context of our own research, model selection via penalization methods are to be tested on this model selection problem.

4.4. Analysis of genomic data

For many years now, SELECT collaborates with Marie-Laure Martin-Magniette (URGV) for the analysis of genomic data. An important theme of this collaboration is using statistically sound model-based clustering methods to discover groups of co-expressed genes from microarray and high-throughput sequencing data. In particular, identifying biological entities that share similar profiles across several treatment conditions, such as co-expressed genes, may help identify groups of genes that are involved in the same biological processes.

Yann Vasseur is completing a thesis co-supervised by Gilles Celeux and Marie-Laure Martin-Magniette on this topic, which is also an interesting investigation domain for the latent block model developed by SELECT. For this work, Yann Vasseur is dealing with high-dimensional ill-posed problems where the number of variable is almost equal to the number of observations. He has designed heuristic tools using regularized regression methods to circumvent this difficulty.

SELECT collaborates with Anavaj Sakuntabhai and Benno Schwikowski (Pasteur Institute) on prediction of dengue fever severity from high-dimensional gene expression data. One project involves using/finding new and computationally efficient methods (e.g., 2d isotonic regression, lasso regression) for predicting dengue severity. Due to the high-dimensional nature of the data and low-dimensional nature of the number of individuals, false discovery rate (FDR) methods are used to provide statistical justification of results. A second project aims to predict dengue severity using only low-dimensional clinical data obtained at hospital arrival. A third project involves statistical meta-analysis of newly collected dengue gene expression data along with recently published data sets from other groups.

SELECT is involved in the ANR "jeunes chercheurs" MixStatSeq directed by Cathy Maugis (INSA Toulouse), which is concerned with statistical analysis and clustering of RNASeq genomics data.

4.5. Pharmacovigilance

A collaboration is ongoing with Pascale Tubert-Bitter, Ismael Ahmed and Mohamed Sedki (Pharmacoepidemiology and Infectious Diseases, PhEMI) for the analysis of pharmacovigilance data. In this framework, the goal is to detect, as soon as possible, potential associations between certain drugs and adverse effects, which appeared after the authorized marketing of these drugs. Instead of working on aggregate data (contingency table) like is usually the case, the approach developed aims to deal with individual's data, which perhaps gives more information. Valerie Robert is completing a thesis co-supervised by Gilles Celeux and Christine Keribin on this topic, which involves the development of a new model-based clustering method, inspired by latent block models. Morever, she has defined new tools to estimate and assess the block clustering involved in these models.

4.6. Spectroscopic imaging analysis of ancient materials

Ancient materials, encountered in archaeology and paleontology are often complex, heterogeneous and poorly characterized before physico-chemical analysis. A popular technique to gather as much physico-chemical information as possible, is spectro-microscopy or spectral imaging, where a full spectra, made of more than a thousand samples, is measured for each pixel. The produced data is tensorial with two or three spatial dimensions and one or more spectral dimensions, and requires the combination of an "image" approach

with a "curve analysis" approach. Since 2010 SELECT, collaborates with Serge Cohen (IPANEMA) on the development of conditional density estimation through GMM, and non-asymptotic model selection, to perform stochastic segmentation of such tensorial datasets. This technique enables the simultaneous accounting for spatial and spectral information, while producing statistically sound information on morphological and physico-chemical aspects of the studied samples.

5. New Software and Platforms

5.1. BlockCluster

Block Clustering
KEYWORDS: Mixture model - Block cluster analysis
SCIENTIFIC DESCRIPTION

Blockcluster is software devoted to model-based block clustering. It is developed in partnership with the MODAL team (Inria Lille). This year, some major bugs have been fixed, and the Bayesian point of view has been reinforced by including Gibbs sampling for binary and categorial data. This Gibbs sampler, coupled with the variational Bayes algorithm, provides solutions which are more stable and less dependent on the initial values of the algorithm. An exact expression of the ICL criterion has also been provided. This non-asymptotic criterion appears to be more relevant than the BIC-like approximation of ICL.

FUNCTIONAL DESCRIPTION

BlockCluster is an R package for co-clustering of binary, contingency and continuous data based on mixture models.

- Participants: Gilles Celeux, Christine Keribin, Christophe Biernacki and Serge Iovleff
- Contact: Gilles Celeux
- URL: http://cran.r-project.org/web/packages/blockcluster/index.html

5.2. Mixmod

Multi-purpose software for model-based clustering and classification with continuous and categorical variables.
KEYWORDS: Mixture model - cluster analysis - discriminant analysis
FUNCTIONAL DESCRIPTION

MIXMOD is being developed in collaboration with Christophe Biernacki, Florent Langrognet (Université de Franche-Comté) and Gérard Govaert (Université de Technologie de Compiègne). MIXMOD (MIXture MODeling) software fits mixture models to a given data set, with either a clustering or a discriminant analysis purpose. MIXMOD uses a large variety of algorithms to estimate mixture parameters, e.g., EM, Classification EM, and Stochastic EM. They can be combined to create different strategies that lead to a sensible maximum of the likelihood (or completed likelihood) function. Moreover, different information criteria for choosing a parsimonious model, e.g. the number of mixture components, some of them favoring either a cluster analysis or a discriminant analysis point of view, are included. Many Gaussian models for continuous variables and multinomial models for discrete variable are included. Written in C++, MIXMOD is interfaced with MATLAB. The software, statistical documentation, and user guide are available here: http://www.mixmod.org.

Since 2010, MIXMOD has a proper graphical user interface. A version of MIXMOD in R is now available: http://cran.r-project.org/web/packages/Rmixmod/index.html.

Josselin Demont and Benjamin Auder have contributed to software improvement in MIXMOD. They have implemented an interface to test any mathematical library (Armadillo, Eigen, etc.) to replace NEWMAT. They have contributed to the continuous integration setup using Jenkins tools, and prepared an automated testing framework for unit and non-regression tests.

Jonas Renault, an engineer, is in charge of developing a web version of MIXMOD.

- Participants: Christophe Biernacki, Gilles Celeux, Gérard Govaert, Florent Langrognet and Benjamin Auder
- Partners: CNRS - HEUDIASYC - Laboratoire Paul Painlevé - LIFL - LMB - Université Lille 1
- Contact: Gilles Celeux
- URL: http://www.mixmod.org

5.3. MASSICCC

Massive Clustering with Cloud Computing
KEYWORDS: Statistic analysis - Big data - Machine learning - Web Application
SCIENTIFIC DESCRIPTION

The web application let users use several software packages developed by Inria directly in a web browser. Mixmod is a classification library for continuous and categorical data. MixtComp allows for missing data and a larger choice of data types developed by MODAL team (Inria Lille). BlockCluster is a library for co-clustering data. When using the web application, the user can first upload a data set, then configure a job using one of the libraries mentioned and start the execution of the job on a cluster. The results are then displayed directly in the browser allowing for rapid understanding and interactive visualisation.
FUNCTIONAL DESCRIPTION

The MASSICCC web application offers a simple and dynamic interface for analysing heterogeneous data with a web browser. Various software packages for statistical analysis are available (Mixmod, MixtComp, BlockCluster) which allow for supervised and supervised classification of large data sets.

- Participants: Christophe Biernacki, Gilles Celeux, Benjamin Auder, Josselein Demont, Jonas Renault
- Contact: Jonas Renault
- URL: https://massiccc.lille.inria.fr

6. New Results

6.1. Model selection in Regression and Classification

Participants: Gilles Celeux, Serge Cohen, Pascal Massart, Sylvain Arlot, Jean-Michel Poggi, Kevin Bleakley.

The well-documented and consistent variable selection procedure in model-based cluster analysis and classification that Cathy Maugis (INSA Toulouse) designed during her PhD thesis in SELECT, makes use of stepwise algorithms which are painfully slow in high dimensions. In order to circumvent this drawback, Gilles Celeux, in collaboration with Mohammed Sedki (Université Paris XI) and Cathy Maugis), have proposed to sort variables using a lasso-like penalization adapted to the Gaussian mixture model context. Using this ranking to select variables, they avoid the combinatory problem of stepwise procedures. The performances on challenging simulated and real data sets are similar to the standard procedure, with a CPU time divided by a factor of more than a hundred.

In collaboration with Jean-Michel Marin (Université de Montpellier) and Olivier Gascuel (LIRMM), Gilles Celeux has continued research aiming to select a short list of models rather a single model. This short list 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. Furthermore, the Kullback-Leibler distances at hand are estimated through nonparametric and parametric bootstrap procedures. Different strategies are compared through numerical experiments on simulated and real data sets. This year their method has been compared favorably to competing methods.

Sylvain Arlot, in collaboration with Damien Garreau (Inria Paris, Sierra team), studied the kernel change-point algorithm (KCP) proposed by Arlot, Celisse and Harchaoui, that aims at locating an unknown number of change-points in the distribution of a sequence of independent data taking values in an arbitrary set. The change-points are selected by model selection with a penalized kernel empirical criterion. They provide a non-asymptotic result showing that, with high probability, the KCP procedure retrieves the correct number of change-points, provided that the constant in the penalty is well-chosen; in addition, KCP estimates the change-points location at the minimax rate $\log(n)/n$. As a consequence, when using a characteristic kernel, KCP detects all kinds of change in the distribution (not only changes in the mean or the variance), and it is able to do so for complex structured data (not necessarily in $\mathbb{R}^d$). Most of the analysis is conducted assuming that the kernel is bounded; part of the results can be extended when we only assume a finite second-order moment.

Emilie Devijver, Yannig Goude and Jean-Michel Poggi have proposed a new methodology for customer segmentation, in the context of load profiles in energy consumption. The method is based on high-dimensional regression models which perform clustering and model selection at the same time. They have focused on uncovering classes corresponding to different regression models, and compute clustering and model identification in each cluster simultaneously. They have shown the feasibility of the approach on a real data set of Irish customers. Benjamin Goehry is completing a thesis co-supervised by P. Massart and J-M. Poggi, aiming at extending this scheme by introducing the use of time series forecasting models adapted to each cluster.

J-M. Poggi, with J. Cugliari, Y. Goude, have proposed building clustering tools useful for forecasting load consumption. The idea is to disaggregate the global signal in such a way that the sum of disaggregated forecasts significantly improves the prediction of the whole global signal. The strategy has three steps: first they cluster curves defining super-consumers, then they build a hierarchy of partitions from which the best one is selected with respect to a disaggregated forecast criterion. The proposed strategy is applied to a dataset of individual consumers from the French electricity provider EDF.

V. Thouvenot and J-M. Poggi, with A. Pichavant, A. Antoniadis, Y. Goude, consider electricity forecasting using multi-stage estimators of nonlinear additive models. An automatic procedure for variable selection is used to correct middle term forecasting errors for short term forecasting. An application to the EDF customer load demand at an aggregate level is considered as well as an application on load demand from the GEFCom 2012 competition; this is a local application.

6.2. Estimator selection

Participants: Claire Lacour, Pascal Massart.

Estimator selection has become a crucial issue in nonparametric estimation. Two widely used methods are penalized empirical risk minimization (such as penalized log-likelihood estimation) and pairwise comparison (such as Lepski's method). C. Lacour, P. Massart and V. Rivoirard have developed a new method for bandwidth selection which is in some sense intermediate between these two main methods mentioned above, and is called "Penalized Comparison to Overfitting". They have first provided some theoretical results (oracle bounds, minimal penalty) within the framework of kernel density estimation, which leads to some fully data-driven selection strategies. Currently, S. Varet is implementing this method, making a thorough comparison with other selection methods, and tackling the multivariate case. Theoretical work is also in progress, in order to expand the method to other loss functions, such as the Hellinger loss.

6.3. Statistical learning methodology and theory

Participants: Gilles Celeux, Christine Keribin, Michel Prenat, Kaniav Kamary, Sylvain Arlot, Benjamin Auder, Jean-Michel Poggi, Neska El Haouij, Kevin Bleakley.

Gaussian graphical models are widely used to infer and visualize networks of dependencies between continuous variables. However, inferring the graph is difficult when the sample size is small compared to the number of variables. To reduce the number of parameters to estimate in the model, the past PhD. students Emilie Devijver (supervisors: Pascal Massart and Jean-Michel Poggi) and Mélina Gallopin (supervisor: Gilles Celeux)

proposed a non-asymptotic model selection procedure supported by strong theoretical guarantees based on an oracle inequality and a minimax lower bound. The covariance matrix of the model is approximated by a block-diagonal matrix. The structure of this matrix is detected by thresholding the sample covariance matrix, where the threshold is selected using the slope heuristic. Based on the block-diagonal structure of the covariance matrix, the estimation problem is divided into several independent problems: subsequently, the network of dependencies between variables is inferred using the graphical lasso algorithm in each block. The performance of the procedure has been illustrated on simulated data. An application to a real gene expression dataset with a limited sample size has been achieved: the dimension reduction allows attention to be objectively focused on interactions among smaller subsets of genes, leading to a more parsimonious and interpretable modular network. This work has been accepted for publication in the *Journal of the American Statistical Association*.

J-M. Poggi, with A. Bar-Hen, have focused on individual observation diagnosis issues for graphical models. The use of an influence measure is a classical diagnostic method to measure the perturbation induced by single elements. The stability issue is here considered using jackknife. For a given graphical model, tools to perform diagnosis on observations are provided. In the second step, a filtering of the dataset to obtain a stable network is proposed.

Latent Block Models (LBM) are a model-based method to cluster simultaneously the d columns and n rows of a data matrix. The Blockcluster package estimates such LBMs. Parameter estimation in LBM is a difficult and multifaceted problem. Although various estimation strategies have been proposed and are now well-understood empirically, theoretical guarantees about their asymptotic behavior is rather rare. Christine Keribin, in collaboration with Mahendra Mariadassou (INRA) and Vincent Brault (Université de Grenoble) have shown that under some mild conditions on the parameter space, and in an asymptotic regime where $\log(d)/n$ and $\log(n)/d$ go to 0 when n and d go to $+\infty$, (1) the maximum likelihood estimate of the complete model (with known labels) is consistent and (2) the log-likelihood ratios are equivalent under the complete and observed (with unknown labels) models. This equivalence allows us to transfer the asymptotic consistency to the maximum likelihood estimate under the observed model. Moreover, the variational estimator is also consistent. These results extends the results of Bickel et al. (2013) on stochastic block models, and detail the case where the parameter exhibits symmetry.

For the same LBM, Valérie Robert and Yann Vasseur have extended the popular Adjusted Rand Index (ARI) to the task of simultaneous clustering of the rows and columns of a given matrix. This new index, called the Coclustering Adjusted Rand Index (CARI), overcomes the label switching phenomenon while remaining useful and competitive with respect to other indices. Indeed, partitions with high numbers of clusters can be considered, and no convention is required when the numbers of clusters in partitions are different. They are now exploring links with other indices.

Gilles Celeux continued his collaboration with Jean-Patrick Baudry on model-based clustering. This year, they proposed to consider the model selection criterion ICL as a validity index. They show how it can be coupled with a null model of homogeneity focusing on clustering. This null model, which includes the Gaussian distributions, can be difficult to analyze. They find an explicit representation for simple models and show how the parametric bootstrap test can be applied in such situations. In more general situations, they propose a solution for applying this approach involving an "acceptance-rejection" procedure which explores the parameter space to approximate the maximum likelihood estimator inside the null model of homogeneity. The uncovering of this null model highlights the notion of class underlying ICL, and confirms the results of earlier results which show that ICL is consistent for a loss function taking clustering into account.

In collaboration with Arthur White and Jason Wyse (Trinity College, Dublin) Gilles Celeux has evaluated for multivariate Poisson mixture models the performance of a greedy search method compared to the expectation maximization (EM) algorithm, to optimize the ICL model selection criterion, which can be computed exactly for such models. It appears that EM gives often slightly better results, but the greedy search is computationally is more efficient.

The Dutch and French schools of data analysis differ in their approaches to the question: How does one understand and summarize the information contained in a data set? Julie Josse, in collaboration with François Husson (Agro Rennes) and Gibert Saporta (CNAM, Paris), explored the shared factors and differences between

the schools, with a focus on methods dedicated to the analysis of categorical data, which are known either as homogeneity analysis (HOMALS) or multiple correspondence analysis (MCA). MCA is a dimension-reduction method which plays a large role in the analysis of tables with categorical nominal variables such as survey data. Though it is usually motivated and derived using geometric considerations, they proved that it amounts to a single proximal Newton step of a natural bilinear exponential family model for categorical data: the multinomial logit bilinear model. They compared and contrasted the behavior of MCA with that of the model on simulations, and discussed new insights into the properties of both exploratory multivariate methods and their cognate models. The main conclusion is to recommend approximating the multilogit model parameters using MCA. Indeed, estimating the parameters of the model is not a trivial task, whereas MCA has the great advantage of being easily solved by a singular value decomposition, as well as being scalable to large datasets.

Julie Josse, with Sobczyk and Bogdan, have discussed the problem of estimating the number of principal components in Principal Components Analysis (PCA). They address this issue by presenting an approximate Bayesian approach based on Laplace approximation, and introduce a general method for building the model selection criteria, called PEnalized SEmi-integrated Likelihood (PESEL). This general framework encompasses a variety of existing approaches based on probabilistic models, like e.g., Bayesian Information Criterion for the Probabilistic PCA (PPCA), and allows for construction of new criteria, depending on the size of the data set at hand. Specifically, they define PESEL when the number of variables substantially exceeds the number of observations. Numerical simulations show that PESEL-based criteria can be quite robust against deviations from probabilistic model assumptions. Selected PESEL-based criteria for estimation of the number of principal components are implemented in the R package varclust, which is available on Github.

Gillies Celeux and Julie Josse have started research on missing data for model-based clustering in collaboration with Christophe Biernacki (Modal, Inria Lille). The aim of this research is to propose appropriate and efficient tools for the packages Mixmod and Mixtcomp.

In collaboration with Jean-Michel Marin (Université de Montpellier) and Christian Robert (Université Paris 9-Dauphine), Gilles Celeux and Kaniav Kamary investigated the ability of Bayesian inference to properly estimate the parameters of Gaussian mixtures in high dimensions. Their study shows how the choice of the prior distributions is important. In particular, independent prior distributions give much better performances. Moreover, when the dimension d becomes very large (say $d > 40$) Bayesian inference becomes questionable. The results of this study will be gathered in a chapter of a book on mixture models that Gilles Celeux is preparing with Christian Robert and Sylvia Fruhwirth Schnatter.

Sylvain Arlot, in collaboration with Robin Genuer (ISPED), studied the reasons why random forests work so well in practice. Focusing on the problem of quantifying the impact of each ingredient of random forests on their performance, they showed that such a quantification is possible for a simple pure forest, leading to conclusions that could apply more generally. Then, they considered "hold-out" random forests, which are a good midpoint between "toy" pure forests and Breiman's original random forests.

J.-M. Poggi and N. El Haouij (with R. Ghozi, S. Sevestre Ghalila and M. Jaïdane) provide a random forest-based method for the selection of physiological functional variables in order to classify stress levels during a real-world driving experience. The contribution of this study is twofold: on the methodological side, it considers physiological signals as functional variables and offers a procedure for data processing and variable selection. On the applied side, the proposed method provides a "blind" procedure of driver's stress level classification that does not depend on expert-based studies of physiological signals.

J-M. Poggi (with R. Genuer, C. Tuleau-Malot, N. Villa-Vialaneix), have focused on random forests in Big Data classification problems, and have performed a review of available proposals about random forests in parallel environments as well as on online random forests. Three variants involving subsampling, Big Data-bootstrap and MapReduce respectively are tested on two massive datasets, one simulated one, and the other, real-world data.

B. Auder and J-M. Poggi (with M. Bobbia, B. Portier) have tested some methods for sequential aggregation for forecasting PM10 concentrations for the next day, in the context of air quality monitoring in Normandy

(France). The main originality is that the set of experts contains at the same time statistical models built by means of various methods and groups of predictors, as well as experts coming from deterministic chemical models of prediction. The obtained results show that such a strategy clearly improves the performances of the best expert both in terms of prediction errors and in terms of alerts. What is more, it obtains, for the non-convex weighting strategy, the "unbiasedness" of observed-forecasted scatterplots, which is extremely difficult to obtain.

J-M. Poggi (with A. Antoniadis, I. Gijbels, S. Lambert-Lacroix) have considered the joint estimation and variable selection for mean and dispersion in proper dispersion models. They used recent results on Bregman divergence for establishing theoretical results for the proposed estimators in fairly general settings, and also studied variable selection when there is a large number of covariates, with this number possibly tending to infinity with the sample size. The proposed estimation and selection procedure is investigated via a simulation study, and illustrated via some real data applications.

6.4. Estimation for conditional densities in high dimension

Participants: Claire Lacour, Jeanne Nguyen.

Jeanne Nguyen is working on estimation for conditional densities in high dimension. Much more informative than the regression function, conditional densities are of high interest in recent methods, particularly in the Bayesian framework (studying the posterior distribution). Considering a specific family of kernel estimators, she is studying a greedy algorithm for selecting the bandwidth. Her method addresses several issues: avoiding the curse of high dimensionality under some suitably defined sparsity conditions, being computationally efficient using iterative procedures, and early variable selection, providing theoretical guarantees on the minimax risk.

6.5. Reliability

Participants: Gilles Celeux, Florence Ducros, Patrick Pamphile.

Since June 2015, in the framework of a CIFRE convention with Nexter, Florence Ducros has begun a thesis on the modeling of aging of vehicles, supervised by Gilles Celeux and Patrick Pamphile. This thesis should lead to designing an efficient maintenance strategy according to vehicle use profiles. It involves the estimation of mixtures and competing risk models in a highly-censored setting. Moreover, she can deduce from these models operational tools to estimate the number of spare parts to be stocked in a given period. These tools are defined to take vehicle use patterns into account.

6.6. Statistical analysis of genomic data

Participants: Gilles Celeux, Mélina Gallopin, Christine Keribin, Yann Vasseur, Kevin Bleakley.

The subject of Yann Vasseur's PhD Thesis, supervised by Gilles Celeux and Marie-Laure Martin-Magniette (INRA URGV), is the inference of a regulatory network for Transcriptions Factors (TFs), which are specific genes, of *Arabidopsis thaliana*. For this, a transcriptome dataset with a similar number of TFs and statistical units is available. The first aim consists of reducing the dimension of the network to avoid high-dimensional difficulties. Representing this network with a Gaussian graphical model, the following procedure has been defined:

1. *Selection step*: choose the set of TF regulators (supports) of each TF.
2. *Classification step*: deduce co-factor groups (TFs with similar expression levels) from these supports.

Thus, the reduced network would be built on the co-factor groups. Currently, several selection methods based on Gauss-LASSO and resampling procedures have been applied to the dataset. The study of stability and parameter calibration of these methods is in progress. The TFs are clustered with the Latent Block Model into a number of co-factor groups, selected with BIC or the exact ICL criterion. Since these models are built in an ad hoc way, Yann Vasseur has defined complex simulation tools to asses their performances in a proper way.

In a collaboration with Marie-Laure Martin-Magniette, Cathy Maugis and Andrea Rau, Gilles Celeux has studied gene expression obtained from high-throughput sequencing technology. The focus is on the question of clustering gene expression profiles as a means to discover groups of co-expressed genes. A Poisson mixture model is proposed, using a rigorous framework for parameter estimation, as well as for the choice of the appropriate number of clusters. They illustrate co-expression analyses using this approach on two real RNA-seq datasets. A set of simulation studies also compares the performance of the proposed model with that of several related approaches developed to cluster RNA-seq and serial analysis of gene expression data. The proposed method is implemented in the open-source R package `HTSCluster`, available on CRAN. It can now be compared with Gaussian mixtures obtained after relevant data transformations. Moreover, the performance of `HTSCluster` is compared with $k$means-like algorithms using the χ^2 distance.

In collaboration with Benno Schwikowski, Iryna Nikolayeva and A Anavaj Sakuntabhai (Pasteur Institute, Paris), Kevin Bleakley works on using 2-d isotonic regression to predict dengue fever severity at hospital arrival using high-dimensional microarray gene expression data. Important marker genes for dengue severity have been detected, some of which now have been validated in external lab trials.

6.7. Model based-clustering for pharmacovigilance data

Participants: Gilles Celeux, Christine Keribin, Valérie Robert.

In collaboration with Pascale Tubert-Bitter, Ismael Ahmed and Mohamed Sedki, Gilles Celeux and Christine Keribin have started research concerning the detection of associations between drugs and adverse events in the framework of the PhD of Valerie Robert. At first, this team developed model-based clustering inspired by latent block models, which consists of co-clustering rows and columns of two binary tables, imposing the same row ranking. This enables it to highlight subgroups of individuals sharing the same drug profile, and subgroups of adverse effects and drugs with strong interactions. Furthermore, some sufficient conditions are provided to obtain identifiability of the model, and some results are shown for simulated data. The exact ICL criterion has been extended to this double block latent model. Through computer experiments, Valérie Robert has demonstrated the interest of the proposed model, compared with standard contingency table analysis, to detect co-prescription and masking effects.

6.8. Statistical rating and ranking of scientific journals

Participants: Gilles Celeux, Julie Josse, Simon Grah.

In collaboration with Jean-Louis Foulley (université of Montpellier), Gilles Celeux and Julie Josse have started research on statistical rating and ranking of scientific journals. This research was the subject of the internship of Simon Grah (Université Paris-Sud). Simon Grah compared many models on a set of 47 statistical journals. His study showed that the Row-Column (RC) models appears to be the most relevant. In the future, Bayesian inference for different approaches, including PageRank, will be considered.

7. Bilateral Contracts and Grants with Industry

7.1. Contract with SNECMA

Participants: Gilles Celeux, Florence Ducros, Patrick Pamphile.

SELECT has a contract with Nexter regarding modeling the reliability of vehicles.

8. Partnerships and Cooperations

8.1. Regional Initiatives

Gilles Celeux and Christine Keribin have a collaboration with the Pharmacoepidemiology and Infectious Diseases (PhEMI, INSERM) groups.

Christine Keribin is treasurer of the Société Française de Statistique (SFdS).

Sylvain Arlot and Pascal Massart co-organize a working group at ENS (Ulm) on statistical learning.

8.2. National Initiatives

8.2.1. ANR

SELECT is part of the ANR funded MixStatSeq.

8.3. International Initiatives

Gilles Celeux is one of the co-organizers of the international working group on model-based clustering. This year this workshop took place in Paris.

Julie Josse was chair of useR!2016, Stanford, CA, USA, July 2016. http://user2016.org/

Juile Josse is member of the R foundation.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

Sylvain Arlot:

- Sylvain Arlot organized a workshop at IHES on statistics and learning in the Paris-Saclay area.
- Sylvain Arlot co-organized (with Francis Bach, Inria Paris, and Alain Celisse, Univ. Lille 1) a 2-day workshop at IHES about computational and statistical trade-offs in learning.

Jean-Michel Poggi:

- Organization Special Invited Session entitled "Advances in Random Forests", 22nd International Conference on Computational Statistics, Oviedo, Spain, 23-26 August 2016.

9.1.1.2. Member of the Organizing Committees

Gilles Celeux is one of the co-organizers of the international working group on model-based clustering. This year the workshop took place in Paris.

Sylvain Arlot co-organized the 1st Junior Conference on Data Science and Engineering at Paris-Saclay (at LAL, Orsay).

9.1.2. Scientific Events Selection

9.1.2.1. Chair of Conference Program Committees

Jean-Michel Poggi was:

- President of the Scientific Programme Committee, ENBIS 2017, Naples, 10-14 June 2017

9.1.2.2. Member of the Conference Program Committees

Jean-Michel Poggi was:

- Member of the Scientific Committee of CESS 2016, Conference of European Statistics Stakeholders, Hungarian Academy of Sciences, Budapest, 20-21 October 2016
- Member of SPC ENBIS-2016, Sheffield, UK, 11-15 September 2016
- Member of the Scientific committee of the journées MAS 2016, Grenoble
- Member of SPC COMPSTAT 2016, 22nd International Conference on Computational Statistics, Oviedo, Spain, 23-26 August 2016.

9.1.3. *Journal*

9.1.3.1. Member of the Editorial Boards

Gilles Celeux is Editor-in-Chief of the *Journal de la SFdS*. He is Associate Editor of *Statistics and Computing*, *CSBIGS*.

Pascal Massart is Associate Editor of *Annals of Statistics*, *Confluentes Mathematici*, and *Foundations and Trends in Machine Learning*.

Jean-Michel Poggi is Associate Editor of *Journal of Statistical Software*, *Journal de la SFdS* and *CSBIGS*.

9.1.3.2. Reviewer - Reviewing Activities

The members of the team have reviewed numerous papers for numerous international journals.

9.1.4. *Invited Talks*

The members of the team have given many invited talks on their research in the course of 2016.

9.1.5. *Leadership within the Scientific Community*

Jean-Michel Poggi is:

- Vice-President ENBIS (European Network for Business and Industrial Statistics), 2015-18
- Vice-President FENStatS (Federation of European National Statistical Societies) since 2012
- Council Member of the ISI (2015-19)
- Member of the Board of Directors of the ERS of IASC (since 2014)

9.1.6. *Scientific Expertise*

Jean-Michel Poggi is member of the EMS Committee for Applied Mathematics (since 2014).

9.1.7. *Research Administration*

Jean-Michel Poggi is the president of ECAS (European Courses in Advanced Statistics) since 2015.

Sylvain Arlot coordinates (jointly with Marc Schoenauer, Inria Saclay) the math-STIC program of the Labex Mathématique Hadamard.

9.2. Teaching - Supervision - Juries

9.2.1. *Teaching*

SELECT members teach various courses at several different universities, and in particular the Master 2 "Mathématique de l'aléatoire" of Université Paris-Saclay.

9.2.2. *Supervision*

PhD in progress: Valérie Robert, 2013, Gilles Celeux and Christine Keribin

PhD in progress: Yann Vasseur, 2013, Gilles Celeux and Marie-Laure Martin-Magniette (URGV)

PhD in progress: Neska El Haouij, 2014, Jean-Michel Poggi and Meriem Jaïdane, Raja Ghozi (ENIT Tunisie) and Sylvie Sevestre-Ghalila (CEA LinkLab), Thesis ENITUPS

PhD in progress: Florence Ducros, 2015, Gilles Celeux and Patrick Pamphile

PhD in progress: Claire Brécheteau, 2015, Pascal Massart

PhD in progress: Eddie Aamari, 2015, Pascal Massart and Frédéric Chazal

PhD in progress: Damien Garreau, 2016, Sylvain Arlot and Gérard Biau

PhD in progress: Guillaume Maillard, 2016, Sylvain Arlot and Matthieu Lerasle

PhD in progress: Jeanne Nguyen, 2015, Claire Lacour

PhD in progress: Benjamin Goehry, 2015, Pascal Massart and Jean-Michel Poggi

9.2.3. Juries

- Ph.D. Jérémy Bensadon: Sylvain Arlot (president)
- Ph.D. Gwenaelle Mabon: Sylvain Arlot (member)
- Ph.D. Mokhtar Alaya: Sylvain Arlot (president)
- Ph.D. Marie-Liesse Cauwet: Sylvain Arlot (member)

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] S. ARLOT, R. GENUER.*Comments on: "A Random Forest Guided Tour" by G. Biau and E. Scornet*, in "Test", 2016, vol. 25, n⁰ 2, p. 228–238, The final publication is available at Springer: http://dx.doi.org/10.1007/s11749-016-0484-4 [*DOI : 10.1007/S11749-016-0484-4*], https://hal.archives-ouvertes.fr/hal-01297557.

[2] G. BIAU, K. BLEAKLEY, B. CADRE.*The Statistical Performance of Collaborative Inference*, in "Journal of Machine Learning Research", 2016, https://hal.archives-ouvertes.fr/hal-01170254.

[3] F. CHAZAL, P. MASSART, B. MICHEL.*Rates of convergence for robust geometric inference*, in "Electronic journal of statistics ", 2016, vol. 10, n⁰ 2, 44, https://hal.inria.fr/hal-01336913.

[4] C. LACOUR, P. MASSART.*Minimal penalty for Goldenshluger-Lepski method*, in "Stochastic Processes and their Applications", December 2016, vol. 126, n⁰ 12, p. 3774–3789 [*DOI : 10.1016/J.SPA.2016.04.015*], https://hal.archives-ouvertes.fr/hal-01121989.

International Conferences with Proceedings

[5] M. MARIADASSOU, V. BRAULT, C. KERIBIN.*Normalité asymptotique de l'estimateur du maximum de vraisemblance dans le modèle de blocs latents*, in "48èmes Journées de Statistique de la SFdS", Montpellier, France, 48èmes Journées de Statistique de la SFdS, May 2016, https://hal.inria.fr/hal-01440084.

Conferences without Proceedings

[6] J.-P. BAUDRY, G. CELEUX.*What does ICL tell us about homogeneity for Model-Based Clustering?*, in "Workshop on Model-Based Clustering and Classification", Catane, Italy, Salvatore Ingrassia, September 2016, https://hal.inria.fr/hal-01437468.

[7] G. CELEUX, A. WHITE, J. WYSE.*Comparing EM with a greedy search algorithm to optimize ICL for Poisson mixture models* , in "Worshop on Model-Based Clustering", Paris, France, July 2016, https://hal.inria.fr/hal-01437459.

Other Publications

[8] E. AAMARI, C. LEVRARD.*Stability and Minimax Optimality of Tangential Delaunay Complexes for Manifold Reconstruction*, June 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01245479.

[9] S. ARLOT, A. CELISSE, Z. HARCHAOUI.*A kernel multiple change-point algorithm via model selection*, March 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-00671174.

[10] N. EL HAOUIJ, J.-M. POGGI, R. E. GHOZI, S. SEVESTRE-GHALILA, M. JAÏDANE.*Random Forest-Based Approach for Physiological Functional Variable Selection: Towards Driver's Stress Level Classification*, January 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01426752.

[11] D. GARREAU, S. ARLOT.*Consistent change-point detection with kernels*, December 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01416704.

[12] C. LACOUR, P. MASSART, V. RIVOIRARD.*Estimator selection: a new method with applications to kernel density estimation*, July 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01346081.

[13] P. SOBCZYK, M. BOGDAN, J. JOSSE.*Bayesian dimensionality reduction with PCA using penalized semi-integrated likelihood*, 2016, 31 pages, 7 figures, https://hal.archives-ouvertes.fr/hal-01342815.

Project-Team SMIS

Secured and Mobile Information Systems

IN COLLABORATION WITH: Parallelisme, réseaux, systèmes, modélisation (PRISM)

IN PARTNERSHIP WITH:
CNRS

Université Versailles Saint-Quentin

RESEARCH CENTER
Saclay - Île-de-France

THEME
Data and Knowledge Representation and Processing

Table of contents

Project-Team SMIS

Creation of the Project-Team: 2004 September 01, end of the Project-Team: 2016 December 31

Keywords:

Computer Science and Digital Science:

       1.1.6. - Cloud

       1.1.8. - Security of architectures

       1.4. - Ubiquitous Systems

       3.1.2. - Data management, quering and storage

       3.1.3. - Distributed data

       3.1.5. - Control access, privacy

       3.1.6. - Query optimization

       3.1.8. - Big data (production, storage, transfer)

       3.1.9. - Database

       4.3. - Cryptography

       4.7. - Access control

       4.8. - Privacy-enhancing technologies

Other Research Topics and Application Domains:

       2. - Health

       6.4. - Internet of things

       6.5. - Information systems

       6.6. - Embedded systems

       8.2. - Connected city

       8.5. - Smart society

       9.8. - Privacy

       9.10. - Ethics

1. Members

Research Scientists

Nicolas Anciaux [Inria, Researcher, HDR]

Luc Bouganim [Inria, Senior Researcher, HDR]

Faculty Members

Philippe Pucheral [Team leader, Univ. Versailles, Professor, HDR]

Iulian Sandu Popa [Univ. Versailles, Associate Professor]

Guillaume Scerri [Univ. Versailles, Associate Professor, from Sept. 2016]

Technical Staff

Aydogan Ersoz [Inria]

Oana Manea [Inria, from Nov. 2016]

PhD Students

Athanasia Katsouraki [Inria, until Sept. 2016]

Saliha Lallali [Inria, granted by ANR KISS project, until Jan. 2016]

Paul Tran Van [CozyCloud, CIFRE]

Axel Michel [INSA-CVL, from Sept. 2015]
Julien Loudet [CozyCloud, CIFRE, from April 2016]
Riad Ladjel [IDEX grant, from Oct. 2016]
Visiting Scientist
Benjamin Nguyen [INSA CVL, Professor, HDR]
Administrative Assistant
Emmanuelle Perrot [Inria]

2. Overall Objectives

2.1. Overall Objectives

The research work within the project-team is devoted to the design and analysis of core database techniques dedicated to the definition of secured and mobile information systems.

Ubiquitous computing and ambient intelligence entail embedding data in increasingly light and specialized devices (chips, sensors and electronic appliances for smart buildings, telephony, transportation, health, etc.). These devices exhibit severe hardware constraints to match size, security, power consumption and also production costs requirements. At the same time, they could highly benefit from embedded database functionalities to store data, analyze it, query it and protect it. This raises a first question "Q_1: *How to make powerful data management techniques compatible with highly constrained hardware platforms?*". To tackle this question, SMIS contributes to the design and validation of new storage and indexing models, query execution and optimization techniques, and transaction protocols. The relevance of this research goes beyond embedded databases and may have potential applications for database servers running on advanced hardware.

By making information more accessible and by multiplying –often transparently– the means of acquiring it, ubiquitous computing involves new threats for data privacy. The second question addressed by the project-team is then "Q_2: *How to make smart objects less intrusive?*". New access and usage control models have to be devised to help individuals keep a better control on the acquisition and sharing conditions of their data. This means integrating privacy principles like user's consent, limited collection and limited retention in the access and usage control policy definition. This also means designing appropriate mechanisms to enforce this control and provide accountability with strong security guarantees.

In parallel, thanks to a high degree of decentralization and to the emergence of low cost tamper-resistant hardware, ubiquitous computing contains the seeds for new ways of managing personal/sensitive data. The third question driving the research of the project-team is therefore "Q_3: *How to build privacy-by-design architectures based on trusted smart objects?*". The objective is to capitalize on embedded data management techniques, privacy-preserving mechanisms, trusted devices and cryptographic protocols to define an integrated framework dedicated to the secure management of personal/sensitive data. The expectation is showing that credible alternatives to a systematic centralization of personal/sensitive data on servers can be devised and validating the approach through real case experiments.

3. Research Program

3.1. Embedded Data Management

The challenge tackled is this research action is twofold: (1) to design embedded database techniques matching the hardware constraints of (current and future) smart objects and (2) to set up co-design rules helping hardware manufacturers to calibrate their future platforms to match the requirements of data driven applications. While a large body of work has been conducted on data management techniques for high-end servers (storage, indexation and query optimization models minimizing the I/O bottleneck, parallel DBMS, main memory

DBMS, etc.), less research efforts have been placed on embedded database techniques. Light versions of popular DBMS have been designed for powerful handheld devices; yet DBMS vendors have never addressed the complex problem of embedding database components into chips. Proposals dedicated to databases embedded on chip usually consider small databases, stored in the non-volatile memory of the microcontroller –hundreds of kilobytes– and rely on NOR Flash or EEPROM technologies. Conversely, SMIS is pioneering the combination of microcontrollers and NAND Flash constraints to manage Gigabyte(s) size embedded databases. We present below the positioning of SMIS with respect to international teams conducting research on topics which may be connected to the addressed problem, namely work on electronic stable storage, RAM consumption and specific hardware platforms.

Major database teams are investigating data management issues related to hardware advances (EPFL: A. Ailamaki, CWI: M. Kersten, U. Of Wisconsin: J. M. Patel, Columbia: K. Ross, UCSB: A. El Abbadi, IBM Almaden: C. Mohan, etc.). While there are obvious links with our research on embedded databases, these teams target high-end computers and do not consider highly constrained architectures with non traditional hardware resources balance. At the other extreme, sensors (ultra-light computing devices) are considered by several research teams (e.g., UC Berkeley: D. Culler, ITU: P. Bonnet, Johns Hopkins University: A. Terzis, MIT: S. Madden, etc.). The focus is on the processing of continuous streams of collected data. Although the devices we consider share some hardware constraints with sensors, the objectives of both environments strongly diverge in terms of data cardinality and complexity, query complexity and data confidentiality requirements. Several teams are looking at efficient indexes on flash (HP LABS: G. Graefe, U. Minnesota: B. Debnath, U. Massachusetts: Y. Diao, Microsoft: S. Nath, etc.). Some studies try to minimize the RAM consumption, but the considered RAM/stable storage ratio is quite large compared to the constraints of the embedded context. Finally, a large number of teams have focused on the impact of flash memory on database system design (we presented an exhaustive state of the art in a VLDB tutorial [34]). The work conducted in the SMIS team on bi-modal flash devices takes the opposite direction, proposing to influence the design of flash devices by the expression of database requirements instead of running after the constantly evolving flash device technology.

3.2. Access and Usage Control Models

Access control management has been deeply studied for decades. Different models have been proposed to declare and administer access control policies, like DAC, MAC, RBAC, TMAC, and OrBAC. While access control management is well established, new models are being defined to cope with privacy requirements. Privacy management distinguishes itself from traditional access control is the sense that the data to be protected is personal. Hence, the user's consent must be reflected in the access control policies, as well as the usage of the data, its collection rules and its retention period, which are principles safeguarded by law and must be controlled carefully.

The research community working on privacy models is broad, and involves many teams worldwide including in France ENST-B, LIRIS, Inria LICIT, and LRI, and at the international level IBM Almaden, Purdue Univ., Politecnico di Milano and Univ. of Milano, George Mason Univ., Univ. of Massachusetts, Univ. of Texas and Colorado State Univ. to cite a few. Pioneer attempts towards privacy wary systems include the P3P Platform for Privacy Preservation [36] and Hippocratic databases [29]. In the last years, many other policy languages have been proposed for different application scenarios, including EPAL [40], XACML [39] and WSPL [32]. Hippocratic databases are inspired by the axiom that databases should be responsible for the privacy preservation of the data they manage. The architecture of a Hippocratic database is based on ten guiding principles derived from privacy laws.

The trend worldwide has been to propose enhanced access control policies to capture finer behavior and bridge the gap with privacy policies. To cite a few, Ardagna *et al.* (Univ. Milano) enables actions to be performed after data collection (like notification or removal), purpose binding features have been studied by Lefevre *et al.* (IBM Almaden), and Ni *et al.* (Purdue Univ.) have proposed obligations and have extended the widely used RBAC model to support privacy policies.

The positioning of the SMIS team within this broad area is rather (1) to focus on intuitive or automatic tools helping the individual to control some facets of her privacy (e.g., data retention, minimal collection) instead of increasing the expressiveness but also the complexity of privacy models and (2) to push concrete models enriched by real-case (e.g., medical) scenarios and by a joint work with researchers in Law.

3.3. Tamper-resistant Data Management

Tamper-resistance refers to the capacity of a system to defeat confidentiality and integrity attacks. This problem is complementary to access control management while being (mostly) orthogonal to the way access control policies are defined. Security surveys regularly point out the vulnerability of database servers against external (i.e., by intruders) and internal (i.e., by employees) attacks. Several attempts have been made in commercial DBMSs to strengthen server-based security, e.g., by separating the duty between DBA and DSA (Data Security Administrator), by encrypting the database footprint and by securing the cryptographic material using Hardware Security Modules (HSM) [35]. To face internal attacks, client-based security approaches have been investigated where the data is stored encrypted on the server and is decrypted only on the client side. Several contributions have been made in this direction, notably by U. of California Irvine (S. Mehrotra, Database Service Provider model), IBM Almaden (R. Agrawal, computation on encrypted data), U. of Milano (E. Damiani, encryption schemes), Purdue U. (E. Bertino, XML secure publication), U. of Washington (D. Suciu, provisional access) to cite a few seminal works. An alternative, recently promoted by Stony Brook Univ. (R. Sion), is to augment the security of the server by associating it with a tamper-resistant hardware module in charge of the security aspects. Contrary to traditional HSM, this module takes part in the query computation and performs all data decryption operations. SMIS investigates another direction based on the use of a tamper-resistant hardware module on the client side. Most of our contributions in this area are based on exploiting the tamper-resistance of secure tokens to build new data protection schemes.

While our work on Privacy-Preserving data Publishing (PPDP) is still related to tamper-resistance, a complementary positioning is required for this specific topic. The primary goal of PPDP is to anonymize/sanitize microdata sets before publishing them to serve statistical analysis purposes. PPDP (and privacy in databases in general) is a hot topic since 2000, when it was introduced by IBM Research (IBM Almaden: R. Agrawal, IBM Watson: C.C. Aggarwal), and many teams, mostly north American universities or research centres, study this topic (e.g., PORTIA DB-Privacy project regrouping universities such as Stanford with H. Garcia-Molina). Much effort has been devoted by the scientific community to the definition of privacy models exhibiting better privacy guarantees or better utility or a balance of both (such as differential privacy studied by C. Dwork: Microsoft Research or D. Kifer: Penn-State Univ and J. Gehrke: Cornell Univ) and thorough surveys exist that provide a large overview of existing PPDP models and mechanisms [37]. These works are however orthogonal to our approach in that they make the hypothesis of a trustworthy central server that can execute the anonymization process. In our work, this is not the case. We consider an architecture composed of a large population of tamper-resistant devices weakly connected to an untrusted infrastructure and study how to compute PPDP problems in this context [1]. Hence, our work has some connections with the works done on Privacy Preserving Data Collection (Stevens Institute of Tech. / Rutgers Univ,NJ: R.N.Wright, Univ Austin Texas: V. Shmatikov), on Secure Multi-party Computing for Privacy Preserving Data Mining (Rutgers Univ: J. Vaidya, Purdue Univ: C. Clifton) and on distributed PPDP algorithms (Univ Wisconsin: D. DeWitt, Univ Michigan: K. Lefevre, Rutgers Univ: J. Vaidya, Purdue Univ: C. Clifton) while none of them share the same architectural hypothesis as us.

4. Application Domains

4.1. Application Domains

Our work addresses varied application domains. Typically, data management techniques on chip are required each time data-driven applications have to be embedded in ultra-light computing devices. This situation occurs for example in healthcare applications where medical folders are embedded into smart tokens (e.g.,

smart cards, secured USB keys), in telephony applications where personal data (address book, agenda, etc.) is embedded into cellular phones, in sensor networks where sensors log row measurements and perform local computation on them, in smart-home applications where a collection of smart appliances gather information about the occupants to provide them a personalized service, and more generally in most applications related to ambient intelligence.

Safeguarding data confidentiality has become a primary concern for citizens, administrations and companies, broadening the application domains of our work on access control policies definition and enforcement. The threat on data confidentiality is manifold: external and internal attacks on the data at rest, on the data on transit, on the data hosted in untrusted environments (e.g., Database Service Providers, Web-hosting companies) and subject to illegal usage, insidious gathering of personal data in an ambient intelligence surrounding. Hence, new access control models and security mechanisms are required to accurately declare and safely control who is granted access to which data and for which purpose.

While the application domain mentioned above is rather large, two applications are today more specifically targeted by the SMIS team. The first one deals with privacy preservation in EHR (Electronic Health Record) systems and PCEHR (Personally Controlled EHR) [3]. We are developing technologies tackling this issue and experiment them in the field. The second application area deals with privacy preservation in the context of personal Cloud, that is personal data hosted in dedicated servers staying under the holder's control (e.g., in a personal internet box or in a home automation box).

5. New Software and Platforms

5.1. PLUG-DB ENGINE

FUNCTIONAL DESCRIPTION: PlugDB is a complete platform dedicated to a secure and ubiquitous management of personal data. It aims at providing an alternative to a systematic centralization of personal data. The PlugDB engine is a personal database server capable of storing data (tuples and documents) in tables and BLOBs, indexing them, querying them in SQL, sharing them through assertional access control policies and enforcing transactional properties (atomicity, integrity, durability) [4]. The PlugDB engine is embedded in a tamper-resistant hardware device combining the security of smartcard with the storage capacity of NAND Flash. The personal database is hosted encrypted in NAND Flash and the PlugDB engine code runs in the microcontroller. Complementary modules allow to pre-compile SQL queries for the applications, communicate with the DBMS from a remote Java program, synchronize local data with remote servers (typically used for recovering the database in the case of a broken or lost devices) and participate in distributed computation (e.g., global queries). PlugDB runs both on secure devices provided by Gemalto and on specific secure devices designed by SMIS and assembled by electronic SMEs. Mastering the hardware platform opens up new research and experiment opportunities (e.g., we have recently integrated a Bluetooth module to communicate wirelessly with PlugDB and a fingerprint module to strongly authenticate users) and allows us to engage ourselves in an open-source/open hardware initiative. Open-SW/open-HW contributes to the trust the community of users can put in any privacy preserving solution and is key to enable a diversity of solutions, hence decreasing the risk of class attacks. PlugDB engine has been registered first at APP (Agence de Protection des Programmes) in 2009 - a new version being registered every two years and the hardware datasheets in 2015. PlugDB has been experimented in the field, notably in the healthcare domain. We also recently set up an educational platform on top of PlugDB, named SIPD (Système d'Information privacy-by-Design) and used at ENSIIE, INSA CVL and UVSQ through the Versailles Sciences Lab fablab, to raise students awareness of privacy protection problems and embedded programming. As a conclusion, PlugDB combines several research contributions from the team, at the crossroads of flash data management, embedded data processing and secure distributed computations. It then strongly federates all members of our team (permanent members, PhD students and engineers). It is also a vector of visibility, technological transfer and dissemination and gives us the opportunity to collaborate with researchers from other disciplines around a concrete privacy enhancing platform.

- Participants: Nicolas Anciaux, Luc Bouganim, Philippe Pucheral and Aydogan Ersoz

- Contact: Nicolas Anciaux
- URL: https://project.inria.fr/plugdb/

5.2. Privacy Preserving Mobile Laboratory

FUNCTIONAL DESCRIPTION: We have started to design a privacy preserving mobile laboratory used as an experimental platform for multidisciplinary research launched 'in vivo'. The goal is to conduct reliable surveys and avoid the privacy paradox (what users declare on their privacy behavior is far from what they effectively do). The platform, built on top of PlugDB, includes two android applications, a "server" which takes as input a questionnaire description and broadcast it on demand to the client applications. Users interact with the questionnaire on the client applications, storing the detailed answers in their PlugDB personal server. Then a secure distributed computation takes place (between users' PlugDB servers) and computes non-sensitive global statistics based on potentially sensitive raw answers. A beta-version of this platform was developed during the PhD of Athanasia Katsouraki and was used for a pre-experimentation targeting 140 students. While the experiment was successful, it showed the limitation and complexity of the initial setting (laptops, required Internet access, complexity in the questionnaire deployment). We designed and implement a second platform running on android tablets with a local router and automatic questionnaire deployment. The platform has been demonstrated in several forums and very recently at the Sénat in Paris. This platform represents a backing for two PhD theses on privacy (the first one in economics, the second one in our team) funded in 2016 by the interdisciplinary doctoral program at UPSay (IDI 2016).

- Participants: Nicolas Anciaux, Luc Bouganim, Aydogan Ersoz, Athanasia Katsouraki, Riad Ladjel, Benjamin Nguyen, Remy Pasquion, Paul Tran Van
- Contact: Luc Bouganim
- URL: https://project.inria.fr/plugdb/en/PPML

6. New Results

6.1. Embedded Data Management

Participants: Nicolas Anciaux, Saliha Lallali, Philippe Pucheral, Iulian Sandu Popa [correspondent].

Embedded keyword indexing: In this work, we revisit the traditional problem of information retrieval queries over large collections of files in an embedded context. A file can be any form of document, picture or data stream, associated with a set of terms. A query can be any form of keyword search using a ranking function (e.g., TF-IDF) identifying the top-k most relevant files. The proposed search engine can be used in sensors to search for relevant objects in their surroundings, in cameras to search pictures by using tags, in personal smart dongles to secure the querying of documents and files hosted in an untrusted Cloud, or in a personal cloud securely managed using a tamper resistant smart object. A search engine is usually based on a (large) inverted index and queries are traditionally evaluated by allocating one container in RAM per document to aggregate its score, making the RAM consumption linear with the size of the document corpus. To tackle this issue, we designed a new form of inverted index which can be accessed in a pure pipeline manner to evaluate search queries without materializing any intermediate result. Successive index partitions are written once in Flash and maintained in the background by timely triggering merge operations while files are inserted or deleted from the index. This work was initially published at VLDB'15 [5] and demonstrated at SIGMOD'15 [38]. It constitutes the main contribution of the PhD thesis of Saliha Lallali defended in January 2016. In 2016, we extended this work to demonstrate at EDBT'16 [22] its applicability to set up a secure distributed search engine for the Personal Cloud. We also complemented this work with (1) a thorough analysis of the RAM consumption linked to the main algorithms implementing the solution, (2) the support of conditional top-k queries in a personal Cloud context that we consider as a killer application domain today and (3) new performance measurements with a real dataset (ENRON), representative of this personal Cloud context. These new contributions have been submitted to Information Systems journal.

6.2. Secure Global Computing on Asymmetric Architecture

Participants: Benjamin Nguyen [correspondent], Axel Michel, Philippe Pucheral, Iulian Sandu Popa.

Asymmetric Architecture Computing: This research direction studies the secure execution of various algorithms on data stored in an unstructured network of Trusted Cells (i.e., personal trusted device) so that each user can keep control over her data. The data could be stored locally in a trusted cell or encrypted on some external cloud. Execution takes place on a specific infrastructure called the Asymmetric Architecture (AA): the network of trusted cells, supported by an untrusted cloud supporting IaaS or PaaS. Our objective is to show that many different algorithms and computing paradigms can be executed on AA, thus achieving secure and private computation. Our first contribution in this area was to study the execution of Privacy Preserving Data Publishing algorithms on such an architecture (T. Allard's PhD Thesis). Then we studied general SQL queries in this same execution context. We concentrated on the subset of SQL queries without joins, but including Group By and aggregates, and show how to secure their execution in the presence of honest-but-curious attackers. This work, named SQL-AA and notably published at EDBT'14 [8] and demonstrated at VLDB'15, was part of Quoc-Cuong To's Ph.D defended in 2015. We have extended this framework through a collaboration with INSA Centre Val de Loire, LIFO Lab and University of Paris Nord, LIPN lab and have shown in CoopIS'15 [9] that it is possible to achieve seamless integration of distributed MapReduce processing using trusted cells, while maintaining reasonable performance. In 2016, we added three novel contributions to SQL-AA: (i) an extended privacy analysis in which we consider stronger adversaries with more background knowledge, (ii) an extended threat model in which we consider malicious attacker and propose safety properties to prevent malicious attacks and (iii) we tackled practical issues like exchanging securely shared keys among trusted cells and Querier (GKE protocol) and enforcing access control at query execution time. These new contributions have been published in TODS'16 [15]. In parallel, we are starting a new study in the line of our previous work on Privacy Preserving Data Publishing (PPDP) with the objective to inject individualized privacy requirements in the PPDP protocol. A preliminary contribution has been published at BDA'16 [25] to compute SQL aggregate queries under k-anonymity constraints where each individual contributing to the query may define her own k constraint, thereby letting each one weighting differently the sensitiveness of a given piece of information according to her own situation.

Secure spatio-temporal distributed processing: Mobile participatory sensing could be used in many applications such as vehicular traffic monitoring, pollution tracking, or even health surveying (e.g., to allow measuring in real-time the individual exposure to environmental risk factors or the propagation of an epidemic). However, its success depends on finding a solution for querying a large number of users which protects user location privacy and works in real-time [10]. We addressed these issues and proposed PAMPAS, a privacy-aware mobile distributed system for efficient data aggregation in mobile participatory sensing. In PAMPAS, mobile devices enhanced with secure hardware, called secure probes, perform distributed query processing, while preventing users from accessing other users' data. Secure probes exchange data in encrypted form with help from an untrusted supporting server infrastructure. PAMPAS uses two efficient, parallel, and privacy-aware protocols for location-based aggregation and adaptive spatial partitioning of secure probes. Our experimental results and security analysis demonstrate that these protocols are able to collect, aggregate and share statistics or derived data in real-time, without any privacy leakage. This work is part of Dai Hai Ton That's Ph.D. thesis defended in January 2016, co-supervised by Iulian Sandu Popa. The system implementation was demonstrated in [41], while two papers describing the technical details of the system have been published in 2016 [23], [16].

6.3. Personal Cloud

Participants: Nicolas Anciaux [correspondent], Luc Bouganim, Julien Loudet, Benjamin Nguyen, Philippe Pucheral, Iulian Sandu Popa, Guillaume Scerri, Paul Tran Van.

We are witnessing an exponential increase in the acquisition of personal data about the individuals or produced by them. Today, this information is managed using Web applications, centralizing this data in cloud data servers, under the control of few Web majors [2]. However, it has now become clear that (1) centralizing millions of personal records exposes the data to very sophisticated attacks, linked to a very high potential benefit in case of success (millions of records being revealed), and (2) delegating the management of personal

records without any tangible guarantee for the individuals leads to privacy violations, the data being potentially made accessible to other organizations (e.g., governments, commercial partners) and being subject to lucrative secondary usages (not advertised to the individuals). To face this situation, many recent initiatives push towards the emergence of the Personal Cloud paradigm. A personal cloud can be viewed as a personal server, owned by a given individual, which gives to its owner the ability to store her complete digital environment, synchronize it among various devices and share it with other individuals and applications under control. In the SMIS team, we claim the need of a Secure Personal Cloud, and promote the introduction of a secure (tamper resistant) data engine in the architecture [30]. On this basis, we investigate new data sharing and dissemination models, where usage and access control rules endorsed by the individuals could be enforced and have presented this vision at EDBT'14 [6] and at ADBIS'15 [31]. We have started a cooperation with the startup CozyCloud at the end of 2014. A contract was signed at the end of 2014 to integrate PlugDB in a CozyCloud instance and two CIFRE PhD thesis have been launched so far. Paul Tran Van's PhD thesis explores a new data sharing paradigm dedicated to the personal cloud context. This paradigm, called SWYSWYK (Share What You See with Who You Know), allows to automatically derive intuitive sharing rules from a personal cloud content, to share rules among a community of users and to let each user physically visualize the net effects of these rules on her own Personal Cloud. We propose a reference architecture providing the users with tangible guarantees about the enforcement of SWYSWYK policies and demonstrate through a performance evaluation conducted on a real personal cloud platform that the approach is practical. This work constitutes the core of Paul Tran Van's thesis and is being submitted for publication at VLDB. Preliminary ideas related to this work are presented in ERCIM news'16 [27]. Julien Loudet's PhD thesis is just starting with the objective to explore privacy-preserving distributed computations over personal clouds.

More generally, the personal cloud context gain in importance in our research work. It is even at the heart of our future project-team named PETRUS (PErsonal and TRUSted cloud). PETRUS is expected to take over from the SMIS team beginning of 2017.

6.4. Interdisciplinary study on Privacy-by-Design

Participants: Nicolas Anciaux, Luc Bouganim [correspondent], Athanasia Katsouraki, Benjamin Nguyen, Philippe Pucheral.

The objective of this research action is to study the reciprocal entanglements between economic, legal, societal and technological aspects of the management and exploitation of personal data. Indeed, devising new ways of protecting data privacy cannot be done in isolation; it requires also identifying alternative economic models that are both viable and regulatory compliant. We started an interdisciplinary research work with economists (RITM Lab) and jurists (CERDI and DANTE labs) in the privacy axis of ISN (Institut de la Société Numérique) and plan to pursue it in two projects in preparation: the Convergence Institute I2DRIVE (Interdisciplinary Institute for Data Research: Intelligence, Values and Ethics) and the CNRS Federation SIHS (Sciences Informatiques, Humaines et Sociales) at UVSQ. A first interdisciplinary work conducted in 2016 concerns the design of a privacy preserving platform needed to conduct privacy studies "in vivo". Such platforms are required to validate the effectiveness of privacy preserving solutions, in terms of technical feasibility, lawfulness, acceptability and benefits. To this end, we have designed a privacy preserving mobile lab, derived from the personal cloud platform developed by the team (see 'Software' section). In her PhD thesis, Athanasia Katsouraki developed a beta-version of that platform and used it to perform a pre-experimentation in the context of online form based survey, targeting 140 students. The goal was threefold: (1) to test the effectiveness of the proposed platform, (2) to test the adequacy of the questionnaire and experimentation protocol (a result for the experimental economist), and (3) to check the impact of the use of a secure platform on the student's answers. The pre-experimentation showed several improvement axis and led to the actual design of the privacy preserving mobile lab described in the Software section.

Another joint work is related to the design of technical means to help individuals perceive how their personal life is exposed compared to others and to make appropriate protection choices. This work led to the definition of a new principle called Privacy-by-Using [20], that we introduced to try to circumvent the limits of the privacy-by-design principle promoted by the regulator. The confrontation of the Privacy-by-Using principle with Big Data processing [26] has also been studied with jurists and economists.

Finally, we conducted a scientific expertise on behalf of DGCCRF (Direction Générale de la Concurrence, de la Consommation et de la Répression des Fraudes) and of the European Council regarding the draft proposal of "Directive of the European Parliament and of the Council on certain aspects concerning contracts for the online and other distance sales of goods" regulating the payment of numeric goods and services by means of personal data. This led us to a cross-analysis, with researchers in Law and computer scientists, of technical, societal and economic issues linked to the smart disclosure principle, that is, under which conditions and formats individuals can get their data back from service providers [17], [19], [18].

6.5. Formal guarantees

Participant: Guillaume Scerri.

The aim of the action is to investigate the changes required for the PlugDB architecture to be amenable to formal security proofs.

More precisely we started exploring the precise formal guarantees that are desirable for a personal data server. Following work started in Bristol [7], relating to formal guarantees provided by secure hardware, we started studying how one could leverage the low level guarantees provided by secure hardware (PlugDB for example) to cover the more complex operations and guarantees required of a personal data server. The first finding of the action is that a modular architecture is required for formal proofs to be obtainable. This is reflected in the architectural concerns presented in the PETRUS project.

Additionally, we started studying how to leverage secure hardware guarantees in order to perform secure computations on distributed data. A first result in this direction is presented in [33], and submitted to Financial Cryptography 2017.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

7.1.1. Cozy Cloud bilateral contract (Dec 2014 - Nov. 2015)

Partners: Cozy Cloud, Inria-SMIS
SMIS funding: 50k€.
While this initial contract is over, we mention it to explain the increasing relationship being built between Cozy Cloud and our team. Cozy Cloud is a French startup providing a personal Cloud platform. The Cozy product is a software stack that anyone can deploy to run his personal server in order to host his personal data and web services. While centralizing all personal data in the holder's hand is a natural way to reestablish his control on his privacy, this represents an unprecedented threat in case of attacks by an intruder, especially for individuals who are not security experts. The objective of this bilateral contract is to address this issue by integrating the PlugDB solution into the Cozy stack. Roughly speaking, the Cozy data system will be modified in such a way to store only encrypted files and each file access will be intercepted and routed to PlugDB. PlugDB will act as a doorkeeper for the whole individual dataspace by managing the files' metadata, the access control rules defined on these metadata, the decryption keys and the user/application authentication.

7.1.2. Cozy Cloud CIFRE - Tran Van contract (Oct 2014 - Sept 2017)

Partners: Cozy Cloud, Inria-SMIS
SMIS funding: 30k€
In relation with the bilateral contract mentioned above, a CIFRE PhD thesis has been started by Paul Tran Van. The objective is to capitalize on the Cozy-PlugDB platform to devise new access and usage control models to exchange data among devices of the same user (devices may have different levels of trustworthiness) and among different users thanks to a user-friendly sharing model (see the work on the SWYSWYK - Share What You See with Who You Know - model presented above) .

7.1.3. Cozy Cloud CIFRE - Loudet contract (Apr 2016 - Apr 2019

Partners: Cozy Cloud, Inria-SMIS
SMIS funding: 45k€
In relation with the bilateral contract mentioned above, a second CIFRE PhD thesis has been started by Julien Loudet. The objective is to allow for a secure execution of distributed queries on a set of personal clouds associated to users, depending on social links, user's localization or user's profile. The general idea is to build secure indexes, distributed on the users' personal cloud and to devise a secure execution protocol revealing solely the query result to the querier. Such highly distributed secure queries potentially enable new (social) applications fed by user's personal data which could be developed on the Cozy-PlugDB platform.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR PerSoCloud (Jan. 2017 - Jan. 2020)

Partners: Orange Labs (coordinator), Inria-SMIS, Cozy Cloud, Univ. of Versailles.
SMIS funding: 170k€.

The objective of PerSoCloud is to design, implement and validate a fullfledged Privacy-by-Design Personal Cloud Sharing Platform. One of the major difficulties linked to the concept of personal cloud lies in organizing and enforcing the security of the data sharing while the data is no longer under the control of a central server. We identify three dimensions to this problem. Devices-sharing: assuming that the primary copy of user U1's personal data is hosted in a secure place, how to share and synchronize it with U1's multiple (mobile) devices without compromising security? Peers-sharing: how user U1 could exchange a subset of his-her data with an identified user U2 while providing to U1 tangible guarantees about the usage made by U2 of this data? Community-sharing: how user U1 could exchange a subset of his-her data with a large community of users and contribute to personal big data analytics while providing to U1 tangible guarantees about the preservation of his-her anonymity? In addition to tackling these three scientific and technical issues, a legal analysis will guarantee compliance of this platform with the security and privacy French and UE regulation, which firmly promotes the Privacy by Design principle, including the current reforms of personal data regulation.

8.1.2. ANR KISS (Dec. 2011 - Feb. 2016)

Partners: Inria-SMIS (coordinator), Inria-SECRET, LIRIS, Univ. of Versailles, CryptoExperts, Gemalto, Yvelines district.
SMIS funding: 230k€.

The idea promoted in KISS is to embed, in trusted devices, software components capable of acquiring, storing and managing securely various forms of personal data (e.g., salary forms, invoices, banking statements, geolocation data, depending on the applications). These software components form a Personal Data Server which can remain under the holder's control. The scientific challenges include: embedded data management issues tackling regular, streaming and spatio-temporal data (e.g., geolocation data), data provenance-based privacy models, crypto-protected distributed protocols to implement private communications and secure global computations.

8.1.3. PIA - PDP SECSi (May. 2016 - Dec. 2017)

Partners: Cozy Cloud (coordinator), Qwant, Inria-SMIS, FING.
SMIS funding: 149k€.

The objective of this PIA-PDP (Programme Investissement d'Avenir - Protection des Données Personnelles) SECSi project is to build a concrete Personal Cloud platform which can support a large scale deployment of Self Data services. Three major difficulties are identified and will be tackled in this project: (1) how to implement and enforce a fine control of the data flow when personal data are exploited by third party applications, (2) how to protect these same applications when processing is delegated to the personal cloud platform itself and (3) how to implement personalized search on the web without hurting user's privacy.

8.1.4. CAPPRIS Project-Lab (Dec. 2011 - Dec. 2016)

Inria Partners: PRIVATICS (coordinator), SMIS, PLANETE, CIDRE, COMETE.
External partners: Univ. of Namur, Eurecom, LAAS.
Funding: not associated to individual project-teams.

An Inria Project Lab (IPL) is a long-term multi-disciplinary project launched by Inria to sustain large scale risky research actions in line with its own strategic plan. CAPPRIS stands for "Collaborative Action on the Protection of Privacy Rights in the Information Society". The key issues that are addressed are: (1) the identification of existing and future threats to privacy, (2) the definition of formally grounded measures to assess and quantify privacy, (3) the definition of the fundamental principles underlying privacy by design and methods to apply them in concrete situations and (4) The integration of the social and legal dimensions. To assess the relevance and significance of the research results, they are confronted to three classes of case studies CAPPRIS partners are involved in: namely Online Social Networks, Location Based Services and Electronic Health Record Systems.

8.1.5. CityLab@Inria, Inria Project Lab (May 2014 -).

Inria Partners: CLIME, DICE, FUN, MIMOVE, MYRIADS, SMIS, URBANET, WILLOW.
External partners: UC Berkeley.
Funding: not associated to individual project teams.

CityLab@Inria studies ICT solutions toward smart cities that promote both social and environmental sustainability. A strong emphasis of the Lab is on the undertaking of a multi-disciplinary research program through the integration of relevant scientific and technology studies, from sensing up to analytics and advanced applications, so as to actually enact the foreseen smart city Systems of Systems. SMIS contributes to Privacy-by-Design architectures for trusted smart objects so as to ensure privacy to citizens, which is critical for ensuring that urbanscale sensing contributes to social sustainability and does not become a threat. https://citylab.inria.fr/

8.1.6. VALDO (Valorisation et monétisation des données personnelles à l'ère du Big Data), Digital Society Institute (DSI) (May 2015 - Sept. 2016).

Partners: DANTE and SMIS (co-organizers), CERDI, RITM.
SMIS funding: 50K€.

The objective of this project is to study with a multidisciplinary approach (i.e., computer science, law and economics) the impact of putting a certain (e.g., monetary) value on personal data, over the behavior of individuals (that are the rightful owners of the data) and market companies (that make usage of the personal data) in terms of data protection practices and data usage.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

- Philippe Pucheral: Co-founder of the bi-annual French Summer School 'Masses de Données Distribuées' and co-organizer of this school in 2016

9.1.1.2. Member of the Organizing Committees

- Benjamin Nguyen: Steering commitee of 'Atelier sur la Protection de la Vie Privée' (APVP) in 2016
- Nicolas Anciaux: co-organizer (with Fabrice Le Guel and Ulysse Roux) of the 'Journée de restitution d'ISN' in 2016

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

- Philippe Pucheral: DATA'16, DBKDA'16, BDA'16, EDBT'17
- Luc Bouganim: VLDB'17, EDBT'16, BDA'16 (Demo)
- Benjamin Nguyen: ECML/PKDD'16
- Nicolas Anciaux: SIGMOD'17, EDBT'16, DATA'16, ICT4AWE'16, BDA'16, RESSI'16
- Iulian Sandu Popa: IEEE MobileCloud'16, DATA'16, MobilWare'16, MCSMS'16

9.1.2.2. Reviewer

- Luc Bouganim: EDBT'17
- Iulian Sandu Popa: MDM'16
- Guillaume Scerri : SIGMOD'17, EDBT'17, STACS'17

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Nicolas Anciaux: Associate Editor of the VLDB Journal (since 2015)
- Benjamin Nguyen: Techniques et Sciences Informatiques (TSI) - Revue Française (Rédacteur adjoint, domaine Sécurité et Vie Privée depuis 2016-05)
- Luc Bouganim: Area Editor for PVLDB 2018

9.1.3.2. Reviewer - Reviewing Activities

- Iulian Sandu Popa: IEEE TKDE, ACM Transactions on Storage, International Journal of Digital Earth

9.1.4. Invited Talks

- Invited tutorial: N. Anciaux, L. Bouganim, Tutorial: Decentralized Personal Data Management, using secure devices, BDA summer school: "Masses de données distribuées", Urugne, June 2016 [28].
- Invited talk: N. Anciaux, La valorisation de la donnée par l'usager : approche technique. Centre français de droit comparé, Colloque 'Propriété(s) et données', Paris, 2016 https://linc.cnil.fr/fr/le-centre-francais-de-droit-compare-organise-un-colloque-proprietes-et-donnees
- Invited talk: L. Bouganim, Remy Pasquion, A Secure and Mobile Platform for Experimental Economy, Multidisciplinary Research Seminar at Telecom Paris Tech, Paris, June 2016.

- Invited talk: N. Anciaux, Contrôler ses données personnelles par la décentralisation. Journée Protection de la vie privée : Nouveaux instruments techniques et juridiques, CNRS, November 2016. http://www.univ-orleans.fr/lifo/evenements/CAPPRIS/?page_id=125

- Invited demonstration: L. Bouganim and R. Pasquion, Présentation d'une Plateforme de Questionnaires Respectueux de la Vie Privée, Sommet « Partenariat pour un Gouvernement ouvert », Agora numérique « Pour un Sénat ouvert », December, 8, 2016.

- Invited demonstrations: N. Anciaux, L. Bouganim and P. Tran Van, Partager et disposer de ses données en en gardant la maîtrise, trois démonstrations sur la base de PlugDB: Mobile Platform for Experimental Economy, Secure personal cloud (with Cozy Cloud), Personal Social-Medical Folder. Journée Protection de la vie privée : Nouveaux instruments techniques et juridiques, CNRS, November 2016. http://www.univ-orleans.fr/lifo/evenements/CAPPRIS/?page_id=125

- Invited talk: Iulian Sandu Popa, Distributed Architectures for Privacy-Aware Mobile Participatory Sensing, BIS'16 Workshop, June 2016 https://project.inria.fr/siliconvalley/bis2016-day-2-full-day-working-session-on-smart-cities/

- Invited talk: N. Anciaux, A new approach for secure personal cloud, CITI Talk, Lyon, May 2016 http://www.citi-lab.fr/2016/05/17/citi-talk-a-new-approach-for-secure-personal-cloud-by-nicolas-anciaux-on-26th-may/

- Panel: Nicolas Anciaux, Enjeux et opportunités des modèles ouverts en santé, Paris Open Source Summit, 2016. Trackleader : Olivier de Fresnoye. Panelists : Nicolas Anciaux (Inria), Mobin Yasini(MD, MPH, PhD, mHealth Quality R&D Director), Stéphane Gigandet (Open Food Facts), Clémentine Langlois (Fongwama) http://empoweringopeninnovation.org/enjeux-et-opportunites-des-modeles-ouverts-en-sante/

9.1.5. Scientific Expertise

- N. Anciaux, P. Pucheral : scientific expertise on behalf of DGCCRF (Direction Générale de la Concurrence, de la Consommation et de la Répression des Fraudes) and of the European Council regarding the draft proposal of "Directive of the European Parliament and of the Council on certain aspects concerning contracts for the online and other distance sales of goods" regulating the payment of numeric goods and services by means of personal data

- Benjamin Nguyen: Natural Sciences and Engineering Research Council of Canada (NSERC), Research grants

- Benjamin Nguyen: Membre du comité d'éthique de TeraLab, du Groupe des Ecoles Nationales d'Economie et de Statistique (GENES) dont l'ENSAE.

9.1.6. Research Administration

- Philippe Pucheral: Member of the HDR committee of the STV doctoral school (UVSQ) since 2014

- Philippe Pucheral: Member of the steering committee of the ED STIC doctoral school of University Paris-Saclay, 'Data, Knowledge and Interactions' committee (about 250 PhD students) since 2014

- Philippe Pucheral: Representative of Inria in the 'Comité des Tutelles Formation' of Paris-Saclay University since 2016

- Philippe Pucheral: Member of the 'Commission de Sélection' of UVSQ and of the 'Commission des Appellations' of Télécom SudParis in 2016

- Luc Bouganim: Reviewer for the ANR programs (Evaluation committee CE23), 2016

- Luc Bouganim: Member of the 'Commission de Sélection' of the INSA CVL, 2016

- Nicolas Anciaux: Co-director of the 'Privacy and digital identity' WG at Digital Society Institute (DSI), until June 2016

- Nicolas Anciaux: Reviewer of the CHIST-ERA Call 2015 - SPTIoT

- Nicolas Anciaux: Member of the 'Commission de Sélection' of UVSQ, 2016

- Benjamin Nguyen: Director of LIFO (EA 4022) since July 2016
- Benjamin Nguyen: Responsible of CNRS privacy working group "Sécurité" since september 2016
- Benjamin Nguyen: Director of Digital Affairs (INSA CVL)
- Iulian Sandu Popa: Reviewer for the ANR programs (evaluation committee CE23), 2016
- Iulian Sandu Popa: Member of the 'Commission de Sélection' of the UVSQ, 2016

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master : Philippe Pucheral, co-director of the DataScale master (UPSay), courses in M1 and M2 at UVSQ and at ENSIIE

Master: Nicolas Anciaux, Courses on database internal mechanisms and database security, 80, in Master1 and Master2 (AFTI, Orsay) and in engineering school (ENSTA ParisTech, Telecom Paristech)

Master : Luc Bouganim, Bases de données, 24, niveau M2, CFA AFTI/UVSQ, France

Master : Luc Bouganim, Systèmes d'information Privacy by Design, 54, niveau M2, ENSIIE, INSA CVL, France

Master : Luc Bouganim, Sécurité des bases de données, 10, niveau M2, Télécom ParisTech, France

Licence : Iulian Sandu Popa, Initiation aux bases de données (niveau L2), Bases de données (niveau L3), 96, UVSQ, France

Master : Iulian Sandu Popa, Bases de données relationnelles (niveau M1), Gestion des données spatiotemporelles (niveau M2), Sécurité des bases de données (niveau M2), 96, UVSQ, France

Master : Benjamin Nguyen, Databases, IA, Security, 192, INSA CVL, France

Licence : Guillaume Scerri, Fondements de l'informatique, 36, niveau L1, UVSQ, France

Licence : Guillaume Scerri, Initiation aux bases de données, 43.5, niveau L2, UVSQ, France

Master : Guillaume Scerri, Bases de données, 30, niveau M1, UVSQ France

E-learning

MOOC created in 2016: "Villes intelligentes : défis technologiques et sociétaux", coordinated by V. Issarny and N. Mitton. N. Anciaux, S.Grumbach, V. Issarny, N. Mitton, C. Morin, A. Pathak, H. Rivano. Start date 25 Jan. 2016. In 2016, more than 8000 students have registered to follow the courses. A new edition is planned for 2017. https://www.fun-mooc.fr/courses/inria/41009/session01/about

MOOC created in 2016: Bases de données relationnelles : comprendre pour maîtriser (Serge Abiteboul, Benjamin Nguyen, Philippe Rigaux) https://www.fun-mooc.fr/courses/inria/41008/session01/about

9.2.2. Supervision

PhD: Athanasia Katsouraki, Sharing and Usage Control of Personal Information, UVSQ, September 2016, Luc Bouganim and Benjamin Nguyen

PhD: Saliha Lallali, A Secure Search Engine for the Personal Cloud, UVSQ, January 2016, Nicolas Anciaux, Philippe Pucheral, and Iulian Sandu Popa

PhD: Dai Hai Ton That, Efficient Management and Secure Sharing of Mobility Traces, UVSQ, January 2016, Iulian Sandu Popa and Karine Zeitouni

PhD in progress : Paul Tran Van, Partage de documents sécurisé dans le Cloud Personnel, October 2014, Nicolas Anciaux and Philippe Pucheral

PhD in progress : Axel Michel, Secure Distributed Computations, October 2015, Benjamin Nguyen and Philippe Pucheral

PhD in progress : Julien Loudet, Personal Queries on Personal Clouds, July 2016, Luc Bouganim and Iulian Sandu Popa

PhD in progress : Riad Ladjel, Secure Distributed Computation for the Personal Cloud, October 2016, Nicolas Anciaux and Philippe Pucheral

9.2.3. Juries

Philippe Pucheral: member of the PhD jury of Mahsa Najafzadeh (Paris VI, 22/04/2016)

Benjamin Nguyen: reviewer of the PhD of Tarek SAYAH, Exposition sélective et problème de fuite d'inférence dans le Linked Data, Université Claude Bernard Lyon I, 2016/09/08

Benjamin Nguyen: reviewer of the PhD of Germain JOLLY, Evaluation of payment applications on smart cards, Université de Caen/ENSICAEN, 2016/07/08

Benjamin Nguyen: reviewer of the PhD of Julien LOLIVE, Entrelacement des mécanismes d'identification et de respect de la vie privée pour la protection des contenus externalisés, Télécom Bretagne, 2016/05/13

Benjamin Nguyen: president and reviewer of the PhD jury of Karina SOKOLOVA, Bridging the gap between Privacy by Design and mobile systems by patterns, Université Technologique de Troyes, 2016/04/27

9.3. Popularization

General public (large audience magazines, television, videos):

- "Santé connectée. Une médecine sans médecin ?", interview of N. Anciaux, La Recherche, N°510 Avril 2016. (http://fliphtml5.com/ggwf/ynja p.84)
- "PlugDB, coffre-fort numérique personnel", interview of P.Pucheral, Le Monde Science & Techno, November 28th, 2016.
- "Sur Internet, vos données personnelles valent de l'or", interview of P.Pucheral, Ca m'intéresse, November 2016.
- "Santé connectée, santé partagée", interview of P.Pucheral, Le Point, April 7th 2016.

10. Bibliography

Major publications by the team in recent years

[1] T. ALLARD, B. NGUYEN, P. PUCHERAL.*MetaP: Revisiting Privacy-Preserving Data Publishing using Secure Devices*, in "Distributed and Parallel Databases", June 2014, vol. 32, n° 1, p. 191-244 [*DOI :* 10.1007/S10619-013-7122-X], https://hal.archives-ouvertes.fr/hal-00934586.

[2] N. ANCIAUX, P. BONNET, L. BOUGANIM, B. NGUYEN, P. PUCHERAL, I. SANDU POPA.*Trusted Cells : A Sea Change for Personnal Data Services*, in "CIDR 2013 - 6th Biennal Conference on Innovative Database Research", Asilomar, United States, 2013, 4, http://hal.inria.fr/hal-00768379.

[3] N. ANCIAUX, L. BOUGANIM, T. DELOT, S. ILARRI, L. KLOUL, N. MITTON, P. PUCHERAL.*Folk-IS: Opportunistic Data Services in Least Developed Countries*, in "40th International Conference on Very Large Data Bases (VLDB)", Hangzhou, China, Zhejiang University, September 2014, https://hal.inria.fr/hal-00906204.

[4] N. ANCIAUX, L. BOUGANIM, P. PUCHERAL, Y. GUO, L. LE FOLGOC, S. YIN.*MILo-DB: a personal, secure and portable database machine*, in "Distributed and Parallel Databases", March 2014, vol. 32, n° 1, p. 37-63 [*DOI :* 10.1007/S10619-012-7119-X], https://hal.archives-ouvertes.fr/hal-00768355.

[5] N. ANCIAUX, S. LALLALI, I. SANDU POPA, P. PUCHERAL.*A Scalable Search Engine for Mass Storage Smart Objects*, in "Proceedings of the 41th International Conference on Very Large Databases (VLDB)", Kohala Coast, Hawaii, United States, August 2015, vol. 8, n⁰ 9, p. 910-921 [*DOI : 10.14778/2777598.2777600*], https://hal.inria.fr/hal-01176458.

[6] N. ANCIAUX, B. NGUYEN, I. SANDU POPA.*Tutorial: Managing Personal Data with Strong Privacy Guarantees*, March 2014, p. 672-673 [*DOI : 10.5441/002/EDBT.2014.71*], https://hal.inria.fr/hal-01096633.

[7] G. SCERRI, B. WARINSCHI, M. BARBOSA, B. PORTELA.*Foundations of Hardware-Based Attested Computation and Application to SGX*, March 2016, p. 245-260 [*DOI : 10.1109/EUROSP.2016.28*], https://hal.inria.fr/hal-01417137.

[8] Q. C. TO, B. NGUYEN, P. PUCHERAL.*Privacy-Preserving Query Execution using a Decentralized Architecture and Tamper Resistant Hardware*, in "17th International Conference on Extending Database Technology (EDBT)", Athens, Greece, March 2014 [*DOI : 10.5441/002/EDBT.2014.44*], https://hal.inria.fr/hal-01096639.

[9] C. Q. TO, B. NGUYEN, P. PUCHERAL.*TrustedMR: A Trusted MapReduce System based on Tamper Resistance Hardware* , in "Proceedings of the 23rd International Conference on Cooperative Information Systems (COOPIS)", Rhodes, Greece, October 2015, p. 38-56 [*DOI : 10.1007/978-3-319-26148-5_3*], https://hal.inria.fr/hal-01254951.

[10] D. H. TON THAT, I. SANDU POPA, K. ZEITOUNI.*TRIFL: A Generic Trajectory Index for Flash Storage*, in "ACM Transactions on Algorithms", July 2015, vol. 1, n⁰ 2, 44 [*DOI : 10.1145/2786758*], https://hal.inria.fr/hal-01176563.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] A. KATSOURAKI.*Sharing and Usage Control of Personal Information*, Université Paris Saclay, September 2016, https://hal.archives-ouvertes.fr/tel-01425638.

[12] S. LALLALI.*A Scalable Search Engine for the Personal Cloud*, Université Paris-Saclay, January 2016, https://hal.inria.fr/tel-01426486.

[13] D. H. TON THAT.*Efficient management and secure sharing of mobility traces*, Université Paris-Saclay, January 2016, https://tel.archives-ouvertes.fr/tel-01290834.

Articles in International Peer-Reviewed Journal

[14] S. J. PAN, I. SANDU POPA, C. BORCEA.*DIVERT: A Distributed Vehicular Traffic Re-Routing System for Congestion Avoidance*, in "IEEE Transactions on Mobile Computing", January 2017, vol. 16, n⁰ 1, p. 58-72 [*DOI : 10.1109/TMC.2016.2538226*], https://hal.inria.fr/hal-01426424.

[15] C. Q. TO, B. NGUYEN, P. PUCHERAL.*Private and Scalable Execution of SQL Aggregates on a Secure Decentralized Architecture*, in "ACM Transactions on Database Systems", 2016, https://hal.archives-ouvertes.fr/hal-01296432.

[16] D. H. TON THAT, I. SANDU POPA, K. ZEITOUNI, C. BORCEA.*Un protocole basé sur des mobiles sécurisés pour une collecte participative de données spatiales en mobilité réellement anonyme*, in "Revue Internationale de Géomatique", August 2016, vol. 26, n⁰ 2, p. 185-210 [*DOI :* 10.3166/RIG.26.185-210], https://hal.inria.fr/hal-01426358.

Articles in National Peer-Reviewed Journal

[17] N. ANCIAUX, P. PUCHERAL, M. BEHAR-TOUCHAIS, V.-L. BENABOU, G. BRUNAUX, A. LEFEVRE, N. MARTIAL-BRAZ, J. ROCHFELD, N. SAUPHANOR-BROUILLAUD, B. SCHULZ, J. SENECHAL, C. ZOLYNSKI.*Dossier Contenus Numériques Revue Contrats, Concurrence, Consommation - Contenus Numériques*, in "Revue Contrats Concurrence Consommation", February 2017, https://hal.archives-ouvertes.fr/hal-01432544.

[18] C. BERTHET, C. ZOLYNSKI, N. ANCIAUX, P. PUCHERAL.*" Contenus numériques et récupération des données : un nouvel outil d' 'empouvoirement' du consommateur ? "*, in "Dalloz IP/IT", January 2017, vol. IP IT / 10, https://hal.inria.fr/hal-01429939.

[19] P. PUCHERAL, N. ANCIAUX, M. BEHAR-TOUCHAIS, V.-L. BENABOU, N. MARTIAL-BRAZ, J. ROCHFELD, N. SAUPHANOR-BROUILLAUD, B. SCHULZ, J. SENECHAL, C. ZOLYNSKI.*Dossier Contenus Numériques / Données*, in "Contrats concurrence consommation", February 2017, https://hal.inria.fr/hal-01429951.

[20] P. PUCHERAL, A. RALLET, F. ROCHELANDET, C. ZOLYNSKI.*La Privacy by design : une fausse bonne solution aux problèmes de protection des données personnelles soulevés par l'Open data et les objets connectés ?*, in "Legicom", January 2016, n⁰ 56, p. 89-99 [*DOI :* 10.3917/LEGI.056.0089], https://hal.archives-ouvertes.fr/hal-01427983.

International Conferences with Proceedings

[21] L. BOUGANIM, A. KATSOURAKI, B. NGUYEN.*DatShA :A Data Sharing Algebra for access control plans*, in "19th International Conference on Extending Database Technology (EDBT 2016)", Bordeaux, France, Proceedings of the 19th International Conference on Extending Database Technology (short paper), March 2016, https://hal.archives-ouvertes.fr/hal-01289023.

[22] T. B. T. LE, N. ANCIAUX, S. GILLOTON, S. LALLALI, P. PUCHERAL, I. SANDU POPA, C. CHEN.*Distributed Secure Search in the Personal Cloud*, in "19th International Conference on Extending Database Technology (EDBT 2016)", Bordeaux, France, Proceedings of EDBT'16 (Demo paper), March 2016, p. 652-655, https://hal.inria.fr/hal-01293409.

[23] D. H. TON THAT, I. SANDU POPA, K. ZEITOUNI, C. BORCEA.*PAMPAS: Privacy-Aware Mobile Participatory Sensing Using Secure Probes*, in "International Conference on Scientific and Statistical Database Management (SSDBM '16)", Budapest, Hungary, Proceedings of the 28th International Conference on Scientific and Statistical Database Management, July 2016 [*DOI :* 10.1145/1235], https://hal.inria.fr/hal-01426375.

Conferences without Proceedings

[24] A. MICHEL, B. NGUYEN.*Managing distributed queries under anonymity constraints*, in "7e Atelier sur la Protection de la Vie Privée", Toulouse, France, July 2016, https://hal.archives-ouvertes.fr/hal-01424993.

[25] A. MICHEL, B. NGUYEN, P. PUCHERAL.*Exécution de requêtes distribuées sous contraintes d'anonymat*, in "32ème Conférence sur la Gestion de Données - Principes, Technologies et Applications", Poitiers, France, November 2016, https://hal.archives-ouvertes.fr/hal-01424989.

Scientific Books (or Scientific Book chapters)

[26] P. PUCHERAL, A. RALLET, C. ZOLYNSKI.*Privacy by Design et Big Data*, in "Les big data à découvert", 2016, https://hal.archives-ouvertes.fr/hal-01429075.

Scientific Popularization

[27] N. ANCIAUX, B. ANDRÉ, P. PUCHERAL, P. TRAN-VAN.*A Root of Trust for the Personal Cloud*, in "ERCIM News", July 2016, https://hal.archives-ouvertes.fr/hal-01427725.

[28] N. ANCIAUX, L. BOUGANIM.*Decentralized Personal Data Management, using secure devices (Tutorial)*, June 2016, BDA summer school: "Masses de données distribuées", Urugne, France, https://hal.archives-ouvertes.fr/hal-01425656.

References in notes

[29] R. AGRAWAL, J. KIERNAN, R. SRIKANT, Y. XU.*Hippocratic Databases*, in "Proc. of the Int. Conf. on Very Large Data Bases (VLDB)", 2002.

[30] T. ALLARD, N. ANCIAUX, L. BOUGANIM, Y. GUO, L. LE FOLGOC, B. NGUYEN, P. PUCHERAL, I. RAY, I. RAY, S. YIN.*Secure Personal Data Servers: a Vision Paper*, in "Proc. of the 36th Int. Conf. on Very Large Databases (VLDB)", 2010.

[31] N. ANCIAUX, B. NGUYEN, I. SANDU POPA.*Towards an Era of Trust in Personal Data Management*, in "Proceedings of the 19th East-European Conference on Advances in Databases and Information Systems (ADBIS '15). Tutorial", Poitiers, France, 2015, https://hal.inria.fr/hal-01176512.

[32] A. ANDERSON.*An introduction to the web services policy language (WSPL)*, in "IEEE Computer Society", 2004.

[33] R. BAHMANI, M. BARBOSA, F. BRASSER, B. PORTELA, A. SADEGHI, G. SCERRI, B. WARINSCHI.*Secure Multiparty Computation from SGX*, in "IACR Cryptology ePrint Archive", 2016, vol. 2016, 1057.

[34] P. BONNET, L. BOUGANIM, I. KOLTSIDAS, S. VIGLAS.*System Co-Design and Data Management for Flash Devices*, in "Very Large Data Bases (Tutorial)", 2011.

[35] L. BOUGANIM, Y. GUO.*Database Encryption*, in "Encyclopedia of Cryptography and Security", S. JAJODIA, H. VAN TILBORG (editors), Springer, 2009, p. 307-312.

[36] L. CRANOR.*Web Privacy with P3P*, O'Reilly Media, 2002.

[37] B. FUNG, K. WANG, R. CHEN, P. YU.*Privacy-preserving data publishing: A survey of recent developments*, in "ACM Computing Surveys (CSUR)", 2010, vol. 42, n[o] 4.

[38] S. LALLALI, N. ANCIAUX, I. SANDU POPA, P. PUCHERAL.*A Secure Search Engine for the Personal Cloud*, in "Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data (SIGMOD '15). Demo paper", Melbourne, Australia, 2015, p. 1445-1450 [*DOI :* 10.1145/2723372.2735376], https://hal.inria.fr/hal-01176473.

[39] T. MOSES.*Extensible access control markup language (XACML) version 2.0*, in "Oasis Standard 200502", 2005.

[40] M. SCHUNTER, C. POWERS.*Enterprise privacy authorization language (EPAL 1.1)*, in "IBM", 2003.

[41] D. H. TON THAT, I. SANDU POPA, K. ZEITOUNI.*PPTM: Privacy-aware Participatory Traffic Monitoring Using Mobile Secure Probes*, in "Proceedings of the 16th IEEE International Conference on Mobile Data Management (MDM '15). Demo paper", Pittsburgh, United States, 2015, 4, https://hal.inria.fr/hal-01176486.

Project-Team SPECFUN

Symbolic Special Functions : Fast and Certified

Table of contents

Project-Team SPECFUN

Creation of the Team: 2012 November 01, updated into Project-Team: 2014 July 01

Keywords:

Computer Science and Digital Science:

2.1.10. - Domain-specific languages
2.1.11. - Proof languages
2.4.3. - Proofs
4.5. - Formal methods for security
7.2. - Discrete mathematics, combinatorics
7.4. - Logic in Computer Science
7.6. - Computer Algebra
7.7. - Number theory
7.11. - Performance evaluation

Other Research Topics and Application Domains:

9.4.2. - Mathematics
9.4.3. - Physics

1. Members

Research Scientists

Frédéric Chyzak [Team leader, Inria, Researcher, HDR]
Alin Bostan [Inria, Researcher]
Philippe Dumas [Inria, Researcher]
Georges Gonthier [Inria, Senior Researcher, since October 2016]
Pierre Lairez [Inria, Researcher, since October 2016]
Assia Mahboubi [Team co-leader, Inria, Researcher]

Technical Staff

Maxence Guesdon [Inria, 40%]

PhD Students

Louis Dumont [École Polytechnique, until October 2016]
Thomas Sibut-Pinote [École Polytechnique]

Post-Doctoral Fellow

Suzy S. Maddah [Inria]

Visiting Scientist

Marc Mezzarobba [CNRS]

Administrative Assistant

Christine Biard [Inria]

2. Overall Objectives

2.1. Scientific challenges, expected impact

The general orientation of our team is described by the short name given to it: *Special Functions*, that is, particular mathematical functions that have established names due to their importance in mathematical analysis, physics, and other application domains. Indeed, we ambition to study special functions with the computer, by combined means of computer algebra and formal methods.

Computer-algebra systems have been advertised for decades as software for "doing mathematics by computer" [72]. For instance, computer-algebra libraries can uniformly generate a corpus of mathematical properties about special functions, so as to display them on an interactive website. This possibility was recently shown by the computer-algebra component of the team [26]. Such an automated generation significantly increases the reliability of the mathematical corpus, in comparison to the content of existing static authoritative handbooks. The importance of the validity of these contents can be measured by the very wide audience that such handbooks have had, to the point that a book like [21] remains one of the most cited mathematical publications ever and has motivated the 10-year-long project of writing its successor [23]. However, can the mathematics produced "by computer" be considered as *true* mathematics? More specifically, whereas it is nowadays well established that the computer helps in discovering and observing new mathematical phenomenons, can the mathematical statements produced with the aid of the computer and the mathematical results computed by it be accepted as valid mathematics, that is, as having the status of mathematical *proofs*? Beyond the reported weaknesses or controversial design choices of mainstream computer-algebra systems, the issue is more of an epistemological nature. It will not find its solution even in the advent of the ultimate computer-algebra system: the social process of peer-reviewing just falls short of evaluating the results produced by computers, as reported by Th. Hales [50] after the publication of his proof of the Kepler Conjecture about sphere packing.

A natural answer to this deadlock is to move to an alternative kind of mathematical software and to use a proof assistant to check the correctness of the desired properties or formulas. The success of large-scale formalization projects, like the Four-Color Theorem of graph theory [45], the above-mentioned Kepler Conjecture [50], and the Odd Order Theorem of group theory [0], have increased the understanding of the appropriate software-engineering methods for this peculiar kind of programming. For computer algebra, this legitimates a move to proof assistants now.

The Dynamic Dictionary of Mathematical Functions [0] (DDMF) [26] is an online computer-generated handbook of mathematical functions that ambitions to serve as a reference for a broad range of applications. This software was developed by the computer-algebra component of the team as a project [0] of the MSR–INRIA Joint Centre. It bases on a library for the computer-algebra system Maple, Algolib [0], whose development started 20 years ago in ÉPI Algorithms [0]. As suggested by the constant questioning of certainty by new potential users, DDMF deserves a formal guarantee of correctness of its content, on a level that proof assistants can provide. Fortunately, the maturity of special-functions algorithms in Algolib makes DDMF a stepping stone for such a formalization: it provides a well-understood and unified algorithmic treatment, without which a formal certification would simply be unreachable.

The formal-proofs component of the team emanates from another project of the MSR–INRIA Joint Centre, namely the Mathematical Components project (MathComp) [0]. Since 2006, the MathComp group has endeavoured to develop computer-checked libraries of formalized mathematics, using the Coq proof assistant [68]. The methodological aim of the project was to understand the design methods leading to successful large-scale formalizations. The work culminated in 2012 with the completion of a formal proof of the Odd Order Theorem, resulting in the largest corpus of algebraic theories ever machine-checked with a proof assistant and a whole methodology to effectively combine these components in order to tackle complex formalizations. In particular, these libraries provide a good number of the many algebraic objects needed to reason about special functions and their properties, like rational numbers, iterated sums, polynomials, and a rich hierarchy of algebraic structures.

The present team takes benefit from these recent advances to explore the formal certification of the results collected in DDMF. The aim of this project is to concentrate the formalization effort on this delimited area, building on DDMF and the Algolib library, as well as on the Coq system [68] and on the libraries developed by the MathComp project.

---

[0] http://www.msr-inria.inria.fr/news/the-formalization-of-the-odd-order-theorem-has-been-completed-the-20-septembre-2012/
[0] http://ddmf.msr-inria.inria.fr/1.9.1/ddmf
[0] http://www.msr-inria.inria.fr/projects/dynamic-dictionary-of-mathematical-functions/
[0] http://algo.inria.fr/libraries/
[0] http://algo.inria.fr/
[0] http://www.msr-inria.fr/projects/mathematical-components/

2.1.1. *Use computer algebra but convince users beyond reasonable doubt*

The following few opinions on computer algebra are, we believe, typical of computer-algebra users' doubts and difficulties when using computer-algebra systems:

- Fredrik Johansson, expert in the multi-precision numerical evaluation of special functions and in fast computer-algebra algorithms, writes on his blog [56]: "Mathematica is great for cross-checking numerical values, but it's not unusual to run into bugs, so *triple checking is a good habit*." One answer in the discussion is: "We can claim that Mathematica has [...] *an impossible to understand semantics*: If Mathematica's output is wrong then change the input. If you don't like the answer, change the question. That seems to be the philosophy behind."

- A professor's advice to students [64] on using Maple: "You may wish to use Maple to check your homework answers. If you do then keep in mind that Maple sometimes gives the *wrong answer, usually because you asked incorrectly, or because of niceties of analytic continuation*. You may even be bitten by an occasional Maple bug, though that has become fairly unlikely. Even with as powerful a tool as Maple you will still *have to devise your own checks* and you will still have to think."

- Jacques Carette, former head of the maths group at Maplesoft, about a bug [22] when asking Maple to take the limit `limit(f(n) * exp(-n), n = infinity)` for an undetermined function `f`: "The problem is that there is an *implicit assumption in the implementation* that unknown functions do not 'grow too fast'."

As explained by the expert views above, complaints by computer-algebra users are often due to their misunderstanding of what a computer-algebra systems is, namely a purely syntactic tool for calculations, that the user must complement with a semantics. Still, robustness and consistency of computer-algebra systems are not ensured as of today, and, whatever Zeilberger may provocatively say in his Opinion 94 [73], a firmer logical foundation is necessary. Indeed, the fact is that many "bugs" in a computer-algebra system cannot be fixed by just the usual debugging method of tracking down the faulty lines in the code. It is sort of "by design": assumptions that too often remain implicit are really needed by the design of symbolic algorithms and cannot easily be expressed in the programming languages used in computer algebra. A similar certification initiative has already been undertaken in the domain of numerical computing, in a successful manner [54], [29]. It is natural to undertake a similar approach for computer algebra.

2.1.2. *Make computer algebra and formal proofs help one another*

Some of the mathematical objects that interest our team are still totally untouched by formalization. When implementing them and their theory inside a proof assistant, we have to deal with the pervasive discrepancy between the published literature and the actual implementation of computer-algebra algorithms. Interestingly, this forces us to clarify our computer-algebraic view on them, and possibly make us discover holes lurking in published (human) proofs. We are therefore convinced that the close interaction of researchers from both fields, which is what we strive to maintain in this team, is a strong asset.

For a concrete example, the core of Zeilberger's creative telescoping manipulates rational functions up to simplifications. In summation applications, checking that these simplifications do not hide problematic divisions by 0 is most often left to the reader. In the same vein, in the case of integrals, the published algorithms do not check the convergence of all integrals, especially in intermediate calculations. Such checks are again left to the readers. In general, we expect to revisit the existing algorithms to ensure that they are meaningful for genuine mathematical sequences or functions, and not only for algebraic idealizations.

Another big challenge in this project originates in the scientific difference between computer algebra and formal proofs. Computer algebra seeks speed of calculation on *concrete instances* of algebraic data structures (polynomials, matrices, etc). For their part, formal proofs manipulate symbolic expressions in terms of *abstract variables* understood to represent generic elements of algebraic data structures. In view of this, a continuous challenge is to develop the right, hybrid thinking attitude that is able to effectively manage concrete and abstract values simultaneously, alternatively computing and proving with them.

2.1.3. Experimental mathematics with special functions

Applications in combinatorics and mathematical physics frequently involve equations of so high orders and so large sizes, that computing or even storing all their coefficients is impossible on existing computers. Making this tractable is an extraordinary challenge. The approach we believe in is to design algorithms of good—ideally quasi-optimal—complexity in order to extract precisely the required data from the equations, while avoiding the computationally intractable task of completely expanding them into an explicit representation.

Typical applications with expected high impact are the automatic discovery and algorithmic proof of results in combinatorics and mathematical physics for which human proofs are currently unattainable.

2.2. Research axes

The implementation of certified symbolic computations on special functions in the Coq proof assistant requires both investigating new formalization techniques and renewing the traditional computer-algebra viewpoint on these standard objects. Large mathematical objects typical of computer algebra occur during formalization, which also requires us to improve the efficiency and ergonomics of Coq. In order to feed this interdisciplinary activity with new motivating problems, we additionally pursue a research activity oriented towards experimental mathematics in application domains that involve special functions. We expect these applications to pose new algorithmic challenges to computer algebra, which in turn will deserve a formal-certification effort. Finally, DDMF is the motivation and the showcase of our progress on the certification of these computations. While striving to provide a formal guarantee of the correctness of the information it displays, we remain keen on enriching its mathematical content by developing new computer-algebra algorithms.

2.2.1. Computer algebra certified by the Coq system

Our formalization effort consists in organizing a cooperation between a computer-algebra system and a proof assistant. The computer-algebra system is used to produce efficiently algebraic data, which are later processed by the proof assistant. The success of this cooperation relies on the design of appropriate libraries of formalized mathematics, including certified implementations of certain computer-algebra algorithms. On the other side, we expect that scrutinizing the implementation and the output of computer-algebra algorithms will shed a new light on their semantics and on their correctness proofs, and help clarifying their documentation.

2.2.1.1. Libraries of formalized mathematics

The appropriate framework for the study of efficient algorithms for special functions is *algebraic*. Representing algebraic theories as Coq formal libraries takes benefit from the methodology emerging from the success of ambitious projects like the formal proof of a major classification result in finite-group theory (the Odd Order Theorem) [43].

Yet, a number of the objects we need to formalize in the present context has never been investigated using any interactive proof assistant, despite being considered as commonplaces in computer algebra. For instance there is up to our knowledge no available formalization of the theory of non-commutative rings, of the algorithmic theory of special-functions closures, or of the asymptotic study of special functions. We expect our future formal libraries to prove broadly reusable in later formalizations of seemingly unrelated theories.

2.2.1.2. Manipulation of large algebraic data in a proof assistant

Another peculiarity of the mathematical objects we are going to manipulate with the Coq system is their size. In order to provide a formal guarantee on the data displayed by DDMF, two related axes of research have to be pursued. First, efficient algorithms dealing with these large objects have to be programmed and run in Coq. Recent evolutions of the Coq system to improve the efficiency of its internal computations [24], [27] make this objective reachable. Still, how to combine the aforementioned formalization methodology with these cutting-edge evolutions of Coq remains one of the prospective aspects of our project. A second need is to help users *interactively* manipulate large expressions occurring in their conjectures, an objective for which little has been done so far. To address this need, we work on improving the ergonomics of the system in two ways:

first, ameliorating the reactivity of Coq in its interaction with the user; second, designing and implementing extensions of its interface to ease our formalization activity. We expect the outcome of these lines of research to be useful to a wider audience, interested in manipulating large formulas on topics possibly unrelated to special functions.

2.2.1.3. Formal-proof-producing normalization algorithms

Our algorithm certifications inside Coq intend to simulate well-identified components of our Maple packages, possibly by reproducing them in Coq. It would however not have been judicious to re-implement them inside Coq in a systematic way. Indeed for a number of its components, the output of the algorithm is more easily checked than found, like for instance the solving of a linear system. Rather, we delegate the discovery of the solutions to an external, untrusted oracle like Maple. Trusted computations inside Coq then formally validate the correctness of the a priori untrusted output. More often than not, this validation consists in implementing and executing normalization procedures *inside* Coq. A challenge of this automation is to make sure they go to scale while remaining efficient, which requires a Coq version of non-trivial computer-algebra algorithms. A first, archetypal example we expect to work on is a non-commutative generalization of the normalization procedure for elements of rings [49].

2.2.2. Better symbolic computations with special functions

Generally speaking, we design algorithms for manipulating special functions symbolically, whether univariate or with parameters, and for extracting algorithmically any kind of algebraic and analytic information from them, notably asymptotic properties. Beyond this, the heart of our research is concerned with parametrised definite summations and integrations. These very expressive operations have far-ranging applications, for instance, to the computation of integral transforms (Laplace, Fourier) or to the solution of combinatorial problems expressed via integrals (coefficient extractions, diagonals). The algorithms that we design for them need to really operate on the level of linear functional systems, differential and of recurrence. In all cases, we strive to design our algorithms with the constant goal of good theoretical complexity, and we observe that our algorithms are also fast in practice.

2.2.2.1. Special-function integration and summation

Our long-term goal is to design fast algorithms for a general method for special-function integration (*creative telescoping*), and make them applicable to general special-function inputs. Still, our strategy is to proceed with simpler, more specific classes first (rational functions, then algebraic functions, hyperexponential functions, D-finite functions, non-D-finite functions; two variables, then many variables); as well, we isolate analytic questions by first considering types of integration with a more purely algebraic flavor (constant terms, algebraic residues, diagonals of combinatorics). In particular, we expect to extend our recent approach [32] to more general classes (algebraic with nested radicals, for example): the idea is to speed up calculations by making use of an analogue of Hermite reduction that avoids considering certificates. Homologous problems for summation will be addressed as well.

2.2.2.2. Applications to experimental mathematics

As a consequence of our complexity-driven approach to algorithms design, the algorithms mentioned in the previous paragraph are of good complexity. Therefore, they naturally help us deal with applications that involve equations of high orders and large sizes.

With regard to combinatorics, we expect to advance the algorithmic classification of combinatorial classes like walks and urns. Here, the goal is to determine if enumerative generating functions are rational, algebraic, or D-finite, for example. Physical problems whose modelling involves special-function integrals comprise the study of models of statistical mechanics, like the Ising model for ferro-magnetism, or questions related to Hamiltonian systems.

Number theory is another promising domain of applications. Here, we attempt an experimental approach to the automated certification of integrality of the coefficients of mirror maps for Calabi–Yau manifolds. This could also involve the discovery of new Calabi–Yau operators and the certification of the existing ones. We also plan to algorithmically discover and certify new recurrences yielding good approximants needed in irrationality proofs.

It is to be noted that in all of these application domains, we would so far use general algorithms, as was done in earlier works of ours [31], [36], [34]. To push the scale of applications further, we plan to consider in each case the specifics of the application domain to tailor our algorithms.

2.2.3. *Interactive and certified mathematical web sites*

In continuation of our past project of an encyclopedia at http://ddmf.msr-inria.inria.fr/1.9.1/ddmf, we ambition to both enrich and certify the formulas about the special functions that we provide online. For each function, our website shows its essential properties and the mathematical objects attached to it, which are often infinite in nature (numerical evaluations, asymptotic expansions). An interactive presentation has the advantage of allowing for adaption to the user's needs. More advanced content will broaden the encyclopedia:

- the algorithmic discussion of equations with parameters, leading to certified automatic case analysis based on arithmetic properties of the parameters;
- lists of summation and integral formulas involving special functions, including validity conditions on the parameters;
- guaranteed large-precision numerical evaluations.

3. Research Program

3.1. Studying special functions by computer algebra

Computer algebra manipulates symbolic representations of exact mathematical objects in a computer, in order to perform computations and operations like simplifying expressions and solving equations for "closed-form expressions". The manipulations are often fundamentally of algebraic nature, even when the ultimate goal is analytic. The issue of efficiency is a particular one in computer algebra, owing to the extreme swell of the intermediate values during calculations.

Our view on the domain is that research on the algorithmic manipulation of special functions is anchored between two paradigms:

- adopting linear differential equations as the right data structure for special functions,
- designing efficient algorithms in a complexity-driven way.

It aims at four kinds of algorithmic goals:

- algorithms combining functions,
- functional equations solving,
- multi-precision numerical evaluations,
- guessing heuristics.

This interacts with three domains of research:

- computer algebra, meant as the search for quasi-optimal algorithms for exact algebraic objects,
- symbolic analysis/algebraic analysis;
- experimental mathematics (combinatorics, mathematical physics, ...).

This view is made explicit in the present section.

3.1.1. *Equations as a data structure*

Numerous special functions satisfy linear differential and/or recurrence equations. Under a mild technical condition, the existence of such equations induces a finiteness property that makes the main properties of the functions decidable. We thus speak of *D-finite functions*. For example, 60 % of the chapters in the handbook [21] describe D-finite functions. In addition, the class is closed under a rich set of algebraic operations. This makes linear functional equations just the right data structure to encode and manipulate special functions. The power of this representation was observed in the early 1990s [74], leading to the design of many algorithms in computer algebra. Both on the theoretical and algorithmic sides, the study of D-finite functions shares much with neighbouring mathematical domains: differential algebra, D-module theory, differential Galois theory, as well as their counterparts for recurrence equations.

3.1.2. Algorithms combining functions

Differential/recurrence equations that define special functions can be recombined [74] to define: additions and products of special functions; compositions of special functions; integrals and sums involving special functions. Zeilberger's fast algorithm for obtaining recurrences satisfied by parametrised binomial sums was developed in the early 1990s already [75]. It is the basis of all modern definite summation and integration algorithms. The theory was made fully rigorous and algorithmic in later works, mostly by a group in RISC (Linz, Austria) and by members of the team [63], [71], [39], [37], [38], [57]. The past ÉPI Algorithms contributed several implementations (*gfun* [66], *Mgfun* [39]).

3.1.3. Solving functional equations

Encoding special functions as defining linear functional equations postpones some of the difficulty of the problems to a delayed solving of equations. But at the same time, solving (for special classes of functions) is a sub-task of many algorithms on special functions, especially so when solving in terms of polynomial or rational functions. A lot of work has been done in this direction in the 1990s; more intensively since the 2000s, solving differential and recurrence equations in terms of special functions has also been investigated.

3.1.4. Multi-precision numerical evaluation

A major conceptual and algorithmic difference exists for numerical calculations between data structures that fit on a machine word and data structures of arbitrary length, that is, *multi-precision* arithmetic. When multi-precision floating-point numbers became available, early works on the evaluation of special functions were just promising that "most" digits in the output were correct, and performed by heuristically increasing precision during intermediate calculations, without intended rigour. The original theory has evolved in a twofold way since the 1990s: by making computable all constants hidden in asymptotic approximations, it became possible to guarantee a *prescribed* absolute precision; by employing state-of-the-art algorithms on polynomials, matrices, etc, it became possible to have evaluation algorithms in a time complexity that is linear in the output size, with a constant that is not more than a few units. On the implementation side, several original works exist, one of which (*NumGfun* [62]) is used in our DDMF.

3.1.5. Guessing heuristics

"Differential approximation", or "Guessing", is an operation to get an ODE likely to be satisfied by a given approximate series expansion of an unknown function. This has been used at least since the 1970s and is a key stone in spectacular applications in experimental mathematics [36]. All this is based on subtle algorithms for Hermite–Padé approximants [25]. Moreover, guessing can at times be complemented by proven quantitative results that turn the heuristics into an algorithm [33]. This is a promising algorithmic approach that deserves more attention than it has received so far.

3.1.6. Complexity-driven design of algorithms

The main concern of computer algebra has long been to prove the feasibility of a given problem, that is, to show the existence of an algorithmic solution for it. However, with the advent of faster and faster computers, complexity results have ceased to be of theoretical interest only. Nowadays, a large track of works in computer algebra is interested in developing fast algorithms, with time complexity as close as possible to linear in their output size. After most of the more pervasive objects like integers, polynomials, and matrices have been endowed with fast algorithms for the main operations on them [44], the community, including ourselves, started to turn its attention to differential and recurrence objects in the 2000s. The subject is still not as developed as in the commutative case, and a major challenge remains to understand the combinatorics behind summation and integration. On the methodological side, several paradigms occur repeatedly in fast algorithms: "divide and conquer" to balance calculations, "evaluation and interpolation" to avoid intermediate swell of data, etc. [30].

3.2. Trusted computer-algebra calculations

3.2.1. Encyclopedias

Handbooks collecting mathematical properties aim at serving as reference, therefore trusted, documents. The decision of several authors or maintainers of such knowledge bases to move from paper books [21], [23], [67] to websites and wikis [0] allows for a more collaborative effort in proof reading. Another step toward further confidence is to manage to generate the content of an encyclopedia by computer-algebra programs, as is the case with the Wolfram Functions Site [0] or DDMF [0]. Yet, due to the lingering doubts about computer-algebra systems, some encyclopedias propose both cross-checking by different systems and handwritten companion paper proofs of their content [0]. As of today, there is no encyclopedia certified with formal proofs.

3.2.2. Computer algebra and symbolic logic

Several attempts have been made in order to extend existing computer-algebra systems with symbolic manipulations of logical formulas. Yet, these works are more about extending the expressivity of computer-algebra systems than about improving the standards of correctness and semantics of the systems. Conversely, several projects have addressed the communication of a proof system with a computer-algebra system, resulting in an increased automation available in the proof system, to the price of the uncertainty of the computations performed by this oracle.

3.2.3. Certifying systems for computer algebra

More ambitious projects have tried to design a new computer-algebra system providing an environment where the user could both program efficiently and elaborate formal and machine-checked proofs of correctness, by calling a general-purpose proof assistant like the Coq system. This approach requires a huge manpower and a daunting effort in order to re-implement a complete computer-algebra system, as well as the libraries of formal mathematics required by such formal proofs.

3.2.4. Semantics for computer algebra

The move to machine-checked proofs of the mathematical correctness of the output of computer-algebra implementations demands a prior clarification about the often implicit assumptions on which the presumably correctly implemented algorithms rely. Interestingly, this preliminary work, which could be considered as independent from a formal certification project, is seldom precise or even available in the literature.

3.2.5. Formal proofs for symbolic components of computer-algebra systems

A number of authors have investigated ways to organize the communication of a chosen computer-algebra system with a chosen proof assistant in order to certify specific components of the computer-algebra systems, experimenting various combinations of systems and various formats for mathematical exchanges. Another line of research consists in the implementation and certification of computer-algebra algorithms inside the logic [70], [49], [59] or as a proof-automation strategy. Normalization algorithms are of special interest when they allow to check results possibly obtained by an external computer-algebra oracle [42]. A discussion about the systematic separation of the search for a solution and the checking of the solution is already clearly outlined in [55].

3.2.6. Formal proofs for numerical components of computer-algebra systems

Significant progress has been made in the certification of numerical applications by formal proofs. Libraries formalizing and implementing floating-point arithmetic as well as large numbers and arbitrary-precision arithmetic are available. These libraries are used to certify floating-point programs, implementations of mathematical functions and for applications like hybrid systems.

---

[0]for instance http://dlmf.nist.gov/ for special functions or http://oeis.org/ for integer sequences
[0]http://functions.wolfram.com/
[0]http://ddmf.msr-inria.inria.fr/1.9.1/ddmf
[0]http://129.81.170.14/~vhm/Table.html

3.3. Machine-checked proofs of formalized mathematics

To be checked by a machine, a proof needs to be expressed in a constrained, relatively simple formal language. Proof assistants provide facilities to write proofs in such languages. But, as merely writing, even in a formal language, does not constitute a formal proof just per se, proof assistants also provide a proof checker: a small and well-understood piece of software in charge of verifying the correctness of arbitrarily large proofs. The gap between the low-level formal language a machine can check and the sophistication of an average page of mathematics is conspicuous and unavoidable. Proof assistants try to bridge this gap by offering facilities, like notations or automation, to support convenient formalization methodologies. Indeed, many aspects, from the logical foundation to the user interface, play an important role in the feasibility of formalized mathematics inside a proof assistant.

3.3.1. *Logical foundations and proof assistants*

While many logical foundations for mathematics have been proposed, studied, and implemented, type theory is the one that has been more successfully employed to formalize mathematics, to the notable exception of the Mizar system [60], which is based on set theory. In particular, the calculus of construction (CoC) [40] and its extension with inductive types (CIC) [41], have been studied for more than 20 years and been implemented by several independent tools (like Lego, Matita, and Agda). Its reference implementation, Coq [68], has been used for several large-scale formalizations projects (formal certification of a compiler back-end; four-color theorem). Improving the type theory underlying the Coq system remains an active area of research. Other systems based on different type theories do exist and, whilst being more oriented toward software verification, have been also used to verify results of mainstream mathematics (prime-number theorem; Kepler conjecture).

3.3.2. *Computations in formal proofs*

The most distinguishing feature of CoC is that computation is promoted to the status of rigorous logical argument. Moreover, in its extension CIC, we can recognize the key ingredients of a functional programming language like inductive types, pattern matching, and recursive functions. Indeed, one can program effectively inside tools based on CIC like Coq. This possibility has paved the way to many effective formalization techniques that were essential to the most impressive formalizations made in CIC.

Another milestone in the promotion of the computations-as-proofs feature of Coq has been the integration of compilation techniques in the system to speed up evaluation. Coq can now run realistic programs in the logic, and hence easily incorporates calculations into proofs that demand heavy computational steps.

Because of their different choice for the underlying logic, other proof assistants have to simulate computations outside the formal system, and indeed fewer attempts to formalize mathematical proofs involving heavy calculations have been made in these tools. The only notable exception, which was finished in 2014, the Kepler conjecture, required a significant work to optimize the rewriting engine that simulates evaluation in Isabelle/HOL.

3.3.3. *Large-scale computations for proofs inside the Coq system*

Programs run and proved correct inside the logic are especially useful for the conception of automated decision procedures. To this end, inductive types are used as an internal language for the description of mathematical objects by their syntax, thus enabling programs to reason and compute by case analysis and recursion on symbolic expressions.

The output of complex and optimized programs external to the proof assistant can also be stamped with a formal proof of correctness when their result is easier to *check* than to *find*. In that case one can benefit from their efficiency without compromising the level of confidence on their output at the price of writing and certify a checker inside the logic. This approach, which has been successfully used in various contexts, is very relevant to the present research project.

3.3.4. Relevant contributions from the Mathematical Component libraries

Representing abstract algebra in a proof assistant has been studied for long. The libraries developed by the MathComp project for the proof of the Odd Order Theorem provide a rather comprehensive hierarchy of structures; however, they originally feature a large number of instances of structures that they need to organize. On the methodological side, this hierarchy is an incarnation of an original work [43] based on various mechanisms, primarily type inference, typically employed in the area of programming languages. A large amount of information that is implicit in handwritten proofs, and that must become explicit at formalization time, can be systematically recovered following this methodology.

Small-scale reflection [46] is another methodology promoted by the MathComp project. Its ultimate goal is to ease formal proofs by systematically dealing with as many bureaucratic steps as possible, by automated computation. For instance, as opposed to the style advocated by Coq's standard library, decidable predicates are systematically represented using computable boolean functions: comparison on integers is expressed as program, and to state that $a \leq b$ one compares the output of this program run on a and b with $true$. In many cases, for example when a and b are values, one can prove or disprove the inequality by pure computation.

The MathComp library was consistently designed after uniform principles of software engineering. These principles range from simple ones, like naming conventions, to more advanced ones, like generic programming, resulting in a robust and reusable collection of formal mathematical components. This large body of formalized mathematics covers a broad panel of algebraic theories, including of course advanced topics of finite group theory, but also linear algebra, commutative algebra, Galois theory, and representation theory. We refer the interested reader to the online documentation of these libraries [69], which represent about 150,000 lines of code and include roughly 4,000 definitions and 13,000 theorems.

Topics not addressed by these libraries and that might be relevant to the present project include real analysis and differential equations. The most advanced work of formalization on these domains is available in the HOL-Light system [51], [52], [53], although some existing developments of interest [28], [61] are also available for Coq. Another aspect of the MathComp libraries that needs improvement, owing to the size of the data we manipulate, is the connection with efficient data structures and implementations, which only starts to be explored.

3.3.5. User interaction with the proof assistant

The user of a proof assistant describes the proof he wants to formalize in the system using a textual language. Depending on the peculiarities of the formal system and the applicative domain, different proof languages have been developed. Some proof assistants promote the use of a declarative language, when the Coq and Matita systems are more oriented toward a procedural style.

The development of the large, consistent body of MathComp libraries has prompted the need to design an alternative and coherent language extension for the Coq proof assistant [48], [47], enforcing the robustness of proof scripts to the numerous changes induced by code refactoring and enhancing the support for the methodology of small-scale reflection.

The development of large libraries is quite a novelty for the Coq system. In particular any long-term development process requires the iteration of many refactoring steps and very little support is provided by most proof assistants, with the notable exception of Mizar [65]. For the Coq system, this is an active area of research.

4. Highlights of the Year

4.1. Highlights of the Year

4.1.1. Awards

Pierre Lairez has received the *ISSAC Distinguished Paper Award* for his joint work with T. Vaccon on p-adic differential equations [58].

5. New Software and Platforms

5.1. DDMF

Dynamic Dictionary of Mathematical Functions
FUNCTIONAL DESCRIPTION

Web site consisting of interactive tables of mathematical formulas on elementary and special functions. The formulas are automatically generated by OCaml and computer-algebra routines. Users can ask for more terms of the expansions, more digits of the numerical values, proofs of some of the formulas, etc.

This year, Maxence Guesdon started to port DDMF to the new DynaMoW. To this end, a special environment has been set up to be able to use the Inria continuous-integration platform.

- Participants: Alexandre Benoit, Frédéric Chyzak, Alexis Darrasse, Stefan Gerhold, Thomas Grégoire, Maxence Guesdon, Christoph Koutschan, Marc Mezzarobba and Bruno Salvy
- Contact: Frédéric Chyzak
- URL: http://ddmf.msr-inria.inria.fr/1.9.1/ddmf

5.2. DynaMoW

Dynamic Mathematics on the Web
FUNCTIONAL DESCRIPTION

DynaMoW is a programming tool for controlling the generation of mathematical websites that embed dynamical mathematical contents generated by computer-algebra calculations. Implemented in OCaml.

After a complete redesign and rewrite last year, to get more reactiveness and configurability, the implementation of DynaMoW was made more robust this year while porting ECS to this new library. It was next further enhanced, in particular in order have informative and reliable traces of execution, to help with the debugging of asynchronous parallel executions of services.

- Participants: Frédéric Chyzak, Alexis Darrasse and Maxence Guesdon
- Contact: Frédéric Chyzak
- URL: http://ddmf.msr-inria.inria.fr/DynaMoW/

5.3. ECS

Encyclopedia of Combinatorial Structures
FUNCTIONAL DESCRIPTION

ECS is an online mathematical encyclopedia with an emphasis on sequences that arise in the context of decomposable combinatorial structures, with the possibility to search by the first terms in the sequence, keyword, generating function, or closed form.

This year, we finalized the port of ECS to the last evolutions of DynaMoW. A new website was setup, and ECS is now again online, after a few years of discontinuation for technical reasons.

- Participants: Stéphanie Petit, Alexis Darrasse, Frédéric Chyzak and Maxence Guesdon
- Contact: Frédéric Chyzak
- URL: http://ecs.inria.fr/

5.4. Math-Components

Mathematical Components library
FUNCTIONAL DESCRIPTION

The Mathematical Components library is a set of Coq libraries that cover the mechanization of the proof of the Odd Order Theorem.

This year we experimented the maintenance of the library using the public repository stored on the github platform since December 2015. This allowed for merging several contributions from external users and improved significantly the communication with the community of users. A new website has also been set up, which includes pointers to various teaching and documentation resources.

- Participants: Andrea Asperti, Jeremy Avigad, Yves Bertot, Cyril Cohen, Francois Garillot, Georges Gonthier, Stéphane Le Roux, Assia Mahboubi, Sidi Ould Biha, Ioana Pasca, Laurence Rideau, Alexey Solovyev, Enrico Tassi and Russell O'connor
- Contact: Assia Mahboubi
- URL: http://www.msr-inria.fr/projects/mathematical-components-2/

5.5. Ssreflect

FUNCTIONAL DESCRIPTION

Ssreflect is a tactic language extension to the Coq system, developed by the Mathematical Components team. This year we improved the manual of the language, in order to document new features and to clarify some older parts of the document.

- Participants: Cyril Cohen, Yves Bertot, Laurence Rideau, Enrico Tassi, Laurent Thery, Assia Mahboubi and Georges Gonthier
- Contact: Yves Bertot
- URL: http://ssr.msr-inria.inria.fr/

6. New Results

6.1. Formally certified computation of definite integrals

Assia Mahboubi and Thomas Sibut-Pinote, in collaboration with Guillaume Melquiond (Toccata), have developed a Coq library for the computation of intervals approximating the value of definite integrals for elementary mathematical functions. This library provides an automated tool which builds automatically a formal proof of the correctness of the output, that is: a formal proof that the interval contains the mathematical values and a formal proof of the integrability of the input function on the input interval. A description of this work was published in the proceeding of the ITP 2016 conference [13]. An extension to domains including singularities of the integrand is in progress.

6.2. Real closed fields

Assia Mahboubi has worked with Henri Lombardi (Université de Franche Comté) on a constructive axiomatization of real closed fields. For this purpose, they have proposed an equational theory based on virtual roots and close to the classical notion of local real closed rings. This is a first step toward a constructive understanding of o-minimal structures. This work has been accepted for publication in Contemporary Mathematics [19].

6.3. Combinatorial walks with small steps in the quarter plane

Alin Bostan and Frédéric Chyzak, together with Mark van Hoeij (Florida State University), Manuel Kauers (Johannes Kepler University), and Lucien Pech (former intern), have applied their algorithms on special functions to generate complete, quantitative results in the enumerative theory of combinatorial walks with small steps in the quarter plane [2]. They gave the first proof that differential equations conjectured years ago by Bostan and Kauers are indeed satisfied by the corresponding generating functions. They also obtained explicit hypergeometric expressions for the latter, and could provably determine which of the generating functions are transcendental or algebraic.

6.4. Multiple binomial sums

Multiple binomial sums form a large class of multi-indexed sequences, closed under partial summation, which contains most of the sequences obtained by multiple summation of products of binomial coefficients and also all the sequences with algebraic generating function. Alin Bostan and Pierre Lairez, together with Bruno Salvy (Inria and ENS Lyon), have studied in [5] the representation of the generating functions of binomial sums by integrals of rational functions. The outcome is twofold. Firstly, we show that a univariate sequence is a multiple binomial sum if and only if its generating function is the diagonal of a rational function. Secondly, we propose algorithms that decide the equality of multiple binomial sums and that compute recurrence relations for them. In conjunction with geometric simplifications of the integral representations, this approach behaves well in practice. The process avoids the computation of certificates and the problem of the appearance of spurious singularities that afflicts discrete creative telescoping, both in theory and in practice.

6.5. Algebraic diagonals and walks

The diagonal of a multivariate power series F is the univariate power series $\mathrm{Diag}F$ generated by the diagonal terms of F. Diagonals form an important class of power series; they occur frequently in number theory, theoretical physics and enumerative combinatorics. In [35], Alin Bostan and Louis Dumont, together with Bruno Salvy (Inria and ENS Lyon), have studied algorithmic questions related to diagonals in the case where F is the Taylor expansion of a bivariate rational function. It is classical that in this case $\mathrm{Diag}F$ is an algebraic function. We propose an algorithm that computes an annihilating polynomial for $\mathrm{Diag}F$. We give a precise bound on the size of this polynomial and show that generically, this polynomial is the minimal polynomial and that its size reaches the bound. The algorithm runs in time quasi-linear in this bound, which grows exponentially with the degree of the input rational function. We then address the related problem of enumerating directed lattice walks. The insight given by our study leads to a new method for expanding the generating power series of bridges, excursions and meanders. We show that their first N terms can be computed in quasi-linear complexity in N, without first computing a very large polynomial equation. An extended version of this work is presented in [3].

6.6. A human proof of the Gessel conjecture

Counting lattice paths obeying various geometric constraints is a classical topic in combinatorics and probability theory. Many recent works deal with the enumeration of 2-dimensional walks with prescribed steps confined to the positive quadrant. A notoriously difficult case concerns the so-called *Gessel walks*: they 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 such 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. Alin Bostan, together with Irina Kurkova (Univ. Paris 6) and Kilian Raschel (CNRS and Univ. Tours), have proposed in [4] the first "human proofs" of these results. They are derived from a new expression for the generating function of Gessel walks in terms of special functions.

6.7. Enumeration of 3-dimensional lattice walks confined to the positive octant

Small step walks in 2D are by now quite well understood, but almost everything remains to be done in higher dimensions. Alin Bostan, together with Mireille Bousquet-Mélou (CNRS and Univ. Bordeaux), Manuel Kauers (Johannes Kepler Univ.) and Stephen Melczer (Univ. of Waterloo and ENS Lyon), have explored in [1] the classification problem for 3-dimensional walks with unit steps confined to the positive octant. The first difficulty is their number: there are 11 074 225 cases (instead of 79 in dimension 2). In our work, we focused on the 35 548 that have at most six steps. We applied to them a combined approach, first experimental and then rigorous. Among the 35 548 cases, we first found 170 cases with a finite group; in the remaining cases, our experiments suggest that the group is infinite. We then rigorously proved D-finiteness of the generating

series in all the 170 cases, with the exception of 19 intriguing step sets for which the nature of the generating function still remains unclear. In two challenging cases, no human proof is currently known, and we derived computer-algebra proofs, thus constituting the first proofs for those two step sets.

6.8. Computation of the similarity class of the p-curvature

The p-curvature of a system of linear differential equations in positive characteristic p is a matrix that measures how far the system is from having a basis of polynomial solutions. Alin Bostan, together with Xavier Caruso (CNRS and Univ. Rennes) and Éric Schost (Univ. Waterloo), have showed in [10] that the similarity class of the p-curvature can be determined without computing the p-curvature itself. More precisely, we have designed an algorithm that computes the invariant factors of the p-curvature in time quasi-linear in $\sqrt{p}$. This is much less than the size of the p-curvature, which is generally linear in p. The new algorithm allowed to answer a question originating from the study of the Ising model in statistical physics.

6.9. Efficient algorithms for mixed creative telescoping

Creative telescoping is a powerful computer algebra paradigm –initiated by Doron Zeilberger in the 90's– for dealing with definite integrals and sums with parameters. Alin Bostan and Louis Dumont, together with Bruno Salvy (Inria and ENS Lyon), have addressed in [12] the mixed continuous–discrete case, and have focussed on the integration of bivariate hypergeometric-hyperexponential terms. We have designed a new creative telescoping algorithm operating on this class of inputs, based on a Hermite-like reduction procedure. The new algorithm has two nice features: it is efficient and it delivers, for a suitable representation of the input, a minimal-order telescoper. Its analysis reveals tight bounds on the sizes of the telescoper it produces.

6.10. Fast computation of the Nth term of an algebraic series over a finite prime field

Alin Bostan and Philippe Dumas, together with Gilles Christol (IMJ), have addressed in [11] the question of computing one selected term of an algebraic power series. In characteristic zero, the best algorithm currently known for computing the Nth coefficient of an algebraic series uses differential equations and has arithmetic complexity quasi-linear in $\sqrt{N}$. We show that over a prime field of positive characteristic p, the complexity can be lowered to $O(\log N)$. The mathematical basis for this dramatic improvement is a classical theorem stating that a formal power series with coefficients in a finite field is algebraic if and only if the sequence of its coefficients can be generated by an automaton. We revisit and enhance two constructive proofs of this result for finite prime fields. The first proof uses Mahler equations, whose sizes appear to be prohibitively large. The second proof relies on diagonals of rational functions; we turn it into an efficient algorithm, of complexity linear in $\log N$ and quasi-linear in p.

6.11. Formal methods for cryptocurrencies

Georges Gonthier and Thomas Sibut-Pinote, along with a team of researchers from Microsoft Research and Inria, participated in a hackathon internal to Microsoft Research with the goal to apply formal methods to the verification of the smart contracts involved in the Ethereum platform. They outlined a framework to analyze and verify both the runtime safety and the functional correctness of Ethereum contracts by translation to F*, a functional programming language aimed at program verification. This work was published in the proceedings of the PLAS 2016 conference [9].

6.12. Computing solutions of linear Mahler equations

Mahler equations relate evaluations of the same function f at iterated bth powers of the variable. They arise in particular in the study of automatic sequences and in the complexity analysis of divide-and-conquer algorithms. Recently, the problem of solving Mahler equations in closed form has occurred in connection with number-theoretic questions. A difficulty in the manipulation of Mahler equations is the exponential blow-up

of degrees when applying a Mahler operator to a polynomial. In [17], Frédéric Chyzak and Philippe Dumas, together with Thomas Dreyfus (Université Claude Bernard Lyon 1) and Marc Mezzarobba (visiting scientist from UPMC), have presented algorithms for solving linear Mahler equations for series, polynomials, and rational functions, and have obtained polynomial-time complexity under a mild assumption.

6.13. Formal solutions of singularly perturbed linear differential systems

Suzy Maddah, together with Boulay Barkatou (Université de Limoges), has obtained algorithms for computing formal invariants of singularly-perturbed linear differential systems [20].

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

- *Mathematical Components* (project of the MSR–INRIA Joint Centre).
  Goal: Investigate the design of large-scale, modular and reusable libraries of formalized mathematics, using the Coq proof assistant. This project successfully formalized the proof of the Odd Order Theorem, resulting in a corpus of libraries related to various areas of algebra.
  Leader: Georges Gonthier (MSR Cambridge). Participants: Georges Gonthier, Assia Mahboubi.
  Website: http://www.msr-inria.fr/projects/mathematical-components/.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

FastRelax (ANR-14-CE25-0018).
Goal: Develop computer-aided proofs of numerical values, with certified and reasonably tight error bounds, without sacrificing efficiency.
Leader: B. Salvy (Inria, ÉNS Lyon). Participants: Assia Mahboubi, Th. Sibut-Pinote.
Website: http://fastrelax.gforge.inria.fr/.

8.2. European Initiatives

8.2.1. Collaborations in European Programs, Except FP7 & H2020

- Program: COST
- Project acronym: EUTYPES (CA15123)
- Project title: The European research network on types for programming and verification
- Duration: October 2015 - October 2019
- Coordinator: Herman Geuvers (Radboud University, Nijmegen, the Netherlands)
- Other partners: Czech Republic, Estonia, Macedonia, Germany, Greece, the Netherlands, Norway, Poland, Serbia, Slovenia, United Kingdom.

- Abstract: Types are pervasive in programming and information technology. A type defines a formal interface between software components, allowing the automatic verification of their connections, and greatly enhancing the robustness and reliability of computations and communications. In rich dependent type theories, the full functional specification of a program can be expressed as a type. Type systems have rapidly evolved over the past years, becoming more sophisticated, capturing new aspects of the behaviour of programs and the dynamics of their execution. This COST Action will give a strong impetus to research on type theory and its many applications in computer science, by promoting: (1) the synergy between theoretical computer scientists, logicians and mathematicians to develop new foundations for type theory, for example as based on the recent development of "homotopy type theory", (2) the joint development of type theoretic tools as proof assistants and integrated programming environments, (3) the study of dependent types for programming and its deployment in software development, (4) the study of dependent types for verification and its deployment in software analysis and verification. The action will also tie together these different areas and promote cross-fertilisation. Europe has a strong type theory community, ranging from foundational research to applications in programming languages, verification and theorem proving, which is in urgent need of better networking. A COST Action that crosses the borders will support the collaboration between groups and complementary expertise, and mobilise a critical mass of existing type theory research.

8.3. International Research Visitors

8.3.1. Research Stays Abroad

- Thomas Sibut-Pinote has spent two months at Microsoft Research Cambridge, visiting Georges Gonthier and working on mathematical libraries for the Lean proof assistant. He also participated in a hackathon internal to Microsoft Research with the goal to apply formal methods to the verification of the smart contracts involved in the Ethereum framework for cryptocurrency.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

- Frédéric Chyzak is member of the steering committee of the *Journées Nationales de Calcul Formel* (JNCF), the annual meeting of the French computer algebra community.
- Frédéric Chyzak has been elected member of the steering committee of the *International Symposium on Symbolic and Algebraic Computation* (ISSAC, 3-year term).
- Alin Bostan is part of the Scientific advisory board of the conference series *Effective Methods in Algebraic Geometry* (MEGA).
- Assia Mahboubi serves in the scientific advisory board of the Mathematics, Algorithms and Proofs community (MAP).
- Assia Mahboubi has served in the scientific advisory board of the École du GDR Informatique – Mathématiques 2016.
- Georges Gonthier is a member of the steering committee of the *Certified Programs and Proofs* conference (CPP).

9.1.1.2. Member of the Organizing Committees

- Assia Mahboubi has co-organized the *MAP'16 conference* at CIRM (Marseille), with B. Spitters (Aarhus University, Denmark) and P. Schuster (University of Verona, Italy): http://scientific-events.weebly.com/1508.html.

- Assia Mahboubi has co-organized, with E. Tassi (Marelle), the workshop *Mathematical Components: an introduction*, satellite of the conference ITP 2016: https://itp2016.inria.fr/workshops/#mc.

- Assia Mahboubi has co-organized, with K. Nakata (FireEye, Germany), the workshop *TTT*, satellite of the POPL'17 conference: http://popl17.sigplan.org/track/TTT-2017.

- Suzy Maddah has co-organized the gathering *Functional Equations in Limoges (FELIM 2016)*: https://indico.math.cnrs.fr/event/919/.

- Suzy Maddah has co-organized a session on software for the symbolic study of functional equations at the *International Congress on Mathematical Software (ICMS 2016)*: http://icms2016.zib.de/.

- Alin Bostan has co-organized, together with Bruno Salvy (Inria and ENS Lyon) and Conrado Martinez (UPC BarcelonaTech), the conference *ALEA 2016* at CIRM (Marseille): http://scientific-events.weebly.com/1406.html.

9.1.2. Scientific Events Selection

9.1.2.1. Chair of Conference Program Committees

- Alin Bostan has served as Symbolic Computation track chair for the international conference SYNASC 2016.

9.1.2.2. Member of the Conference Program Committees

- Assia Mahboubi has served as a member of the program committee for the international conferences with proceedings CPP 2017, ITP 2016, CSL 2016, CICM 2016 and SCSS 2016. She has also served as member of the program committee for the MAP 2016 conference and for the HaTT workshop.

- Alin Bostan has served as a member of the program committee of the ISSAC 2016 and of the SYNASC 2016 international conferences.

9.1.2.3. Reviewer

- Assia Mahboubi has served as reviewer for the proceedings of the international conferences CPP 2017, ITP 2016, CSL 2016, CICM 2016 and SCSS.

- Alin Bostan has served as reviewer for the proceedings of the international conferences FPSAC 2016, ISSAC 2016, AofA 2016 and SYNASC 2016.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Georges Gonthier is a member of the editorial board of the *Journal of Formalized Reasoning*.

9.1.3.2. Reviewer - Reviewing Activities

- Assia Mahboubi has served as a reviewer for the *Journal of Automated Reasoning*.

- Frédéric Chyzak has served as a reviewer for the journals: *Applicable Algebra in Engineering, Communication and Computing*; *Journal of Symbolic Computation*; *Journal of Algebra*; and *Electronic Journal of Combinatorics*.

- Alin Bostan has served as a reviewer for the journals: *Journal of Complexity*; *Mathematics of Computation*; *Linear Algebra and its Applications*; *Journal of Physics A: Mathematical and Theoretical*; *Journal of Algebra and its Applications*; *Journal of Symbolic Computation*; *Advances in Applied Mathematics*.

9.1.4. Invited Talks

- Assia Mahboubi has given an invited talk at the special trimester Mathematics – Computer Science – Philosophy CIPPMI in Toulouse in March 2016.

- Assia Mahboubi has given an invited lecture for the students of École Normale Supérieure Paris-Saclay in September 2016.

- Alin Bostan has been invited to give a series of five lectures at the summer school *Algorithmic and Enumerative Combinatorics* (RISC, Hagenberg, Austria), August 1–5, 2016.
- Philippe Dumas has given an invited lecture about divide-and-conquer recurrences at the *Journées Aléa* (CIRM, Marseille, France), March 7–11, 2016 [18].

9.1.5. Leadership within the Scientific Community

- Assia Mahboubi is leading the working group *Type theory based tools* inside the EUTYPES COST project. She is also M.C. for France for this project and a member of its core managment group.

9.1.6. Research Administration

- Assia Mahboubi is member of the *Commission Scientifique* of Inria Saclay — Île-de-France.
- Georges Gonthier is a member of the board of the *École doctorale de mathématiques Hadamard (EDMH)*.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

License:

    Louis Dumont, two L1 maths courses, 64h, Université Paris-Sud, France.

Master:

    Assia Mahboubi, *Proof Assistants*, 18h, M2, Denis Diderot University (Paris), France.

    Frédéric Chyzak, *Algorithmes efficaces en calcul formel*, 18h, M2, MPRI, France.

    Alin Bostan, *Algorithmes efficaces en calcul formel*, 40.5h, M2, MPRI, France.

    Pierre Lairez, *Algorithmique avancée*, 18h, M1, École polytechnique, France.

9.2.2. Supervision

PhD in progress: Thomas Sibut-Pinote, "Calcul numérique et démonstrations mathématiques: de la rigueur à la preuve formelle", September 2014, Assia Mahboubi

PhD in progress: Louis Dumont, "Algorithmes rapides pour le calcul symbolique de certaines intégrales de contour à paramètre", started in September 2013, supervised by Alin Bostan and B. Salvy.

Master intership (M1): G. Boisseau and Th. Huffschmitt, *Combination of decision procedures in presence of meta-variables*, École Polytechnique, supervised by Assia Mahboubi (jointly with S. Graham-Lengrand from LIX).

9.2.3. Juries

- Alin Bostan has served as a jury member of the French *Agrégation de Mathématiques – épreuve de modélisation, option C*.
- Alin Bostan has served as an examiner in the PhD jury of Aladin Virmaux, *Théorie des représentations combinatoires de tours de monoïdes, Application à la catégorification et aux fonctions de parking*, Université Paris-Saclay, June 13, 2016.

9.3. Popularization

- Assia Mahboubi has written a paper [8] for the quarterly journal of the Royal Dutch Mathematical Society.
- Alin Bostan has given a talk at the *Mathematic Park* seminar at IHP, Paris, on January 23rd 2016.

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] A. BOSTAN, M. BOUSQUET-MÉLOU, M. KAUERS, S. MELCZER.*On 3-dimensional lattice walks confined to the positive octant*, in "Annals of Combinatorics", October 2016, 36, First Online: 14 October 2016 [*DOI :* 10.1007/S00026-016-0328-7], https://hal.archives-ouvertes.fr/hal-01063886.

[2] A. BOSTAN, F. CHYZAK, M. VAN HOEIJ, M. KAUERS, L. PECH.*Hypergeometric Expressions for Generating Functions of Walks with Small Steps in the Quarter Plane*, in "European Journal of Combinatorics", 2016, 29 [*DOI :* 10.1016/J.EJC.2016.10.010], https://hal.inria.fr/hal-01332175.

[3] A. BOSTAN, L. DUMONT, B. SALVY.*Algebraic Diagonals and Walks: Algorithms, Bounds, Complexity*, in "Journal of Symbolic Computation", 2016 [*DOI :* 10.1016/J.JSC.2016.11.006], https://hal.archives-ouvertes.fr/hal-01244914.

[4] A. BOSTAN, I. KURKOVA, K. RASCHEL.*A human proof of Gessel's lattice path conjecture*, in "Transactions of the American Mathematical Society", 2016, vol. 369, n[o] 2, February 2017, p. 1365-1393, Published electronically: April 14, 2016, https://hal.archives-ouvertes.fr/hal-00858083.

[5] A. BOSTAN, P. LAIREZ, B. SALVY.*Multiple binomial sums*, in "Journal of Symbolic Computation", 2016 [*DOI :* 10.1016/J.JSC.2016.04.002], https://hal.archives-ouvertes.fr/hal-01220573.

[6] P. LAIREZ.*Computing periods of rational integrals*, in "Mathematics of Computation", November 2016, vol. 85, p. 1719-1752 [*DOI :* 10.1090/MCOM/3054], https://hal.inria.fr/hal-00981114.

[7] A. MAHBOUBI.*An Induction Principle over Real Numbers*, in "Archive for Mathematical Logic", 2016 [*DOI :* 10.1007/S00153-016-0513-8], https://hal.inria.fr/hal-01376054.

[8] A. MAHBOUBI.*Machine-checked mathematics*, in "Nieuw Archief voor Wiskunde", September 2016, vol. 5/17, n° 3, 5, https://hal.inria.fr/hal-01363284.

International Conferences with Proceedings

[9] K. BHARGAVAN, A. DELIGNAT-LAVAUD, C. FOURNET, A. GOLLAMUDI, G. GONTHIER, N. KOBEISSI, N. KULATOVA, A. RASTOGI, T. SIBUT-PINOTE, N. SWAMY, S. ZANELLA-BÉGUELIN.*Formal Verification of Smart Contracts: Short Paper*, in "ACM Workshop on Programming Languages and Analysis for Security", Vienna, Austria, October 2016 [*DOI :* 10.1145/2993600.2993611], https://hal.inria.fr/hal-01400469.

[10] A. BOSTAN, X. CARUSO, É. SCHOST.*Computation of the Similarity Class of the p-Curvature*, in "ISSAC 2016", Waterloo, ON, Canada, ACM Press, July 2016, p. 111–118 [*DOI :* 10.1145/2930889.2930897], https://hal.archives-ouvertes.fr/hal-01321043.

[11] A. BOSTAN, G. CHRISTOL, P. DUMAS.*Fast Computation of the Nth Term of an Algebraic Series over a Finite Prime Field*, in "ISSAC '16", Waterloo, Ontario, Canada, July 2016 [*DOI :* 10.1145/2930889.2930904], https://hal.inria.fr/hal-01389421.

[12] A. BOSTAN, L. DUMONT, B. SALVY.*Efficient Algorithms for Mixed Creative Telescoping*, in "IS-SAC 2016", Waterloo, Canada, Proceedings ISSAC'16, pp. 127–134, ACM Press, 2016., July 2016, 8 [*DOI :* 10.1145/2930889.2930907], https://hal.inria.fr/hal-01317940.

[13] A. MAHBOUBI, G. MELQUIOND, T. SIBUT-PINOTE.*Formally Verified Approximations of Definite Integrals*, in "Interactive Theorem Proving", Nancy, France, J. C. BLANCHETTE, S. MERZ (editors), Lecture Notes in Computer Science, August 2016, vol. 9807 [*DOI :* 10.1007/978-3-319-43144-4_17], https://hal.inria.fr/hal-01289616.

Scientific Books (or Scientific Book chapters)

[14] A. BOSTAN, F. CHYZAK, M. GIUSTI, R. LEBRETON, G. LECERF, B. SALVY, E. SCHOST.*Algorithmes Efficaces en Calcul Formel*, January 2017, Cette version est la prépublication de l'ouvrage du même nom à paraître dans la collection https://hal.archives-ouvertes.fr/AECF/, https://hal.inria.fr/hal-01431717.

Research Reports

[15] G. GONTHIER, A. MAHBOUBI, E. TASSI.*A Small Scale Reflection Extension for the Coq system*, Inria Saclay Ile de France, 2016, n^o RR-6455, https://hal.inria.fr/inria-00258384.

Other Publications

[16] A. BOSTAN.*Algebraicity and transcendence of power series: combinatorial and computational aspects*, August 2016, Lecture, https://hal.inria.fr/cel-01391907.

[17] F. CHYZAK, T. DREYFUS, P. DUMAS, M. MEZZAROBBA.*Computing solutions of linear Mahler equations*, 2016, working paper or preprint, https://hal.inria.fr/hal-01418653.

[18] P. DUMAS.*Divide-and-Conquer Algebra and Analysis : Lecture given at Journées ALÉA 2016*, March 2016, 80, Lecture, https://hal.inria.fr/cel-01388741.

[19] H. LOMBARDI, A. MAHBOUBI.*Théories géométriques pour l'algèbre des nombres réels*, January 2017, working paper or preprint, https://hal.inria.fr/hal-01426164.

[20] S. MADDAH, M. A. BARKATOU.*Formal Solutions of Singularly Perturbed Linear Differential Systems*, January 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01417265.

References in notes

[21] M. ABRAMOWITZ, I. A. STEGUN (editors). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, Dover, New York, 1992, xiv+1046, Reprint of the 1972 edition.

[22] *Computer Algebra Errors*, Article in mathematics blog MathOverflow, http://mathoverflow.net/questions/11517/computer-algebra-errors.

[23] F. W. J. OLVER, D. W. LOZIER, R. F. BOISVERT, C. W. CLARK (editors). *NIST Handbook of mathematical functions*, Cambridge University Press, 2010.

[24] M. ARMAND, B. GRÉGOIRE, A. SPIWACK, L. THÉRY.*Extending Coq with Imperative Features and its Application to SAT Verication*, in "Interactive Theorem Proving, international Conference, ITP 2010, Edinburgh, Scotland, July 11–14, 2010, Proceedings", Lecture Notes in Computer Science, Springer, 2010.

[25] B. BECKERMANN, G. LABAHN.*A uniform approach for the fast computation of matrix-type Padé approximants*, in "SIAM J. Matrix Anal. Appl.", 1994, vol. 15, n° 3, p. 804–823.

[26] A. BENOIT, F. CHYZAK, A. DARRASSE, S. GERHOLD, M. MEZZAROBBA, B. SALVY.*The Dynamic Dictionary of Mathematical Functions (DDMF)*, in "The Third International Congress on Mathematical Software (ICMS 2010)", K. FUKUDA, J. VAN DER HOEVEN, M. JOSWIG, N. TAKAYAMA (editors), Lecture Notes in Computer Science, 2010, vol. 6327, p. 35–41, http://dx.doi.org/10.1007/978-3-642-15582-6_7.

[27] M. BOESPFLUG, M. DÉNÈS, B. GRÉGOIRE.*Full reduction at full throttle*, in "First International Conference on Certified Programs and Proofs, Taiwan, December 7–9", Lecture Notes in Computer Science, Springer, 2011.

[28] S. BOLDO, C. LELAY, G. MELQUIOND.*Improving Real Analysis in Coq: A User-Friendly Approach to Integrals and Derivatives*, in "Certified Programs and Proofs", C. HAWBLITZEL, D. MILLER (editors), Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2012, vol. 7679, p. 289-304, http://dx.doi.org/10.1007/978-3-642-35308-6_22.

[29] S. BOLDO, G. MELQUIOND.*Flocq: A Unified Library for Proving Floating-point Algorithms in Coq*, in "Proceedings of the 20th IEEE Symposium on Computer Arithmetic", Tübingen, Germany, July 2011, p. 243–252.

[30] A. BOSTAN.*Algorithmes rapides pour les polynômes, séries formelles et matrices*, in "Actes des Journées Nationales de Calcul Formel", Luminy, France, 2010, p. 75–262, Les cours du CIRM, tome 1, numéro 2, http://ccirm.cedram.org:80/ccirm-bin/fitem?id=CCIRM_2010__1_2_75_0.

[31] A. BOSTAN, S. BOUKRAA, S. HASSANI, J.-M. MAILLARD, J.-A. WEIL, N. ZENINE.*Globally nilpotent differential operators and the square Ising model*, in "J. Phys. A: Math. Theor.", 2009, vol. 42, n° 12, 50, http://dx.doi.org/10.1088/1751-8113/42/12/125206.

[32] A. BOSTAN, S. CHEN, F. CHYZAK, Z. LI.*Complexity of creative telescoping for bivariate rational functions*, in "ISSAC'10: Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation", New York, NY, USA, ACM, 2010, p. 203–210, http://doi.acm.org/10.1145/1837934.1837975.

[33] A. BOSTAN, F. CHYZAK, G. LECERF, B. SALVY, É. SCHOST.*Differential equations for algebraic functions*, in "ISSAC'07: Proceedings of the 2007 international symposium on Symbolic and algebraic computation", C. W. BROWN (editor), ACM Press, 2007, p. 25–32, http://dx.doi.org/10.1145/1277548.1277553.

[34] A. BOSTAN, F. CHYZAK, M. VAN HOEIJ, L. PECH.*Explicit formula for the generating series of diagonal 3D rook paths*, in "Sém. Loth. Comb.", 2011, vol. B66a, 27, http://www.emis.de/journals/SLC/wpapers/s66bochhope.html.

[35] A. BOSTAN, L. DUMONT, B. SALVY.*Algebraic Diagonals and Walks*, in "ISSAC'15 International Symposium on Symbolic and Algebraic Computation", Bath, United Kingdom, ACM Press, July 2015, p. 77–84 [*DOI :* 10.1145/2755996.2756663], https://hal.archives-ouvertes.fr/hal-01240729.

[36] A. BOSTAN, M. KAUERS.*The complete generating function for Gessel walks is algebraic*, in "Proceedings of the American Mathematical Society", September 2010, vol. 138, n⁰ 9, p. 3063–3078, With an appendix by Mark van Hoeij.

[37] F. CHYZAK.*An extension of Zeilberger's fast algorithm to general holonomic functions*, in "Discrete Math.", 2000, vol. 217, n⁰ 1-3, p. 115–134, Formal power series and algebraic combinatorics (Vienna, 1997).

[38] F. CHYZAK, M. KAUERS, B. SALVY.*A Non-Holonomic Systems Approach to Special Function Identities*, in "ISSAC'09: Proceedings of the Twenty-Second International Symposium on Symbolic and Algebraic Computation", J. MAY (editor), 2009, p. 111–118, http://dx.doi.org/10.1145/1576702.1576720.

[39] F. CHYZAK, B. SALVY.*Non-commutative elimination in Ore algebras proves multivariate identities*, in "J. Symbolic Comput.", 1998, vol. 26, n⁰ 2, p. 187–227.

[40] T. COQUAND, G. P. HUET.*The Calculus of Constructions*, in "Inf. Comput.", 1988, vol. 76, n⁰ 2/3, p. 95-120, http://dx.doi.org/10.1016/0890-5401(88)90005-3.

[41] T. COQUAND, C. PAULIN-MOHRING.*Inductively defined types*, in "Proceedings of Colog'88", P. MARTIN-LÖF, G. MINTS (editors), Lecture Notes in Computer Science, Springer-Verlag, 1990, vol. 417.

[42] D. DELAHAYE, M. MAYERO.*Dealing with algebraic expressions over a field in Coq using Maple*, in "J. Symbolic Comput.", 2005, vol. 39, n⁰ 5, p. 569–592, Special issue on the integration of automated reasoning and computer algebra systems, http://dx.doi.org/10.1016/j.jsc.2004.12.004.

[43] F. GARILLOT, G. GONTHIER, A. MAHBOUBI, L. RIDEAU.*Packaging Mathematical Structures*, in "Theorem Proving in Higher-Order Logics", S. BERGHOFER, T. NIPKOW, C. URBAN, M. WENZEL (editors), Lecture Notes in Computer Science, Springer, 2009, vol. 5674, p. 327–342.

[44] J. VON ZUR. GATHEN, J. GERHARD.*Modern computer algebra*, 2nd, Cambridge University Press, New York, 2003, xiv+785.

[45] G. GONTHIER.*Formal proofs—the four-colour theorem*, in "Notices of the AMS", 2008, vol. 55, n⁰ 11, p. 1382-1393.

[46] G. GONTHIER, A. MAHBOUBI.*An introduction to small scale reflection in Coq*, in "Journal of Formalized Reasoning", 2010, vol. 3, n⁰ 2, p. 95–152.

[47] G. GONTHIER, A. MAHBOUBI, E. TASSI.*A Small Scale Reflection Extension for the Coq system*, Inria, 2008, n⁰ RR-6455, http://hal.inria.fr/inria-00258384.

[48] G. GONTHIER, E. TASSI.*A language of patterns for subterm selection*, in "ITP", LNCS, 2012, vol. 7406, p. 361–376.

[49] B. GRÉGOIRE, A. MAHBOUBI.*Proving Equalities in a Commutative Ring Done Right in Coq*, in "Theorem Proving in Higher Order Logics, 18th International Conference, TPHOLs 2005, Oxford, UK, August 22-25, 2005, Proceedings", Lecture Notes in Computer Science, Springer, 2005, vol. 3603, p. 98–113.

[50] T. HALES.*Formal proof*, in "Notices of the AMS", 2008, vol. 55, n⁰ 11, p. 1370-1380.

[51] J. HARRISON.*A HOL Theory of Euclidean space*, in "Theorem Proving in Higher Order Logics, 18th International Conference, TPHOLs 2005", Oxford, UK, J. HURD, T. MELHAM (editors), Lecture Notes in Computer Science, Springer-Verlag, 2005, vol. 3603.

[52] J. HARRISON.*Formalizing an analytic proof of the prime number theorem*, in "Journal of Automated Reasoning", 2009, vol. 43, p. 243–261, Dedicated to Mike Gordon on the occasion of his 60th birthday.

[53] J. HARRISON.*Theorem proving with the real numbers*, CPHC/BCS distinguished dissertations, Springer, 1998.

[54] J. HARRISON.*A Machine-Checked Theory of Floating Point Arithmetic*, in "Theorem Proving in Higher Order Logics: 12th International Conference, TPHOLs'99", Nice, France, Y. BERTOT, G. DOWEK, A. HIRSCHOWITZ, C. PAULIN, L. THÉRY (editors), Lecture Notes in Computer Science, Springer-Verlag, 1999, vol. 1690, p. 113–130.

[55] J. HARRISON, L. THÉRY.*A Skeptic's Approach to Combining HOL and Maple*, in "J. Autom. Reason.", December 1998, vol. 21, n⁰ 3, p. 279–294, http://dx.doi.org/10.1023/A:1006023127567.

[56] F. JOHANSSON.*Another Mathematica bug*, Article on personal blog, http://fredrik-j.blogspot.fr/2009/07/another-mathematica-bug.html.

[57] C. KOUTSCHAN.*A fast approach to creative telescoping*, in "Math. Comput. Sci.", 2010, vol. 4, n⁰ 2-3, p. 259–266, http://dx.doi.org/10.1007/s11786-010-0055-0.

[58] P. LAIREZ, T. VACCON.*On p-adic differential equations with separation of variables*, in "ISSAC 2016", Waterloo, Canada, Proceedings ISSAC'16, pp. 127–134, ACM Press, 2016., July 2016, 8, http://dx.doi.org/10.1145/2930889.2930907.

[59] A. MAHBOUBI.*Implementing the cylindrical algebraic decomposition within the Coq system*, in "Mathematical Structures in Computer Science", 2007, vol. 17, n⁰ 1, p. 99–127.

[60] R. MATUSZEWSKI, P. RUDNICKI.*Mizar: the first 30 years*, in "Mechanized Mathematics and Its Applications", 2005, vol. 4.

[61] M. MAYERO.*Problèmes critiques et preuves formelles*, Université Paris 13, novembre 2012, Habilitation à Diriger des Recherches.

[62] M. MEZZAROBBA.*NumGfun: a package for numerical and analytic computation and D-finite functions*, in "ISSAC 2010—Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation", New York, ACM, 2010, p. 139–146, http://dx.doi.org/10.1145/1837934.1837965.

[63] P. PAULE, M. SCHORN.*A Mathematica version of Zeilberger's algorithm for proving binomial coefficient identities*, in "J. Symbolic Comput.", 1995, vol. 20, n⁰ 5-6, p. 673–698, Symbolic computation in combinatorics Δ_1 (Ithaca, NY, 1993), http://dx.doi.org/10.1006/jsco.1995.1071.

[64] B. PETERSEN.*Maple*, Personal web site.

[65] P. RUDNICKI, A. TRYBULEC.*On the Integrity of a Repository of Formalized Mathematics*, in "Proceedings of the Second International Conference on Mathematical Knowledge Management", London, UK, MKM '03, Springer-Verlag, 2003, p. 162–174, http://dl.acm.org/citation.cfm?id=648071.748518.

[66] B. SALVY, P. ZIMMERMANN.*Gfun: a Maple package for the manipulation of generating and holonomic functions in one variable*, in "ACM Trans. Math. Software", 1994, vol. 20, n° 2, p. 163–177.

[67] N. J. A. SLOANE, S. PLOUFFE.*The Encyclopedia of Integer Sequences*, Academic Press, San Diego, 1995.

[68] THE COQ DEVELOPMENT TEAM.*The Coq Proof Assistant: Reference Manual*, http://coq.inria.fr/doc/.

[69] THE MATHEMATICAL COMPONENT TEAM.*A Formalization of the Odd Order Theorem using the Coq proof assistant*, September 2012, http://www.msr-inria.fr/projects/mathematical-components/.

[70] L. THÉRY.*A Machine-Checked Implementation of Buchberger's Algorithm*, in "J. Autom. Reasoning", 2001, vol. 26, n° 2, p. 107-137, http://dx.doi.org/10.1023/A:1026518331905.

[71] K. WEGSCHAIDER.*Computer generated proofs of binomial multi-sum identities*, RISC, J. Kepler University, May 1997, 99.

[72] S. WOLFRAM.*Mathematica: A system for doing mathematics by computer (2nd ed.)*, Addison-Wesley, 1992, I.

[73] D. ZEILBERGER.*Opinion 94: The Human Obsession With "Formal Proofs" is a Waste of the Computer's Time, and, Even More Regretfully, of Humans' Time*, 2009, http://www.math.rutgers.edu/~zeilberg/Opinion94.html.

[74] D. ZEILBERGER.*A holonomic systems approach to special functions identities*, in "J. Comput. Appl. Math.", 1990, vol. 32, n° 3, p. 321–368.

[75] D. ZEILBERGER.*The method of creative telescoping*, in "J. Symbolic Comput.", 1991, vol. 11, n° 3, p. 195–204.

Project-Team TAO

Machine Learning and Optimisation

IN COLLABORATION WITH: Laboratoire de recherche en informatique (LRI)

IN PARTNERSHIP WITH:
CNRS

Université Paris-Sud (Paris 11)

RESEARCH CENTER
Saclay - Île-de-France

THEME
Optimization, machine learning and statistical methods

Table of contents

Project-Team TAO

Creation of the Project-Team: 2004 November 04, end of the Project-Team: 2016 December 31

Keywords:

Computer Science and Digital Science:

- 3.3.3. - Big data analysis
- 3.4. - Machine learning and statistics
- 3.5.2. - Recommendation systems
- 7.3. - Optimization
- 7.8. - Information theory
- 8.2. - Machine learning
- 8.3. - Signal analysis

Other Research Topics and Application Domains:

- 1.3. - Neuroscience and cognitive science
- 4. - Energy
- 7.1.2. - Road traffic
- 7.1.3. - Air traffic
- 9.1.2. - Serious games
- 9.4.3. - Physics
- 9.4.5. - Data science
- 9.5.10. - Digital humanities

1. Members

Research Scientists

Marc Schoenauer [Team leader, Inria, Senior Researcher, HDR]
Anne Auger [Inria, Researcher, until Dec. 2016, HDR]
Guillaume Charpiat [Inria, Researcher]
Cyril Furtlehner [Inria, Researcher]
Gregory Grefenstette [Inria, Advanced Research position, until Oct 2016]
Nikolaus Hansen [Inria, Senior Researcher, until Dec. 2016, HDR]
Odalric Maillard [Inria, Researcher, until Nov 2016]
Yann Ollivier [CNRS, Researcher, HDR]
Michèle Sebag [CNRS, Senior Researcher, HDR]
Olivier Teytaud [Inria, Researcher, until May 2016, HDR]
Paola Tubaro [CNRS, Researcher]

Faculty Members

Isabelle Guyon [Univ. Paris XI, Professor]
Philippe Caillou [Univ. Paris XI, Associate Professor]
Aurélien Decelle [Univ. Paris XI, Associate Professor]
Cécile Germain [Univ. Paris XI, Professor, HDR]

Technical Staff

Mohamed Bouatira [Inria]
Jean-Baptiste Hoock [Inria, until Sep 2016]
Felix Louistisserand [Inria, from Feb 2016]

Gaetan Marceau Caron [Inria, until Oct 2016, granted by Conseil Régional d'Ile-de-France]

Luis Marti Orosa [Inria]

Dejan Tusar [Inria, from Oct 2016]

PhD Students

Sandra Cecilia Astete Morales [Inria, until Nov 2016]

Asma Atamna [Inria, granted by Inria, until Dec. 2016]

Nacim Belkhir [Thalès, granted by CIFRE Thalès]

Vincent Berthier [Inria, granted by University Paris-Sud, until Dec.2016]

Chaouki Boufenar [Univ. Oran – Algerie, invited PhD Student, until Aug 2016]

Marie Liesse Cauwet [Inria, until Aug 2016, granted by Univ. Paris-Sud]

Benjamin Donnot [RTE, granted by CIFRE RTE]

Guillaume Doquet [Univ. Paris XI]

Victor Estrade [Univ. Paris XI]

François Gonard [Inst. de Recherche Technologique SystemX]

Sourava Prasad Mishra [Univ. Paris XI]

Adrian Pol [CERN, from Oct 2016]

Karima Rafes [Univ. Paris XI, from Oct 2014]

Thomas Schmitt [Inria]

Lisheng Sun [Univ. Paris XI, from Oct 2016]

Corentin Tallec [Ecole Polytechnique]

Pierre Wolinski [ENS Paris]

Ouassim Ait Elhara [Univ. Paris XI, until Dec. 2016]

Post-Doctoral Fellows

Berna Bakir Batu [Inria, from Nov 2016]

Alexandre Chotard [Inria, until Sep 2016]

Edgar Galvan Lopez [University College Dublin, Marie Curie Post-Doctoral Fellow]

Lawrence Githiari [Inria, until Jun 2016]

Olivier Goudet [Inria, from Jul 2016]

Duc Manh Nguyen [Inria, until Sep 2016]

Phillipe Rodrigues Sampaio [Inria, until Sep 2016]

Administrative Assistant

Olga Mwana Mobulakani [Inria]

Others

Flora Jay [CNRS, Researcher]

Nicolas Spyratos [Univ. Paris XI, Professor emeritus]

Véronique Ventos [Univ. Paris XI, Associate Professor]

Balász Kégl [CNRS, Senior Researcher]

2. Overall Objectives

2.1. Presentation

TAO main research domains, defined at the birth of the team in 2003, have slowly evolved since then. As TAO has now reached the 12-years lifetime limit of all Inria teams, it will die at the end of 2016 and most TAO researchers will together propose a new team with renewed research themes. This Section is thus meant as a historical remain of the late TAO team.

Data Mining (DM), acknowledged to be one of the main ten challenges of the 21st century [0], aims at building (partial) phenomenological models from the massive amounts of data produced in scientific labs, industrial plants, banks, hospitals or supermarkets. Machine Learning (ML) likewise aims at modeling the complex systems underlying the available data; the main difference between DM and ML disciplines is the emphasis put on the acquisition, storage and management of large-scale data.

DM and ML problems can be set as optimization problems, thus leading to two possible approaches. Note that this alternative has been characterized by H. Simon (1982) as follows. *In complex real-world situations, optimization becomes approximate optimization since the description of the real-world is radically simplified until reduced to a degree of complication that the decision maker can handle. Satisficing seeks simplification in a somewhat different direction, retaining more of the detail of the real-world situation, but settling for a satisfactory, rather than approximate-best, decision.*

The first approach is to simplify the learning problem to make it tractable by standard statistical or optimization methods. The alternative approach is to preserve as much as possible the genuine complexity of the goals (yielding "interesting" models, accounting for prior knowledge): more flexible optimization approaches are therefore required, such as those offered by Evolutionary Computation.

Symmetrically, optimization techniques are increasingly used in all scientific and technological fields, from optimum design to risk assessment. Evolutionary Computation (EC) techniques, mimicking the Darwinian paradigm of natural evolution, are stochastic population-based dynamical systems that are now widely known for their robustness and flexibility, handling complex search spaces (e.g., mixed, structured, constrained representations) and non-standard optimization goals (e.g., multi-modal, multi-objective, context-sensitive), beyond the reach of standard optimization methods.

The price to pay for such properties of robustness and flexibility is twofold. On one hand, EC is tuned, mostly by trials and errors, using quite a few parameters. On the other hand, EC generates massive amounts of intermediate solutions. It is suggested that the principled exploitation of preliminary runs and intermediate solutions, through Machine Learning and Data Mining techniques, can offer sound ways of adjusting the parameters and finding shortcuts in the trajectories in the search space of the dynamical system.

2.2. Context and overall goal of the project

The overall goals of the project are to model, predict, understand, and control physical or artificial systems. The central claim is that Learning and Optimization approaches must be used, adapted and integrated in a seamless framework, in order to bridge the gap between the system under study on the one hand, and the expert's goal as to the ideal state/functionality of the system on the other hand.

Specifically, our research context is based on the following assumptions:

1. The systems under study range from large-scale engineering systems to physical or chemical phenomenons, including robotics and games. Such systems, sometimes referred to as *complex systems*, can hardly be modeled based on first principles due to their size, their heterogeneity and the incomplete information aspects involved in their behavior.

2. Such systems can be observed; indeed selecting the relevant observations and providing a reasonably appropriate description thereof are part of the problem to be solved. A further assumption is that these observations are sufficient to build a reasonably accurate model of the system under study.

3. The available expertise is sufficient to assess the system state, and any modification thereof, with respect to the desired states/functionalities. The assessment function is usually not a well-behaved function (differentiable, convex, defined on a continuous domain, etc.), barring the use of standard optimization approaches and making Evolutionary Computation a better suited alternative.

---

[0]MIT Technological Review, feb. 2001.

In this context, the objectives of TAO are threefold:

1. Investigating how specific prior knowledge and requirements can be accommodated in Machine Learning thanks to evolutionary computation (EC) and more generally Stochastic Optimization;

2. Investigating how statistical Machine Learning can be used to interpret, study and enhance evolutionary computation;

3. Facing diversified and real-world applications, requiring and suggesting new integrated ML/EC approaches.

3. Research Program

3.1. The Five Pillars of TAO

This Section describes TAO main research directions at the crossroad of Machine Learning and Evolutionary Computation. Since 2008, TAO has been structured in several special interest groups (SIGs) to enable the agile investigation of long-term or emerging theoretical and applicative issues. The comparatively small size of TAO SIGs enables in-depth and lively discussions; the fact that all TAO members belong to several SIGs, on the basis of their personal interests, enforces the strong and informal collaboration of the groups, and the fast information dissemination.

The first two SIGs consolidate the key TAO scientific pillars, while the others evolve and adapt to new topics.

The **Stochastic Continuous Optimization** SIG (OPT-SIG) takes advantage of the fact that TAO is acknowledged the best French research group and one of the top international groups in evolutionary computation from a theoretical and algorithmic standpoint. A main priority on the OPT-SIG research agenda is to provide theoretical and algorithmic guarantees for the current world state-of-the-art continuous stochastic optimizer, CMA-ES, ranging from convergence analysis to a rigorous benchmarking methodology. Incidentally, the benchmark platform COCO has been acknowledged since 2009 as "the" international continuous optimization benchmark, and its extension is at the core of the ANR projects NumBBO and NumBBO2. Another priority is to address the current limitations of CMA-ES in terms of high-dimensional or expensive optimization and constraint handling (respectively Ouassim Ait El Hara's and Asma Atamna's PhDs). Note that most members of this SIG have moved to the recently created Inria team RANDOPT by December 2016.

The **Optimal Decision Making under Uncertainty** SIG (UCT-SIG) benefits from the MoGo expertise and its past and present world records in the domain of computer-Go, establishing the international visibility of TAO in sequential decision making. Since 2010, UCT-SIG resolutely moves to address the problems of **energy management** from a fundamental and applied perspective. On the one hand, energy management offers a host of challenging issues, ranging from long-horizon policy optimization to the combinatorial nature of the search space, from the modeling of prior knowledge to non-stationary environment to name a few. On the other hand, the energy management issue can hardly be tackled in a pure academic perspective: tight collaborations with industrial partners are needed to access the true operational constraints. Such international and national collaborations have been started by Olivier Teytaud during his three stays (1 year, 6 months, 6 months) in Taiwan, and witnessed by the FP7 STREP Citines, the ADEME Post contract, and the METIS I-lab with SME Artelys. Note that Olivier Teytaud has left TAO for Google-Zurich on June 6., 2016. The project is continuing in collaboration with RTE under the leadership of Isabelle Guyon and Marc Schoenauer, making connections with Data Science.

The **Data Science** SIG (DS-SIG) includes the activities conducted or started within the CDS and ISN Lidexes in Saclay. On the one hand, it replaces and extends the former *Distributed systems* SIG, that was devoted to the modeling and optimization of (large scale) distributed systems, and itself was extending the goals of the original *Autonomic Computing* SIG, initiated by Cécile Germain-Renaud and investigating the use of statistical Machine Learning for large scale computational architectures (from data acquisition − the Grid Observatory in the European Grid Initiative − to grid management and fault detection). Under the application pressure from natural and social sciences (ranging from High Energy Physics to computational social sciences), this

SIG has evolved. A major result of this theme has been the creation 3 years ago of the Paris-Saclay Center for Data Science, co-chaired by Balázs Kégl, and the organization of the Higgs-ML challenge (http://higgsml.lal. in2p3.fr/), most popular challenge ever on the Kaggle platform. Another large scale data challenge sponsored by Microsoft with USD 60000 in prizes on the theme of Automatic Machine Learning (AutoML) in 2015/2016 was crowned by success: the winners developed a new tool called AutoSKlearn as a wrapper to the scikit-learn library, an open source project lead by Inria team Parietal.

On the other hand, several activities around Computational Social Sciences involving Gregory Grefenstette, Cécile Germain-Renaud, Michèle Sebag, Philippe Caillou, Isabelle Guyon and Paola Tubaro, have widely extended previous work around the modeling of multi-agent systems and the exploitation of simulation results in the SimTools RNSC network frame. A research direction involves adding semantics to underspecified collections of societal information: in an historical perspective (as in the new TAO H2020 project, EHRI-II on holocaust archives, or in the Gregorius project on church history) or an individual perspective (as in the ongoing Personal Semantics project). Another research direction, developed within the Paris-Saclay Institute for Digital Society (ISN Lidex), examines societal questions (frictional unemployment, Th. Schmitt's PhD, or quality of life at work, O. Goudet's post-doc, or scientific institution activities, F. Louistisserand's engineer stint on Cartolabe) in a data-driven perspective. The key challenge here is to use learning algorithms to find structure and extract knowledge from poorly structured or unstructured information, and to provide intelligible results and/or means to interact with the user. Novel approaches involving causal modeling are under exploration.

The **Designing Criteria** SIG (CRI-SIG) focuses on the design of learning and optimization criteria. It elaborates on the lessons learned from the former *Complex Systems* SIG, showing that the key issue in challenging applications often is to design the objective itself. Such targeted criteria are pervasive in the study and building of autonomous cognitive systems, ranging from intrinsic rewards in robotics to the notion of saliency in vision and image understanding, and that of automatic algorithm selection and parameterization. The desired criteria can also result from fundamental requirements, such as scale invariance in a statistical physics perspective, and guide the algorithmic design. Additionally, the criteria can also be domain-driven and reflect the expert priors concerning the structure of the sought solution (e.g., spatio-temporal consistency); the challenge is to formulate such criteria in a mixed non convex/non differentiable objective function, nevertheless amenable to tractable optimization.

The **Deep Learning and Information Theory** SIG (DEEP-SIG) originated from some extensions of the work done in the *Distributed Systems* SIG that have been developed in the context of the TIMCO FUI project (started end 2012 and just ended); the challenge was not only to port ML algorithms on massively distributed architectures, but to see how these architectures can inspire new ML criteria and methodologies. The coincidence of this project with the arrival of Yann Ollivier in TAO gradualy led this work toward Deep Networks. Other research themes of this SIG are concerned with studying various theoretical and practical aspects of deep learning, providing information-theoretic perspectives on the design and optimization of deep learning models, such as using the Fisher information matrix to optimize the parameters, or using minimum description length criteria to choose the right model structure (topology of the neural graph, addition or removal of parameters...) and to provide regularization and model selection. This activity has also branched out into exploring various applications of Deep Learning. Isabelle Guyon has been involved in applications in computer vision, including the study of personality traits in video data and the verification of fingerprints. Energy Management (Section 4.1), Computational Social Sciences (Section 4.2), and anomaly detection are now also steered toward using Deep Networks for different variants of representation learning.

4. Application Domains

4.1. Energy Management

Energy management, our prioritary application field, involves sequential decision making with:

- stochastic uncertainties (typically weather);

- both high scale combinatorial problems (as induced by nuclear power plants) and non-linear effects;

- high dimension (including hundreds of hydroelectric stocks);

- multiple time scales:

    – minutes (dispatching, ensuring the stability of the grid), essentially beyond the scope of our work, but introducing constraints for our time scales;

    – days (unit commitment, taking care of compromises between various power plants);

    – years, for evaluating marginal costs of long term stocks (typically hydroelectric stocks);

    – decades, for investments.

Significant challenges also include:

- spatial distribution of problems; due to capacity limits we can not consider a power grid like Europe + North Africa as a single "production = demand" constraint; with extra connections we can equilibrate excess production by renewables for remote areas, but not in an unlimited manner.

- other uncertainties, which might be modelized by adversarial or stochastic frameworks (e.g. technological breakthroughs, decisions about ecological penalization).

We have had several related projects in the past, many of them together with the SME Artelys, working on optimization in general, and in particular on energy management. In particular, we had with them an Inria ILAB (Metis, ended in end 2014), and are currently working on POST, an ADEME BIA project about investments in power systems that will end in July 2017. Another project has been submitted to ADEME about the optimization of the local grids (at the city level) depending on the demand and the prediction of the market prices.

In 2016, we started to work with RTE, the company that is managing the global electric network in France. They fund Benjamin Donnot's CIFRE PhD thesis about learning the parries to prevent the loss off security of the network in case of material failures or unexpected consumption peaks. This collaboration had several follow-up, including the organization of a large scale challenge funded by the EU http://see4c.eu/, which will be endowed with 2 million euros in prizes (Isabelle Guyon co-organizer). The participants will be asked to predict the power flow on the entire French territory over several years. This challenge will eventually be followed by a challenge in reinforcement learning (RL), in the context of the PhD thesis of Lisheng Sun who just started working on the problem of RL and Automatic Machine Learning (reducing to the largest possible extend thuman intervention in reinforcement learning). Another direction being explored are uses of causal models to improve explainability of predictive models in decision support systems (Inria-funded post-doc Berna Batu). This should allow making more intelligible suggestions of corrective actions to operators to bring network operations back to safety when incidents or stress occur.

Technical challenges: Our work with Artelys focuses on the combination of reinforcement learning tools, with their anytime behavior and asymptotic guarantees, with existing fast approximate algorithms. Our goal is to extend the state of the art by taking into account non-linearities which are often neglected in power systems due to the huge computational cost. We study various modelling errors, such as biases due to finite samples, linearization, and we propose corrections. The work with RTE involves modeling the network itself from archives, because the numerical simulation is both too expensive and not robust, and modeling the client demand in order to be able to predict possible outlier consumptions.

Related Activities:

- Joint team with Taiwan, namely the Indema associate team.

- Organization of various forums and meetings around Energy Management

4.2. Computational Social Sciences

Several projects related to research in social science and humanities and/or research transfer have started in 2015 and continued in 2016:

- Personal semantics (Gregory Grefenstette). In the current digital world, individuals generate increasing amount of personal data. Our work involves discovering semantic axes for organizing and exploiting this data for personal use.

- Gregorius (Cécile Germain & Gregory Grefenstette). An application of semantic structuring and automatic enrichment of existing digital humanities archives.

- Cartolabe (Ph. Caillou, Jean-Daniel Fekete - AVIZ, Gregory Grefenstette, Michèle Sebag). The Cartolabe project applies machine learning techniques to provide a visual, global and dynamic representation of scientific activities from large scale data (HAL at the moment).

- AmiQap (Philippe Caillou, Isabelle Guyon, Michèle Sebag, Paola Tubaro). The multivariate analysis of government questionaire data relative to the quality of life at work, in relation with the socio-economical indicators of firms, aims at investigating the relationship between quality of life and economic performances (conditionally to the activity sector). This will be the topic of the Divyan Kalainathan's PhD, with emphasis on learning causal effect with novel causal discovery algorithms, in collaboration with post-doctoral student Olivier Goudet and researchers at Facebook AI research.

- Collaborative Hiring (Philippe Caillou, Michèle Sebag). Thomas Schmitt's PhD, started in 2014, aims at matching job offers and resumes viewed as a collaborative filtering problem. An alternative approach based on Deep Networks has been started by François Gonard within his IRT PhD.

- Within the U. Paris-Saclay Nutriperso IRS (Philippe Caillou, Flora Jay, Michèle Sebag), we start investigating the relationships between health, diets and socio-demographic features, with the ultimate goal of emitting individual recommendations toward a more healthy diet, such that these recommendations are acceptable.

- Foodtech (Paola Tubaro, Philippe Caillou, Odalric Maillard). An application of agent-based modelling and machine learning to the study of labor conditions in digital platforms. Focus is on online services and mobile applications for food production, delivery, and consumption.

- Sharing Networks (Paola Tubaro). Mapping the "collaborative economy" of internet platforms through social network data and analysis.

- IODS (Wikidata for Science).

Significant challenges include some Big Data problems:

- learning interpretable clusters from bottom-up treatment of heterogeneous textual and quantitative data

- aligning bottom-up clusters with existing manually created top-down structures

- building a unified system integrating the "dire d'experts".

- merging heterogeneous data from different sources.

- moving from predictive to causal discovery algorithms, in line with state-of-the-art research on causality.

Partners:

- Amiqap is funded by the ISN Lidex, with Mines-Telecom SES, RITM (Univ. Paris Sud) and La Fabrique de l'Industrie as partners.

- The collaborative hiring study is funded by the ISN Lidex, in cooperation with J.P. Nadal from EHESS.

- Cartolabe is funded by Inria, in collaboration between TAO and AVIZ.

4.3. High Energy Physics (HEP)

This is joint work with The Laboratoire de l'Accelerateur Lineaire (LAL) https://www.lal.in2p3.fr and the ATLAS and CMS collaborations at CERN. Our principal collaborators at LAL are David Rousseau and Balazs Kegl. The project started in 2015 with the organization of a large world-wide challenge in machine learning that attracted nearly 2000 participants. The theme of the challenge was to improve the statistical significance of the discovery of the Higgs Boson in a particular decay channel, using machine learning. The outcome of the challenge impacted very importantly the methodology used by HEP researchers, introducing new ways of conducting cross-validation to avoid over-fitting and state-of-the-art learning machines, such as XGboost and deep neural networks. The setting of the challenge was purposely simplified to attract easily participants with no prior knowledge of physics. Following the success of the challenge, we decided to dig deeper and re-introduce into the problem more difficulties, including systematic noise.

1. **SystML.** (Cécile Germain, Isabelle Guyon, Michèle Sebag, Victor Estrade, Arthur Pesah): Preliminary explorations were conducted by an intern from ENSTA (Arthur Pesah) and Victor Estrade as an M2 intern. Victor Estrade started in September 2016 his PhD on this subject. The SystML project aims at tackling this problem from 3 angles:

   – calibrating simulators better;

   – using machine learning to train post-hoc correctors of systematic noise;

   – tolerating systematic noise by computing more accurately their effect on the statistical power of tests.

   Exploratory work was performed by Arthur Pesah and Victor Estrade to align the distributions generated by simulators and real data using Siamese networks and adversarial learning. Although good results were obtained on toy data and bioinformatics data, disappointing results were obtained on HEP data. Victor Estrade is now turning to another technique: tangent propagation. This method allows training neural networks, which are robust to "noise" in given directions of feature space.

2. **TrackML.** (Isabelle Guyon): A new challenge is in preparation with LAL and the ATLAS and CMS collaborations. The instantaneous luminosity of the Large Hadron Collider at CERN is expected to increase so that the amount of parasitic collisions can reach a level of 200 interaction per bunch crossing, almost a factor of 10 w.r.t the current luminosity. In addition, the experiments plan a 10-fold increase of the readout rate. This will be a challenge for the ATLAS and CMS experiments, in particular for the tracking, which will be performed with a new all Silicon tracker in both experiments. In terms of software, the increased combinatorial complexity will have to be dealt with within flat budget at best. To reach out to Computer Science specialists, a Tracking Machine Learning challenge (TrackML) is being set up for 2017, building on the experience of the successful Higgs Boson Machine Learning challenge in 2015. The problem setting is to provide participants with coordinates of "hits" that are excitations of detectors along particle trajectories. The goal of the challenge is to devise fast software to "connect the dots" and guess particle trajectories. TAO contributes preparing the challenge platform using Codalab and preparing the challenge protocol and baseline methods.

5. Highlights of the Year

5.1. Highlights of the Year

- Yann Ollivier was invited to contribute to Yann LeCun's official series of talks on Deep Learning at College de France.

- Isabelle Guyon was program chair of the NIPS 2016 conference (in 2017 she will be general chair).

- The TAO team was selected by Microsoft to become the community lead of the competition plaform Codalab. We received a $20 000 Azure for research grant.

- Paola Tubaro co-organized the Second European Social Networks (EUSN) Conference, a major interdisciplinary event for the international research community interested in social networks. Jean-Daniel Fekete (AVIZ) was keynote speaker, and some TAO members contributed papers.

6. New Software and Platforms

6.1. CMA-ES

Covariance Matrix Adaptation Evolution Strategy
KEYWORDS: Numerical optimization - Black-box optimization - Stochastic optimization
SCIENTIFIC DESCRIPTION

The CMA-ES is considered as state-of-the-art in evolutionary computation and has been adopted as one of the standard tools for continuous optimisation in many (probably hundreds of) research labs and industrial environments around the world. The CMA-ES is typically applied to unconstrained or bounded constraint optimization problems, and search space dimensions between three and a hundred. The method should be applied, if derivative based methods, e.g. quasi-Newton BFGS or conjugate gradient, (supposedly) fail due to a rugged search landscape (e.g. discontinuities, sharp bends or ridges, noise, local optima, outliers). If second order derivative based methods are successful, they are usually faster than the CMA-ES: on purely convex-quadratic functions, f(x)=xTHx, BFGS (Matlabs function fminunc) is typically faster by a factor of about ten (in terms of number of objective function evaluations needed to reach a target function value, assuming that gradients are not available). On the most simple quadratic function f(x)=||x||2=xTx BFGS is faster by a factor of about 30.
FUNCTIONAL DESCRIPTION

The CMA-ES is an evolutionary algorithm for difficult non-linear non-convex black-box optimisation problems in continuous domain.

- Participants: Nikolaus Hansen and Emmanuel Benazera
- Contact: Nikolaus Hansen
- URL: https://www.lri.fr/~hansen/cmaesintro.html

6.2. COCO

COmparing Continuous Optimizers
KEYWORDS: Benchmarking - Numerical optimization - Black-box optimization - Stochastic optimization
SCIENTIFIC DESCRIPTION

COmparing Continuous Optimisers (COCO) [65] is a tool for benchmarking algorithms for black-box optimisation. COCO facilitates systematic experimentation in the field of continuous optimization. COCO provides: (1) an experimental framework for testing the algorithms, (2) post-processing facilities for generating publication quality figures and tables, (3) LaTeX templates of articles which present the figures and tables in a single document.

The COCO software is composed of two parts: (i) an interface available in different programming languages (C/C++, Java, Matlab/Octave, Python) which allows to run and log experiments on a suite of test functions. Several testbeds are provided. (ii) a Python tool for generating figures and tables that can be browsed in `html` or used in LaTeX templates.
FUNCTIONAL DESCRIPTION

The COCO platform provides the functionality to automatically benchmark optimization algorithms for bounded or unbounded, (yet) unconstrained optimization problems in continuous domains. Benchmarking is a vital part of algorithm engineering and a necessary path to recommend algorithms for practical applications. The COCO platform releases algorithm developers and practitioners alike from (re-)writing test functions, logging, and plotting facilities by providing an easy-to-handle interface in several programming languages. The COCO platform has been developed since 2007 and has been used extensively within the "Blackbox Optimization Benchmarking (BBOB)" workshop series since 2009. Overall, 140+ algorithms and algorithm variants by contributors from all over the world have been benchmarked with the platform so far and all data is publicly available for the research community. A new test suite of bi-objective problems [74] has been used for the BBOB-2016 workshop at GECCO.

- Participants: Dimo Brockhoff, Arnaud Liefooghe, Thanh-Do Tran, Nikolaus Hansen, Anne Auger, Marc Schoenauer, Ouassim Ait Elhara, Asma Atamna, Tea Tusar and Dejan Tusar
- Partners: Université technique de Dortmund - Université technique de Prague
- Contact: Dimo Brockhoff
- URL: https://github.com/numbbo/coco

6.3. Cartolabe

FUNCTIONAL DESCRIPTION

The goal of Cartolabe is to build a visual map representing the scientific activity of an institution/university/domain from published articles and reports. Using the HAL Database and building upon the AnHALytics processing chain, Cartolabe provides the user with a map of the thematics, authors and articles and their dynamics along time. ML techniques are used for dimensionality reduction, cluster and topics identification; visualisation techniques are used for a scalable 2D representation of the results.

- Participants: Felix Louistisserand, Philippe Caillou, Michèle Sebag, Jean-Daniel Fekete (AVIZ)
- Partners: AVIZ (Inria)
- Contact: Philippe Caillou
- URL: https://cartolabe.lri.fr

6.4. METIS

KEYWORDS: Optimization - Energy
FUNCTIONAL DESCRIPTION

Many works in Energy Optimization, in particular in the case of high-scale sequential decision making, are based on one software per application, because optimizing the software eventually implies losing generality. Our goal is to develop with Artelys a platform, METIS, which can be used for several applications. In 2012 we interfaced existing codes in Artelys and codes developed in the TAO team, experiments have been performed and test cases have been designed. A main further work is the introduction of generic tools for stochastic dynamic programming into the platform, for comparison and hybridization with other tools from the UCT-SIG.

- Participants: Olivier Teytaud, Jeremie Decock, Jean-Joseph Christophe, Vincent Berthier, Marie Liesse Cauwet and Sandra Cecilia Astete Morales
- Partner: Artelys
- Contact: Olivier Teytaud
- URL: https://www.lri.fr/~teytaud/metis.html

6.5. io.datascience

FUNCTIONAL DESCRIPTION

This Data as a Service (DaaS) platform is developed in the context of the Center for Data Science and the TIMCO project. Its overall goals is to exploit the advances in semantic web techniques for efficient sharing and usage of scientific data [41], [73]. A related specific software is the Tester for Triplestore (TFT) software suite, which benchmarks the compliance of sparql databases wrt the RDF standard and publishes the results through the SparqlScore service. The io.datascience platform has been selected for presentation at numerous venues, see section 10.3 for details.

- Contact: Cécile Germain
- URL: https://io.datascience-paris-saclay.fr/

6.6. CodaLab

Keywords Benchmarking, competitions.

Functional Description

Challenges in machine learning and data science are competitions running over several weeks or months to resolve problems using provided datasets or simulated environments. Challenges can be thought of as crowdsourcing, benchmarking, and communication tools. They have been used for decades to test and compare competing solutions in machine learning in a fair and controlled way, to eliminate "inventor-evaluator" bias, and to stimulate the scientific community while promoting reproducible science.

Codalab Competitions (http://competitions.codalab.org) is a project that was started by Microsoft Research in 2013 in which Isabelle Guyon has taken an active part, to promote the use of challenges in Machine Learning and Data Science. The TAO team has been selected to take over the project under Isabelle Guyon's leadership. The transfer has been successfully completed in the fall 2016. New features are being implemented, including developing a Wizard http://staging.chalab.eu/.

With already over 50 public competitions (including this year the Data Science Game, a student Olympiad co-organized by our PhD. student Benjamin Donnot http://www.datasciencegame.com/, the AutoML challenge http://automl.chalearn.org/ [42] and a new contest in the LAP challenge series http://chalearnlap.cvc.uab.es/ [47], co-organized by Isabelle Guyon), Codalab is taking momentum in medical imaging, computer vision, time series prediction, text mining, and other applications. TAO is going to continue expanding Codalab to accommodate new needs. For example, two competitions in preparation – TrackML competition (in High Energy Physics) [72] and the See.4C competition (spatio-temporal time series in collaboration with RTE) [48] – will require code submission, permitting to benchmark methods in a controlled environment. We are re-designing the backend of CodaLab to allow organizers to add more servers to satisfy on-the-fly demands of new competitions. Other features coming soon will be the possibility of interacting with a data generating model (rather than analyzing "canned" data), which enables the organization of reinforcement learning competitions and the possibility of organizing "coopetitions" (a mix of competition and collaboration). Other existing challenge platforms are too restrictive to simulate collaboration between participants and implement "coopetitions". Our starting PhD. student Lisheng Sun designed and implemented a first prototype of coopetition "Beat AutoSKLearn", which was run at the NIPS Challenges in Machine Learning workshop (CiML 2016 http://ciml.chalearn.org/).

- Contact: Isabelle Guyon
- URL: http://competitions.codalab.org

7. New Results

7.1. Optimal Decision Making under Uncertainty

The Tao UCT-SIG is working mainly on mathematical programming tools useful for power systems. In particular, we advocate a data science approach, in order to reduce the model error - which is much more critical than the optimization error, in most cases. Real data are the best way for handling uncertainties. Our main results in 2016 are as follows:

- **Noisy optimization** In the context of stochastic uncertainties, noisy optimization handles the model error by simulation-based optimization. Our results include:

  – It has been conjectured that gradient approximation by finite differences (hence, not a comparison-based method) is necessary for reaching such a simple regret of O(1/N). We answer this conjecture in the negative [32], providing a comparison-based algorithm as good as gradient methods, i.e. reaching O(1/N) - under the condition, however, that the noise is Gaussian.

  – The concept of Regret is widely used in the bandit literature for assessing the performance of an algorithm. The same concept is also used in the framework of optimization algorithms, sometimes under other names or without a specific name. Experimental results on the noisy sphere function show that the approximation of Simple Regret, termed Approximate Simple Regret, used in some optimization testbeds, fails to estimate the Simple Regret convergence rate, and propose a new approximation of Simple Regret, the Robust Simple Regret [22].

- **Capacity Expansion Planning** The optimization of capacities in large scale power systems is a stochastic problem, because the need for storage and connections (i.e. exchange capacities) varies a lot from one week/season to another. It is usually tackled through sample average approximation, i.e. assuming that the system which is optimal on average over the last 40 years (corrected for climate change) is also approximately optimal in general. However, in many cases, data are high-dimensional; the sample complexity, i.e. the amount of data necessary for a relevant optimization of capacities, increases linearly with the number of parameters and can be scarcely available at the relevant scale. This leads to an underestimation of capacities. We suggested the use of bias correction in capacity estimation, and investigated the importance of the bias phenomenon, and the efficiency of both standard and original bias correction tools [53].

- **Multi-armed bandits** We studied the problem of sequential decision making in the context of multi-armed bandits. We provided:

  – An algorithm to handle a non-stationary formulation of the stochastic multi-armed bandit where the rewards are not assumed to be identically distributed, that achieves both a competitive regret and sampling complexity against a best sequence of arms. See [61].

  – An algorithm to handle the task of recommending items (actions) to users sequentially interacting with a recommender system. Users are modeled as latent mixtures of C many representative user classes, where each class specifies a mean reward profile across actions. Both the user features (mixture distribution over classes) and the item features (mean reward vector per class) are unknown a priori. The user identity is the only contextual information available to the learner while interacting. This induces a low-rank structure on the matrix of expected rewards from recommending item a to user b. The problem reduces to the well-known linear bandit when either user-or item-side features are perfectly known. In the setting where each user, with its stochastically sampled taste profile, interacts only for a small number of sessions, we develop a bandit algorithm for the two-sided uncertainty. It combines the Robust Tensor Power Method with the OFUL linear bandit algorithm. We provide the first rigorous regret analysis of this combination. See [63].

- **Confidence intervals for streaming data** We consider, in a generic streaming regression setting, the problem of building a confidence interval (and distribution) on the next observation based on past observed data. The observations may have arbitrary dependency on the past observations and come from some external filtering process making the number of observations itself a random stopping time. In this challenging context, we provide confidence intervals based on self-normalized vector-valued martingale techniques, applied to the estimation of the mean and of the variance. See [69].

- **Forecasting tool for Hydraulic networks** We studied a problem of prediction in the context of the monitoring of an hydraulic network by the French company Prolog-ingenierie. The problem is to predict the value of some specific sensor in the next thirty minutes from the activity of the network

(values of all other sensors) in the recent past. We designed a simple tool for that purpose, based on a random forests. The tool has been tested on data generated from the activity recorded on the Parisian hydraulic network in 2010, 2011 and 2013.

7.2. Continuous Optimization

- **Markov Chain Analysis of Evolution Strategies** The theory of Markov chains with discrete time and continuous state space turns out to be very useful to analyze the convergence of adaptive evolution strategies, including simplified versions of the state-of-the art CMA-ES. Exploiting invariance properties of the objective function and of a wide variety of comparison-based optimisation algorithms, we have developed a general methodology to prove global linear convergence [4]. The constructed Markov chains also show the connection between comparison-based adaptive stochastic algorithms and Markov chain Monte Carlo algorithms. Furthermore, we have continued to work on new theoretical tools that exploit deterministic control models to prove the irreducibility and T-chain property of general Markov chains. These tools promise to trivialise some stability proofs of the Markov chains we are interested in to analyse.

- **Large-scale Optimisation Algorithms** We have been working on (improved) variants of CMA-ES with more favorable scaling properties with the dimension. While computing and using the natural gradient in appropriate subspaces turned out to be considerably more difficult than expected, we explored variants that restrict the covariance via projection, so-called VkD-CMA-ES [21]. We derived a computationally efficient way to update the restricted covariance matrix, where the richness of the model is controlled by the integer parameter k. This parameter provides a smooth transition between the case where only diagonal elements are subject to changes and changes of the full covariance matrix. In the latter case, the update is equivalent with the original CMA-ES. In order to get rid of the control parameter we propose an adaptation of k which turns out to be surprisingly efficient [20].

- **Analysis of Lagrangian based Constraints Handling in Evolution Strategies** We have addressed the question of linear convergence of evolution strategies on constrained optimisation problems with one linear constraint. Based on previous works, we consider an adaptive augmented Lagrangian approach for the simple (1+1)-ES [23] and for the CMA-ES [24]. By design both algorithms derive from a framework with an underlying homogenous Markov chain which paves the way to prove linear convergence on a comparatively large class of functions. For the time being, stability of the Markov chain, associated with linear convergence, has been shown empirically on convex-quadratic and ill-conditioned functions.

- **Benchmarking of continuous optimizers** We have been pursuing our efforts towards improving the standards in benchmarking of continuous optimisers [65], [66], [64]. Three new testbeds have been developed and implemented. (i) A bi-objective testbed [74] where also a corresponding performance assessment procedure has been advised [62]. In this context, a new version of MO-CMA-ES has been developed and benchmarked [44] on this testbed. (ii) A large-scale testbed, as a straight forward extension of the standard tested. The extension is based on a general methodology we have developed to construct non-trivial but scalable test functions [19]. (iii) a constrained testbed (unpublished).

7.3. Data Science

- **High Energy Physics** The focus of the period has been to expand the collaboration with the High Energy Physics experiments started with the success of the 2014 HiggsML challenge [18] to new issues. The subject of V. Estrade Phd is to advance domain adaptation methods in the specific context of uncertainty quantification and calibration. So far, transfer learning has been addressed only with classical, additive and differentiable objective functions as performance criteria. However, learning to discover, exemplified by HEP, relies on more global and difficult criteria, related to the Area Under Roc Curve (AUC) and Neymann-Pearson learning. CERN funds another PhD (A. Pol), on anomaly detection. Another promising theme has emerged with the ongoing organization of a Tracking Challenge (TrackML) [56], [72], which focuses on extreme scaling of ML image processing.

- **Personal Semantics** Our algorithm for inducing a taxonomy from a set of domain terms, that was placed first in the international Taxonomy Induction task, part of the SemEval 2015 conference in Denver, has been improved by the development of a robust technique for discovering the domain vocabulary for a new topic using a directed crawler we created. We have created hundreds of taxonomy for personal themes (hobbies, illnesses) that can be integrated into our Personal Semantics platform PTraces, and have deployed and evaluated the taxonomies. We also have introduced newer machine learning methods, such as Latent Dirichlet Allocation, for better recognition of domain vocabularies [55], [71].

- **Distributed system observation** The work on distributed system automated analaysis and description has been persued thru the continued development of the GAMA multi-agent framework https://github.com/gama-platform/gama/wiki. The simulation framework has been applied to the study of a new protocol for MOOC management [6]. Philippe Caillou is associated to the young researcher ANR ACTEUR, coordinated by Patrick Taillandier (IDEES, Rouen university). With this project, the BDI cognitive agent model has been improved both in term of flexibility and ease of use for the non expert modeler [50].

- **Computational social sciences** Thomas Schmitt's PhD focuses on the matching of job offers and applicant CVs. An informal collaboration with the Qapa agency (FUI proposal underway) gave us access to the 2012-2016 logs of their activity (CVs, job announcements and application clicks). This wealth of data delivered some unexpected findings, e.g., as to the differences between people's practice (the clicks) and their say (the documents). In [49], with Philippe Caillou and Michèle Sebag, a deep NN system MAJORE (MAtching JObs and REsumes) was proposed, trained to match the metric properties extracted from the collaborative filtering matrix, and address the cold start problem. A further research perspective, in collaboration with J.-P. Nadal from EHESS, is to build an observatory of the job demand dynamics.

  The Cartolabe project, started in Feb. 2016 (F. Louistisserand's engineer stint), applies machine learning techniques to build an interpretable representation from vast amounts of scientific articles. The goal is to use raw textual data, and the results of the pre-processing chain achieved by ANHALYTICS, to define a topology on authors, scientific themes, and teams, and enforce its 2D projection in a semantically admissible way. The collaboration with AVIZ is key to enable the scalable and navigable exploitation of this map. The perspective for 2017 is to build a visual interrogation of the map (locating all author names relevant to a given request) and to display the temporal evolution of the research activities.

  Amiqap studies the relation between quality of life at work and company performance, using both survey data on individual workers (collected by DARES, the statistical service of the French Ministry of labor, in 2013) and administrative data on companies provided by SECAFI, a union body. The study is run by a team within TAO (Philippe Caillou, Isabelle Guyon, Michèle Sebag and Paola Tubaro, plus post-doctoral researcher Olivier Goudet and intern Diviyan Kalainathan) in collaboration with Mines ParisTech social science and economics (SES) department, the RITM economics research center (Univ. Paris Sud) and the think-tank La Fabrique de l'Industrie. In its first stage, the exploratory analysis delivered some unexpected results, e.g. as to the existence of a "industry worker cluster", or the non-monotonous relationship between autonomy, salary and subjective satisfaction. A summary of these findings has been released online on the website of La Fabrique de l'Industrie, as a complement to their book on the same topic (published in October 2016). The exploratory analysis of the SECAFI data (yet unpublished) complements the above and shows how workers' satisfaction correlates with companies' financial and social performance indicators, though with marked differences across industries. The key question regards the nature of this relationship: cause, effect or due to a confounder feature (the industrial sector). Further research (Diviyan Kalainathan's PhD, O. Goudet post-doc) will focus on the use and extension of causal modelling algorithms on this issue; these perspectives attract quite some interest from the ministry (DARES) and big industrial players, willing to assess the relevance of their HR policies.

7.4. Designing criteria

- **Algorithm selection and configuration** Two PhD theses are related to the former *Crossing the Chasm* SIG: Nacim Belkhir (CIFRE PhD with Thalès) is working on Per Instance Algorithm Configuration (PIAC) in the context of continuous optimization. He has worked on the use of surrogate models for feature computation in case of expensive objective functions [31] and has validated his work with Differential Evolution applied to BBOB testnehc [30]. Defence planned for March 2017.

  François Gonard's PhD is dedicated to optimization algorithm selection. The original application domain was that of expensive car industry simulations (within the IRT-ROM project). The lack of real test cases made him investigate some combinatorial optimization setting, for which there exist public datasets. François obtained a "Honorable mention from the jury" for his submission to the ICON Challenge (http://iconchallenge.insight-centre.org/), for its original approach coupling a pre-scheduler and an algorithm selector [39]. Defence is planned for November 2017.

  The work done during Mustafa Misir's post-doc stint (ERCIM 2013-2014), regarding the formalization and tackling of the algorithm selection problem in terms of a collaborative filtering problem, was finally published [15].

- **A statistical physics perspective** Our activity on probabilistic model design is progressively moving from static explicit interactions to dynamical ones and to latent variable models, taking inspiration from latent feature representations provided by deep learning techniques. Concerning explicit pairwise interactions models like in [14] initially motivated by traffic applications, a systematic treatment of loop corrections based on a minimal cycle basis [11] has led us to propose: (i) a fast and large scale generalized belief propagation method (GCBP) with more robust convergence properties than bare belief propagation (ii) an inverse approximate MRF with linear scaling of the computational time, compliant with GCBP (iii) a new sampling method based on extracting random sub-graph of tree-width 2 on which GCBP can provide exact marginals. More generally considering effect of problematic i.e. frustrated cycles open the possibility for new criteria in model design. In particular we have started to bridge this work with the analysis of multi-layer restricted Boltzmann machines (RBM). Remarkably these possess a planar dual representation and we are expecting the density of frustrated cycles nodes to play a key role when characterizing an RBM learned from structured data by contrast with purely random instances. Additionally we have identify some properties of the data themeselves that have to be taken into consideration when learning static [9] or dynamical [8] Ising models.

- **Artificial Immune Systems** Within the E-Lucid project with Thalès TERESIS, around anomaly detection in network trafic, a first approach has been developed using Artificial Immune System (AIS) and the concept of Voronoi representation. A first proof of concept was a poster at the GECCO conference [70], before a complete paper was published at the PPSN conference [46]. Note that this work on anomaly detection is on-going using Deep Learning. AIS are also the basis of Chaouki Boufenar's PhD work (visiting TAO from U. Oran, Algérie), with a first work on arabic characters recognition [5].

7.5. Deep Learning and Information Theory

- **Neural networks for computer vision** We continued working on the topic of large-scale image segmentation with multiple object detection. The application target is the analysis of high-resolution multispectral satellite images covering the Earth. Challenges are numerous: finding good features to distinguish objects, obtaining fine-resolution segmentations, while dealing with badly-registered groundtruth, keeping a scalable complexity, while avoiding boundary effects when tiling a big image into small ones, which are processed independently and merged back together. We propose to move to fully convolutional neural networks [45] to avoid artifacts from patch-based approaches. We show the benefits of training first on imprecise groundtruth, which is available in large amounts, and then refining on precise but scarce groundtruth [13]. To further refine the segmentation, as

convolutional networks tend to produce blurry outputs, we use recurrent neural networks to learn the partial differential equation (PDE) which would sharpen the segmentations, i.e. an iterative process taking into account the edges in the original image to locate precisely their boundaries and to sharpen them [67]. Finally, to benefit simultaneously from information at various resolutions, we design a new, more suitable architecture [68].

We also started to work on medical image classification, in the long-term goal of automatic diagnosis, in collaboration with the Necker Hospital and the Inria start-up Therapixel, and on image labelling and representation, with the database editor company Armadillo, through the Adamme project (cf Section 9.2.1).

In collaboration with the University of Barcelona, we organize a series of challenges in video analysis of human behavior (ChaLearn Looking at People series). Looking at People (LAP) is an area of research that deals with the problem of automatically recognizing people in images, detecting and describing body parts, inferring their spatial configuration, performing action/gesture recognition from still images or image sequences, often including multi-modal data. Any scenario where the visual or multi-modal analysis of people takes a key role is of interest to us within the field of Looking at People. We have been leaders in organizing challenges in this area since 2013 [10], [12], [36], organizing events sponsored by DARPA, NSF, Microsoft, Google, Facebook, NVIDIA, and others. In 2016 we organized follow up competitions on gesture recognition [52] and face aging [37] to advance the state-of-the-art in areas we had previously explored. We also organized two rounds of a completely new recognition on personality trait evaluation from short video clips [47], [34]. The purpose of this study is to evaluate whether human first impression judgements are consistent and reproducible. Such research could lead to device coaching curricula to help job applicants present themselves better and hiring managers to overcome unsubstantiated negative biases. The winners of the challenge used Deep Learning methods. The third place winners teamed up with the organizers to put together a demonstration system, which was shown at the NIPS conference(https://nips.cc/Conferences/2016/Schedule?showEvent=6314). Work performed in collaboration with UC Berkeley on fingerprint verification using Deep Learning was also presented in this demonstration.

- **Natural Gradients for Deep Learning** Deep learning is now established as a state-of-the-art technology for performing different tasks such as image or sequence processing. Nevertheless, much of the computational burden is spent on tuning the hyper-parameters. On-going work, started during the TIMCO project, is proposing, in the framework of Riemannian gradient descents, invariant algorithms for training neural networks that effectively reduce the number of arbitrary choices, e.g., affine transformations of the activation functions or shuffling of the inputs. Moreover, the Riemannian gradient descent algorithms perform as well as the state-of-the-art optimizers for neural networks, and are even faster for training complex models. The proposed approach is based on Amari's theory of information geometry and consists in practical and well-grounded approximations for computing the Fisher metric. The scope of this framework, going beyond Deep Learning, encompasses any class of statistical models. This year's contribution is a new, simple framework (both theoretical and practical) that allowed us to release a simpler implementation of these techniques in Torch (one of the main deep learning libraries in use) and demonstrate good performance on real data. We have also started to explore criteria from information geometry criteria for automating the construction and selection of network architectures themselves, a major problem given the current trend towards highly complex, hand-built model architectures (P. Wolinski's PhD).

- **Training dynamical systems online without backtracking** with application to recurrent neural networks. The standard way to train recurrent neural networks and other systems that exhibit a temporal dynamical behavior involves "backpropagation through time", which as the name indicates goes backward in time and is unrealistic. Last year we proposed an algorithm to learn the parameters of a dynamical system in an online, memoryless setting, thus scalable and requiring no backpropagation through time, in a way guaranteed to be unbiased. This year we started to provide full convergence proofs for this algorithm (the first of their kind). Moreover Corentin Tallec (PhD) proposed a considerably simpler version of the algorithm keeping the same key mathematical

properties, which now allows for a simple "black-box" implementation on top of any existing recurrent network model.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- **Thales Research & Technology** 2014-2017 (30 kEuros), related to Nacim Belkhir's CIFRE PhD
  Coordinator: Marc Schoenauer
  Participants: Johann Dréo, Pierre Savéant, Nacim Belkhir
- **Orange** 2013-2016 (30 kEuros), related to Robin Allesiardo's CIFRE PhD
  Coordinator: Michèle Sebag
  Participants: Raphael Feraud, Robin Allesiardo
- **Réseau Transport d'Electricité** 2015-2018 (30 kEuros), related to Benjamin Donnot's CIFRE PhD
  Coordinator: Olivier Teytaud (until May 2016), now Isabelle Guyon and Marc Schoenauer
  Participants: Benjamin Donnot, Antoine Marot

9. Partnerships and Cooperations

9.1. Regional Initiatives

- **PGMO NUMBBER** 2016-2018 (60 kEuros)
  Coordinator: FMJH(Fondation mathématiques Jacques Hadamard - Paris Saclay) & Anne Auger
  Participants: Anne Auger, Nikolaus Hansen

9.2. National Initiatives

- **ROM**: *Model Reduction and Multiphysics Optimization* 2014-2016 (50 Keuros)
  Coordinator: IRT System X
  Participants: Marc Schoenauer, Michèle Sebag, François Gonard (PhD)
- **MAJORE** *A Collaborative Filtering Approach to Matching Job Openings and Job Seekers*, 2013-2016 (105 kEuros)
  Thomas Schmitt's PhD (funded by ISN).
  Participants: Philippe Caillou, Michèle Sebag, Thomas Schmitt (PhD)
- **AutoML** *An empirical approach to Machine Learning* 2014-2017 (104 kEuros)
  Sourava Mishra's PhD
  Participants: Michèle Sebag, Balazs Kégl, Sourava Mishra
- **ReMODeL** *Rewarded Multimodal Online Deep Learning* 2015-2016 (31,5 kEuros)
  This project lies at the junction of reinforcement learning, deep learning, computational neuroscience and developmental robotic fields. It is closely related to the transversal DIGITEO robotic theme, Roboteo.
  Participants: Michèle Sebag, Mathieu Lefort, Alexander Gepperth
- **AMIQAP** 2015-2016 (12 months of Postdoctoral fellow). Project funded by ISN
  Participants: Philippe Caillou, Olivier Goudet, Isabelle Guyon, Michèle Sebag, Paola Tubaro, Diviyan Kalavanathan (2016 intern, 2017 PhD)
- **NUMBBO** 2012-2017 (290kEuros for TAO). Analysis, Improvement and Evaluation of Numerical Blackbox Optimizers, ANR project, Coordinator Anne Auger, Inria. Other partners: Dolphin, Inria Lille, Ecole des Mines de Saint-Etienne, TU Dortmund
  Participants: Anne Auger, Nikolaus Hansen, Marc Schoenauer, Ouassim Ait ElHara

- **ACTEUR** 2014-2018 (236kEuros). Cognitive agent development for urban simulations, ANR project, Coordinator P. Taillandier (IDEES, Univ Rouen).

  Participant: Philippe Caillou

9.2.1. Other

- **POST** 2014-2017 (1,220 MEuros, including 500 kEuros for a 'private' cluster). Platform for the optimization and simulation of trans-continental grids
  ADEME (Agence de l'Environnement et de la Maîtrise de l'Energie)
  Coordinator: ARTELYS
  Participants: Olivier Teytaud, Marie-Liesse Cauwet, Jérémie Decock, Sandra Cecilia Astete Morales, David L. Saint-Pierre, J. Decock
- **E-LUCID** 2014-2017 (194 kEuros)
  Coordinator: Thales Communications & Security S.A.S
  Participants: Marc Schoenauer, Cyril Furtlehner
- **PIA ADAMME** 2015-2018 (258 kEuros)
  Coordinator: Bull SAS
  Participants: Marc Schoenauer, Yann Ollivier, Gaetan Marceau Caron, Guillaume Charpiat, Cécile Germain-Renaud, Michèle Sebag
- **CNES contract** 2015-2017 (70 kEuros)
  Coordinator: Manuel Grizonnet (CNES) & Yuliya Tarabalka (Inria Sophia-Antipolis, Titane team)
  Participant: Guillaume Charpiat
- **ESA Tender** 2016-2017 (52 kEuros)
  Coordinator: Oana Togt (TNO) & Marc Schoenauer
  Participant: Marc Schoenauer

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

EHRI-II 2015-2019 (7 969 kEuros). European Holcaust Research Infrastructure, H2020, Coordinator NIOD, Amsterdam. Digital Humanities.
Participants: Gregory Grefenstette

See.4C 2016-2017 (2 700 kEuros). SpatiotEmporal ForEcasting: Coopetition to meet Current Cross-modal Challenges
Participants: Isabelle Guyon

9.3.2. Collaborations with Major European Organizations

MLSpaceWeather 2015-2019. Coupling physics-based simulations with Artificial Intelligence.
Coordinator: CWI
Participants: Michèle Sebag, Aurélien Decelle, Cyril Furtlehner.

ESA tender 2015-2016, through collaboration with TNO (see Section 9.2.1).

9.4. International Initiatives

9.4.1. CIADM

Title: Computational intelligence and Decision making

International Partner (Institution - Laboratory - Researcher):

  NUTN (Taiwan) - Multimedia Informatics Lab - Chang-Shing Lee

Start year: 2015

See also: http://www.lri.fr/~teytaud/indema.html

The associate team works on computation intelligence for decision making, with different application fields for the various partners: - power systems (Tao) - eLearning (Oase) - games (Ailab)

9.4.2. S3-BBO

Title: Threefold Scalability in Any-objective Black-Box Optimization

International Partner (Institution - Laboratory - Researcher):

> Shinshu (Japan) - Tanaka-Hernan-Akimoto Laboratory - Hernan Aguirre

Start year: 2015

See also: http://francejapan.gforge.inria.fr/doku.php?id=associateteam

This associate team brings together researchers from the TAO and Dolphin Inria teams with researchers from Shinshu university in Japan. Additionally, researchers from the University of Calais are external collaborators to the team. The common interest is on black-box single and multi-objective optimization with complementary expertises ranging from theoretical and fundamental aspects over algorithm design to solving industrial applications. The work that we want to pursue in the context of the associate team is focused on black-box optimization of problems with a large number of decision variables and one or several functions to evaluate solutions, employing distributed and parallel computing resources. The objective is to theoretically derive, analyze, design, and develop scalable black-box stochastic algorithms including evolutionary algorithms for large-scale optimization considering three different axes of scalability: (i) decision space, (ii) objective space, and (iii) availability of distributed and parallel computing resources.

We foresee that the associate team will make easier the collaboration already existing through a proposal funded by Japan and open-up a long term fruitful collaboration between Inria and Shinshu university. The collaboration will be through exchanging researchers and Ph.D. students and co-organization of workshops.

9.4.3. Informal International Partners

Marc Schoenauer partner of the ARC-DP (Australian Research Council Discovery Project) *bio-inspired computing methods for dynamically changing environments*. Coordinator: University of Adelaide (Frank Neumann), 5 years from Nov. 2015, 400 k$-AUS. Visit to Adelaide planned in Feb. 2017.

Isabelle Guyon partner of UC Berkeley *Fingerprint verification with deep siamese neural networks using ultratonic sensor data.* Co-advisor of a master student (Baiyu Chen). Partners: Alyosha Efros, Bernhard Boser.

9.4.4. Participation in Other International Programs

9.4.4.1. Indo-French Center of Applied Mathematics

Contextual multi-armed bandits with hidden structure

Title: Contextual multi-armed bandits with hidden structure

International Partner (Institution - Laboratory - Researcher):

> IISc Bangalore (India) - ECE - Aditya Gopalan

Duration: 12 months - April 2017

Start year: April 2016

Recent advances in Multi-Armed Bandit (MAB) theory have yielded key insights into, and driven the design of applications in, sequential decision making in stochastic dynamical systems. Notable among these are recommender systems, which have benefited greatly from the study of contextual MABs incorporating user-specific information (the context) into the decision problem from a rigorous theoretical standpoint. In the proposed initiative, the key features of (a) sequential interaction between a learner and the users, and (b) a relatively small number of interactions per user with the system, motivate the goal of efficiently exploiting the underlying collective structure of users.

The state-of-the-art lacks a wellgrounded strategy with provably near-optimal guarantees for general, low-rank user structure. Combining expertise in the foundations of MAB theory together with recent advances in spectral methods and low-rank matrix completion, we target the first provably near-optimal sequential low-rank MAB

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- **Edgar Galvan Lopez** University College Dublin, April 2015 - December 2016, funded by the ELEVATE Fellowship, the Irish Research Council's Career Development Fellowship co-funded by Marie Curie Actions.

9.5.1.1. Internships

- **Borja Seijo** Universidade da Coruña, Galicia, Spain, October-November 2016, self-funded. Worked on missing data under the supervision of Isabelle Guyon.
- **Tomas Lungenstrass**, June 2016 - June 2017, self-funded. Worked on magnetic storm prediction under A. Decelle's, C. Furtlehner's and M. Sebag's supervision.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

Isabelle Guyon, Program Chair of NIPS 2016

10.1.1.2. Member of the Organizing Committees

- Anne Auger, co-organizer of the GECCO workshop on Black Box Optimization Benchmarking.
- Cécile Germain, Scientific Committee, DataScience@HEP 2016
- Nikolaus Hansen, co-organizer of the GECCO workshop on Black Box Optimization Benchmarking.
- Marc Schoenauer, Steering Committee, Parallel Problem Solving from Nature (PPSN); Steering Committee, Learning and Intelligent OptimizatioN (LION).
- Michele Sebag, President of Steering Committee, Eur. Conf. on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD).
- Isabelle Guyon, co-organizer of two NIPS workshops (Challenges in Machine Learning and Spatio-temporal time series), co-organizer of AutoML workshop at ICML, co-organizer of LAP challenge workshops (ECCV, ICPR).
- Paola Tubaro, co-organizer of the Second European Social Networks (EUSN) Conference.

10.1.1.3. Member of Conference Program Committees

All TAO members are members of the Program Committees of the main conferences in the fields of Machine Learning, Evolutionary Computation, and Information Processing.

10.1.1.4. Reviewer

All TAO member review papers for the most prestigious journals in the fields of Machine Learning and Evolutionary Computation.

10.1.2. Journal

10.1.2.1. Member of the Editorial Boards

- Anne Auger, member of Editorial Board, *Evolutionary Computation Journal*, MIT Press.
- Gregory Grefenstette, member of Editorial Board, *Journal of Natural Language Engineering*, Cambridge University Press.
- Isabelle Guyon, action editor, *Journal of Machine Learning Research* (JMLR).
- Isabelle Guyon, series editor, *Microtome book series Challenges in Machine Learning* (CiML).
- Nikolaus Hansen, member of Editorial Board, *Evolutionary Computation Journal*, MIT Press.
- Marc Schoenauer, member of Advisory Board, *Evolutionary Computation Journal*, MIT Press; member of Editorial Board, *Genetic Programming and Evolutionary Machines*, Springer Verlag; action editor, *Journal of Machine Learning Research*(JMLR).
- Michèle Sebag: Editorial Board, Machine Learning, Springer Verlag.
- Olivier Teytaud, action editor, *Journal of Machine Learning Research* (JMLR).
- Paola Tubaro, member of Associate Editorial Board, *Sociology*, Sage; member of Editorial Board, *Revue Française de Sociologie*, Presses de Sciences Po.

10.1.2.2. Reviewer - Reviewing Activities

All members of the team reviewed numerous articles for international conferences and journals.

10.1.3. Invited Talks

- Isabelle Guyon. 17 Oct. ENS SMILE seminars. Network Reconstruction: the Contribution of Challenges in Machine Learning.
- Michele Sebag, 21 Jan., Deep Learning, IHP
- Michele Sebag, 24 Mar., Nutrition personnalisée et alimentation sur mesure, AgroParisTech.
- Michele Sebag, 12 Oct., Deep Learning and Artificial Intelligence, Franco-Japanese Symposium. Tokyo
- Paola Tubaro. 8 Mar. University of Insubria (Italy) economics department seminar. Investigating peer effects and performance similarity in organizational networks: a longitudinal study.
- Paola Tubaro. 15 Mar. EHESS seminar. Sociabilité et soutien social dans les communautés en ligne autour des troubles de l'alimentation.
- Paola Tubaro. 2 Nov. UQAM (Montréal), Seminar of the health communication research center (ComSante). Le phénomène 'pro ana' : Troubles alimentaires et réseaux sociaux.
- Paola Tubaro. 8 Nov. Université Laval (Québec), Seminar of the CELAT research center. Le phénomène 'pro ana' : Troubles alimentaires et réseaux sociaux.

10.1.4. Leadership within the scientific community

- Isabelle Guyon, President and co-founder of ChaLearn, a non-for-profit organization dedicated to the organization of challenge. http://chalearn.org
- Marc Schoenauer, elected Chair of ACM-SIGEVO (Special Interest Group on Evolutionary Computation), July 2015 (2-years term).
- Marc Schoenauer, founding President of SPECIES (Society for the Promotion of Evolutionary Computation In Europe and Surroundings), that organizes the yearly series of conferences *EvoStar*.
- Michèle Sebag, elected Chair of Steering Committee, ECML-PKDD; head of the Research Committee of Labex Digicosme.
- Paola Tubaro, convenor of the Social Network Analysis Group of British Sociological Association.

10.1.5. Scientific expertise

- Cécile Germain, evaluator for the H2020 calls: *ICT-2015 Topic ICT-16 – Big Data - research*
- Gregory Grefenstette, evaluator for FU21, Cap Digital, Digiteo (IASI) review board
- Gregory Grefenstette, project reviewer for the H2020 (SemCare): *Information and Communication Technologies ICT*
- Michele Sebag, evaluator for the Swedish Foundation for Strategic Research
- Michele Sebag, member of hiring jury for U. Nancy
- Paola Tubaro, evaluator for ANR, SNSF (Swiss National Science Foundation), and National Science Center (Poland).

10.1.6. Research administration

- Philippe Caillou, elected member of the Scientific Council and Academic Council.
- Cécile Germain, elected member of the U-PSUD Scientific Council and of its board. University officer for scientific computing. Deputy head of the computer science departement, in charge of research, member of the Board of the Lidex *Center for Data Science*.
- Marc Schoenauer, *Délégué Scientifique* (aka VP-Research) for the Inria Saclay Île-de-France branch until June 2016; co-chair (with Sylvain Arlot) of the *Maths-STIC* program of the Labex of Mathematics Hadamard (LMH).
- Michele Sebag, deputy director of LRI, CNRS UMR 8623; elected member of the Research Council of Univ. Paris-Saclay; member of the STIC department council; member of the Board of the Lidexes *Institut de la Société Numérique* and *Center for Data Science*.
- Paola Tubaro, member of the steering team, MSH Paris-Saclay, axis 1 "Power of algorithms".

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence : Philippe Caillou, Computer Science for students in Accounting and Management, 192h, L1, IUT Sceaux, Univ. Paris Sud.

Licence : Aurélien Decelle, Computer Architecture, 60h, L2, Univ. Paris-Sud.

Licence : Aurélien Decelle, Machine Learning and Artificial Life, 55h, L2, Univ. Paris-Sud.

Licence : Aurélien Decelle, Computer Architecture, 41h, L3, Univ. Paris-Sud.

Licence and Polytech : Cécile Germain, Computer Architecture

Licence : Isabelle Guyon, Project: Creation of mini-challenges, M2, Univ. Paris-Sud.

Master : Anne Auger, Optimisation, 12h, M2 Recherche, U. Paris-sud.

Master : Guillaume Charpiat et Gaétan Marceau, Advanced Machine Learning, 34h, M2 Recherche, Centrale-Supélec.

Master : Aurélien Decelle, Machine Learning, 27h, M1, Univ. Paris-Sud.

Master : Aurélien Decelle, Information theory, 39h, M1, Univ. Paris-Sud.

Master : Cécile Germain, Parallel Programming

Master : Isabelle Guyon, Project: Resolution of mini-challenges (created by M2 students), L2, Univ. Paris-Sud.

Master : Odalric-Ambrym Maillard, Machine Learning, 6h, M2 Recherche, Univ. Paris-Sud

Master : Yann Ollivier, Deep learning, 4h, M2 Recherche, Telecom/Polytech.

Master : Michèle Sebag, Machine Learning, 12h; Deep Learning, 6h; Reinforcement Learning, 6h; M2 Recherche, U. Paris-sud.

Master : Paola Tubaro, Sociology of social networks, 24h, M2, EHESS/ENS/ENSAE.

Master : Paola Tubaro, Digital platforms, online socialization and new economic models, 6h, M2, Arts et Métiers ParisTech (ENSAM).

Doctorate: Paola Tubaro, Research Methods, 9h, University of Insubria, Italy.

10.2.2. Supervision

PhD: Jérémy BENSADON, *Applications of Information Theory to Statistical Learning*, Univ. Paris-Saclay, 02/02/2016, Yann Ollivier.

PhD: Marie-Liesse CAUWET, *Artificial intelligence with uncertainties, application to power systems*, Univ. Paris-Saclay, 30/9/2016, Olivier Teytaud.

PhD: Sandra ASTETE-MORALES, *Noisy optimization, with applications to power systems*, Univ. Paris-Saclay, 5/10/2016, Olivier Teytaud.

PhD: Robin ALLESIARDIO, *Multi-armed Bandits on non Stationary Data Streams*, Univ. Paris-Saclay, 19/10/2016, Raphaël Féraud (Orange Labs) and Michèle Sebag.

PhD in progress: Ouassim AIT ELHARA, *Large-scale optimization and Evolution Strategies*, 1/09/2012, Anne Auger and Nikolaus Hansen.

PhD in progress: Asma ATAMNA, *Evolution Strategies and Constrained Optimization*, 1/10/2013, Anne Auger and Nikolaus Hansen.

PhD in progress: Nacim BELKHIR, *On-line parameter tuning*, 1/5/2014, Marc Schoenauer and Johann Dréo (Thalès), CIFRE Thalès.

PhD in progress: Vincent BERTHIER, *Large scale parallel optimization, with application to power systems*, 1/09/2013, Michèle Sebag et Olivier Teytaud.

PhD in progress: Mehdi CHERTI *Learning to discover: supervised discrimination and unsupervised representation learning with applications in particle physics.* 01/10/2014, Balazs Kegl and Cécile Germain.

PhD in progress : Benjamin DONNOT, *Optimisation et méthodes d'apprentissage pour une conduite robuste et efficace du réseau électrique par anticipation sur base de parades topologiques.*, 1/09/2015, Isabelle Guyon and Marc Schoenauer

PhD in progress : Guillaume DOQUET, *ML Algorithm Selection and Domain Adaptation*, 1/09/2015, Michele Sebag

PhD in progress: Victor ESTRADE *Robust domain-adversarial learning, with applications to High Energy Physics,* 01/10/2016, Cécile Germain and Isabelle Guyon.

PhD in progress: François GONARD, *Automatic optimization algorithm selection and configuration*, 1/10/2014, Marc Schoenauer and Michèle Sebag, thèse IRT SystemX.

PhD in progress : Hoang M. LUONG, *Squaring the Circle in Modelling Corporate Governance, Market Structure and Innovation: A Tobin's Q Approach to R&D Investment when Network Effects Are Present*, 01/09/2014, (with M. Ugur and S. Gorgoni, at the University of Greenwich, London, UK).

PhD in progress : Emmanuel MAGGIORI, *Large-Scale Remote Sensing Image Classification*, 1/1/2015, Guillaume Charpiat (with Yuliya Tarabalka and Pierre Alliez, Inria Sophia-Antipolis)

PhD in progress: Pierre-Yves MASSÉ, *Gradient Methods for Statistical Learning*, 1/10/2014, Yann Ollivier

PhD in progress: Sourava MISHRA, *AutoML: An empirical approach to Machine Learning*, 1/10/2014, Balazs Kégl and Michèle Sebag

PhD in progress : Anna PIAZZA, *Inter-Organisational Relationships and Organisational Performance: Network Analysis Applications to a Health Care System*, 01/09/2014, Paola Tubaro (with F. Pallotti and A. Lomi, at the University of Greenwich, London, UK).

PhD in progress: Adrian POL *Machine Learning Anomaly Detection, with application to CMS Data Quality Monitoring,* 01/10/2016, Cécile Germain.

PhD in progress: Karima RAFES *Gestion et sécurité des données personnelles dans le web des objets*. 01/10/2014, Serge Abiteboul and Cécile Germain.

PhD in progress : Yasaman SARABI, *Network Analysis of Private Water Companies, Challenges Collaboration and Competition*, 15/03/2012, Paola Tubaro (at the University of Greenwich, London, UK).

PhD in progress: Thomas SCHMITT, *A Collaborative Filtering Approach to Matching Job Openings and Job Seekers*, 1/11/2014, Philippe Caillou and Michèle Sebag and Jean-Pierre Nadal (EHESS)

PhD in progress : Lisheng SUN, *Apprentissage Automatique: Vers une analyse de données automatisé*, 1/10/2016, Isabelle Guyon and Michèle Sebag

PhD in progress : Corentin TALLEC, *Reinforcement Learning and Recurrent Neural Networks: dynamical approaches*, 1/10/2016, Yann Ollivier

PhD in progress : Pierre WOLINSKI, *Learning the Architecture of Neural Networks*, 1/9/2016, Guillaume Charpiat and Yann Ollivier

10.2.3. Juries

Marc Schoenauer, PhD jury of Jonathan GUERRA (ISAE, Toulouse), External reviewer of HDR of David Gianazza (IRIT, Toulouse).

Cyril Furtlehner, PhD jury of Christophe Schülke (Université Paris Diderot).

Isabelle Guyon, PhD jury of Mathieu Bouyrie (AgroPariTech), November 29, 2016. Restauration d'images de noyaux cellulaires en microscopie 3D par l'introduction de connaissance a priori.

Guillaume Charpiat, PhD jury of Thomas Bonis (Inria Saclay); jury for the SIF best PhD prize (Gilles Kahn).

Nikolaus Hansen, PhD jury of Oswin Krause (University of Copenhagen), January 8, 2016. External reviewer.

Michele Sebag, HdR jury of Matthieu Geist (U. Lille); PhD Jury Jiaxin Kou (Royal Holloway, London); PhD Phong N'guyen (U. Geneve).

10.3. Popularization

- **Yann Ollivier** coordinated, with IHP, numerous activities for the centenary of Claude Shannon, including a public exhibition at the Musee des Arts et Metiers, a cycle of public conferences, a contest for teachers on the best IT class project, and a workshop on the current state of information theory.

- **Yann Ollivier**, co-organizes the European Union Contest for Young Scientists (science fair for high school students from 30+ countries organized by the European Commission).

- **Aurélien Decelle**, participation to "la fête de la science" animating a presentation of arduino to high school students and families at Inria Turing.

- **Paola Tubaro**, invited talk on "Online social networks and eating disorders", ACT eating disorders association, Nîmes, 5 Feb.; invited talk on "Are we all digital laborers?", Autonomy salon of urban mobility, Paris, 8 Oct.; public interview, "L'économie peut-elle être collaborative? Rencontre avec la sociologue et chercheure Paola Tubaro", Montréal, 2 Nov.; panelist at the round table "Big data, que fait-on de nos données?", organized at Museum of civilization, Québec, 3 Nov.; co-animator, workshop on big data in the "International science and society forum" for high school students, Québec, 4-6 Nov.; panelist at the round table "L'économie peut-elle être vraiment collaborative ET sociale et solidaire?", Ministry of the Economy, Paris, 13 Dec.

- **Paola Tubaro**, training on the digital society and its effects on labor and the economy, for union leaders (CGT, 27 Apr., CFDT, 23 Jun).

- **Paola Tubaro**, radio interviews and panels: Radio France Inter, Radio France Culture, ICI Radio Canada, Aligre.fm. Book reviews and mentions in the media: Le Monde, Rue89/L'Obs, Mashable/France24, Slate.fr, Journal International de Médecine, The Conversation.

- **io.datascience** notable presentations at: CNRS-Inria day *data4ist : exploration et analyse des sources de données pour la recherche et ses environnements* (May 2016) ; Futur en Seine 2016 (Juin 2016) ; DGESIP/MiPNES seminar *Normes et échanges de données : où en est-on ?* (September 2016).

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] R. ALLESIARDO.*Multi-armed bandits for non-stationary data streams*, Université Paris-Saclay ; LRI - CNRS, University Paris-Sud, October 2016, https://hal.inria.fr/tel-01420663.

[2] J. BENSADON.*Applications of Information Theory to Machine Learning*, Université Paris-Saclay, February 2016, https://tel.archives-ouvertes.fr/tel-01297163.

[3] M.-L. CAUWET.*Uncertainties in Optimization*, Université Paris Sud - Orsay, September 2016, https://hal.archives-ouvertes.fr/tel-01422274.

Articles in International Peer-Reviewed Journal

[4] A. AUGER, N. HANSEN.*Linear Convergence of Comparison-based Step-size Adaptive Randomized Search via Stability of Markov Chains*, in "SIAM Journal on Optimization", June 2016, https://hal.inria.fr/hal-00877160.

[5] C. BOUFENAR, M. BATOUCHE, M. SCHOENAUER.*An Artificial Immune System for Offline Isolated Handwritten Arabic Character Recognition*, in "Evolving Systems", 2016, p. 1-17 [*DOI :* 10.1007/S12530-016-9169-1], https://hal.inria.fr/hal-01394841.

[6] I. BRIGUI-CHTIOUI, P. CAILLOU.*Multidimensional Decision Model for Classifying Learners: the case of Massive Online Open Courses (MOOCs)* , in "Journal of Decision Systems", January 2017 [*DOI :* 10.1080/12460125.2017.1252235], https://hal.inria.fr/hal-01391011.

[7] A. D'ANGELO, L. RYAN, P. TUBARO.*Visualization in mixed-methods research on social networks*, in "Sociological Research Online", May 2016, vol. 21, n° 2, 15 [*DOI :* 10.5153/SRO.3996], https://hal.archives-ouvertes.fr/hal-01348848.

[8] A. DECELLE, F. RICCI-TERSENGHI.*Solving the inverse Ising problem by mean-field methods in a clustered phase space with many states*, in "Physical Review E ", July 2016, 6 pages, 5 figures [*DOI :* 10.1103/PHYSREVE.94.012112], https://hal.archives-ouvertes.fr/hal-01250824.

[9] A. DECELLE, F. RICCI-TERSENGHI, P. ZHANG.*Data quality for the inverse Ising problem*, in "Journal of Physics A: Mathematical and Theoretical", August 2016, vol. 49, n° 38 [*DOI :* 10.1088/1751-8113/49/38/384001], https://hal.archives-ouvertes.fr/hal-01250822.

[10] S. ESCALERA, V. ATHITSOS, I. GUYON.*Challenges in multimodal gesture recognition*, in "Journal of Machine Learning Research", 2016, vol. 17, n° 72, p. 1-54, https://hal.archives-ouvertes.fr/hal-01381158.

[11] C. FURTLEHNER, A. DECELLE.*Cycle-based Cluster Variational Method for Direct and Inverse Inference*, in "Journal of Statistical Physics", August 2016, vol. 164, n° 3, p. 531–574, https://hal.inria.fr/hal-01214155.

[12] J. GARCIA-RODRIGUEZ, I. GUYON, S. ESCALERA, A. PSARROU, A. LEWIS, M. CAZORLA.*Editorial: special issue on computational intelligence for vision and robotics*, in "Neural Computing and Applications", 2016, p. 1–2 [*DOI :* 10.1007/S00521-016-2330-8], https://hal.archives-ouvertes.fr/hal-01381150.

[13] E. MAGGIORI, Y. TARABALKA, G. CHARPIAT, P. ALLIEZ.*Convolutional Neural Networks for Large-Scale Remote Sensing Image Classification*, in "IEEE Transactions on Geoscience and Remote Sensing", September 2016, https://hal.inria.fr/hal-01369906.

[14] V. MARTIN, J.-M. LASGOUTTES, C. FURTLEHNER.*Latent binary MRF for online reconstruction of large scale systems*, in "Annals of Mathematics and Artificial Intelligence", 2016, vol. 77, n° 1, p. 123-154 [*DOI :* 10.1007/S10472-015-9470-X], https://hal.inria.fr/hal-01186220.

[15] M. MISIR, M. SEBAG.*Alors: An algorithm recommender system*, in "Artificial Intelligence", December 2016, Published on-line, https://hal.inria.fr/hal-01419874.

[16] P. TUBARO, L. RYAN, A. D'ANGELO.*The Visual Sociogram in Qualitative and Mixed-Methods Research*, in "Sociological Research Online", May 2016, vol. 21, n° 2, 1 [*DOI :* 10.5153/SRO.3864], https://hal.archives-ouvertes.fr/hal-01348491.

Invited Conferences

[17] C. DOERR, N. BREDECHE, E. ALBA, T. BARTZ-BEIELSTEIN, D. BROCKHOFF, B. DOERR, G. A. EIBEN, M. EPITROPAKIS, C. FONSECA, A. GUERREIRO, E. HAASDIJK.*Tutorials at PPSN 2016*, in "Parallel Problem Solving from Nature – PPSN XIV", Edinburgh, United Kingdom, Lecture Notes in Computer Science, September 2016, vol. 9921, p. 1012-1022, https://hal.archives-ouvertes.fr/hal-01363919.

International Conferences with Proceedings

[18] C. ADAM-BOURDARIOS, G. COWAN, C. GERMAIN, I. GUYON, B. KÉGL, D. ROUSSEAU.*How Machine Learning won the Higgs Boson Challenge*, in "European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning", Bruges, Belgium, April 2016, https://hal.inria.fr/hal-01423097.

[19] O. AIT ELHARA, A. AUGER, N. HANSEN.*Permuted Orthogonal Block-Diagonal Transformation Matrices for Large Scale Optimization Benchmarking*, in "GECCO 2016", Denver, United States, T. FRIEDRICH, F. NEUMANN (editors), Proceedings GECCO'16, ACM-Press, July 2016, p. 189-196 [*DOI :* 10.1145/2908812.2908937], https://hal.inria.fr/hal-01308566.

[20] Y. AKIMOTO, N. HANSEN.*Online Model Selection for Restricted Covariance Matrix Adaptation*, in "Parallel Problem Solving from Nature – PPSN XIV", Edinburgh, United Kingdom, J. HANDL, E. HART, P. LEWIS, M. LÓPEZ-IBÁÑEZ, G. OCHOA, B. PAECHTER (editors), LNCS, Springer Verlag, September 2016, p. 3-13, https://hal.inria.fr/hal-01333840.

[21] Y. AKIMOTO, N. HANSEN.*Projection-Based Restricted Covariance Matrix Adaptation for High Dimension*, in "Genetic and Evolutionary Computation Conference 2016", Denver, United States, T. FRIEDRICH, F.

NEUMANN (editors), Proc. ACM-GECCO'16, July 2016, p. 197-204 [*DOI :* 10.1145/2908812.2908863], https://hal.inria.fr/hal-01306551.

[22] S. ASTETE-MORALES, M.-L. CAUWET, O. TEYTAUD.*Analysis of Different Types of Regret in Continuous Noisy Optimization*, in "Genetic and Evolutionary Computation Conference 2016", Denver, United States, T. FRIEDRICH, F. NEUMANN (editors), Proc. ACM-GECCO'16, July 2016, p. 205-212, https://hal.archives-ouvertes.fr/hal-01347814.

[23] A. ATAMNA, A. AUGER, N. HANSEN.*Analysis of Linear Convergence of a (1 + 1)-ES with Augmented Lagrangian Constraint Handling*, in "Genetic and Evolutionary Computation Conference (GECCO)", Denver, United States, T. FRIEDRICH, F. NEUMANN (editors), Proc. ACM-GECCO'16, July 2016, p. 213-220, https://hal.inria.fr/hal-01318807.

[24] A. ATAMNA, A. AUGER, N. HANSEN.*Augmented Lagrangian Constraint Handling for CMA-ES—Case of a Single Linear Constraint*, in "Proceedings of the 14th International Conference on Parallel Problem Solving from Nature", Edinburgh, United Kingdom, September 2016, p. 181 - 191 [*DOI :* 10.1007/978-3-319-45823-6_17], https://hal.inria.fr/hal-01390386.

[25] A. AUGER, D. BROCKHOFF, N. HANSEN, D. TUŠAR, T. TUŠAR, T. WAGNER.*Benchmarking MATLAB's gamultiobj (NSGA-II) on the Bi-objective BBOB-2016 Test Suite*, in "Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion", Denver, CO, United States, July 2016, p. 1233-1239 [*DOI :* 10.1145/2908961.2931706], https://hal.inria.fr/hal-01435445.

[26] A. AUGER, D. BROCKHOFF, N. HANSEN, D. TUŠAR, T. TUŠAR, T. WAGNER.*Benchmarking RM-MEDA on the Bi-objective BBOB-2016 Test Suite*, in "Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion", Denver, CO, United States, July 2016, p. 1241-1247 [*DOI :* 10.1145/2908961.2931707], https://hal.inria.fr/hal-01435449.

[27] A. AUGER, D. BROCKHOFF, N. HANSEN, D. TUŠAR, T. TUŠAR, T. WAGNER.*Benchmarking the Pure Random Search on the Bi-objective BBOB-2016 Testbed*, in "Proceedings of the 2016 Genetic and Evolutionary Computation Conference Companion", Denver, CO, United States, July 2016, p. 1217-1223 [*DOI :* 10.1145/2908961.2931704], https://hal.inria.fr/hal-01435455.

[28] A. AUGER, D. BROCKHOFF, N. HANSEN, D. TUŠAR, T. TUŠAR, T. WAGNER.*The Impact of Search Volume on the Performance of RANDOMSEARCH on the Bi-objective BBOB-2016 Test Suite*, in "Proceedings of the 2016 Genetic and Evolutionary Computation Conference Companion", Denver, CO, United States, July 2016, p. 1257 - 1264 [*DOI :* 10.1145/2908961.2931709], https://hal.inria.fr/hal-01435453.

[29] A. AUGER, D. BROCKHOFF, N. HANSEN, D. TUŠAR, T. TUŠAR, T. WAGNER.*The Impact of Variation Operators on the Performance of SMS-EMOA on the Bi-objective BBOB-2016 Test Suite*, in "Proceedings of the 2016 Genetic and Evolutionary Computation Conference Companion (GECCO '16 Companion)", Denver, CO, United States, July 2016, p. 1225 - 1232 [*DOI :* 10.1145/2908961.2931705], https://hal.inria.fr/hal-01435456.

[30] N. BELKHIR, J. DRÉO, P. SAVÉANT, M. SCHOENAUER.*Feature Based Algorithm Configuration: A Case Study with Differential Evolution*, in "Parallel Problem Solving from Nature – PPSN XIV", Edinburgh, France, J. HANDL, E. HART, P. LEWIS, M. LÓPEZ-IBÁÑEZ, G. OCHOA, B. PAECHTER (editors), LNCS, Springer Verlag, September 2016, vol. 9921, p. 156-165 [*DOI :* 10.1007/978-3-319-45823-6_15], https://hal.inria.fr/hal-01359539.

[31] N. BELKHIR, J. DRÉO, P. SAVÉANT, M. SCHOENAUER.*Surrogate Assisted Feature Computation for Continuous Problems*, in "LION 10 Learning and Intelligent OptimizatioN Conference", Ischia, Italy, P. FESTA (editor), Proc. Learning and Intelligent OptimizatioN Conference, Springer Verlag, May 2016, forthcoming, https://hal.inria.fr/hal-01359543.

[32] M.-L. CAUWET, O. TEYTAUD.*Noisy Optimization: Fast Convergence Rates with Comparison-Based Algorithms*, in "Genetic and Evolutionary Computation Conference", Denver, United States, T. FRIEDRICH, F. NEUMANN (editors), Proc. ACM-GECCO'16, July 2016, p. 1101-1106, https://hal.archives-ouvertes.fr/hal-01306636.

[33] T. CAZENAVE, J. LIU, F. TEYTAUD, O. TEYTAUD.*Learning opening books in partially observable games: using random seeds in Phantom Go*, in "Computer intelligence and Games (CIG 2016)", Santorini, Greece, Computer intelligence and Games, September 2016, https://hal.inria.fr/hal-01413229.

[34] B. CHEN, S. ESCALERA, I. GUYON, V. PONCE-LOPEZ, N. SHAH, M. OLIU.*Overcoming Calibration Problems in Pattern Labeling with Pairwise Ratings: Application to Personality Traits*, in "European Conference on Computer Vision (ECCV 2016) Workshops", 2016, https://hal.archives-ouvertes.fr/hal-01381148.

[35] A. ERRAQABI, M. VALKO, A. CARPENTIER, O.-A. MAILLARD.*Pliable rejection sampling*, in "International Conference on Machine Learning", New York City, United States, June 2016, https://hal.inria.fr/hal-01322168.

[36] H. J. ESCALANTE, V. PONCE-LOPEZ, J. WAN, M. A. RIEGLER, B. CHEN, A. CLAPES, S. ESCALERA, I. GUYON, X. BARO, P. HALVORSEN, H. MÜLLER, M. LARSON.*ChaLearn Joint Contest on Multimedia Challenges Beyond Visual Analysis: An overview*, in "International Conference on Pattern Recognition (ICPR 2016) Workshops", 2016, https://hal.archives-ouvertes.fr/hal-01381144.

[37] S. ESCALERA, M. TORRES, B. MARTINEZ, X. BARO, H. J. ESCALANTE, I. GUYON, G. TZIMIROPOULOS, C. CORNEOU, M. OLIU, M. A. BAGHERI, M. VALSTAR.*ChaLearn Looking at People and Faces of the World: Face Analysis Workshop and Challenge 2016*, in "The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops", 2016, https://hal.archives-ouvertes.fr/hal-01381152.

[38] E. GALVÁN-LÓPEZ, E. MEZURA-MONTES, O. AIT ELHARA, M. SCHOENAUER.*On the Use of Semantics in Multi-objective Genetic Programming*, in "14th International Conference Parallel Problem Solving from Nature – PPSN XIV", Edinburgh, United Kingdom, J. HANDL, E. HART, P. R. LEWIS, M. LÓPEZ-IBÁÑEZ, G. OCHOA, B. PAECHTER (editors), LNCS - Lecture Notes in Computer Science, Springer Verlag, September 2016, vol. 9921, p. 353 - 363 [*DOI :* 10.1007/978-3-319-45823-6_33], https://hal.inria.fr/hal-01387632.

[39] F. GONARD, M. SCHOENAUER, M. SEBAG.*Algorithm Selector and Prescheduler in the ICON challenge*, in "International Conference on Metaheuristics and Nature Inspired Computing (META'2016)", Marrakech, Morocco, October 2016, https://hal.inria.fr/hal-01378745.

[40] G. GREFENSTETTE, L. MUCHEMI.*On the Place of Text Data in Lifelogs, and Text Analysis via Semantic Facets*, in "iConference 2016 SIE on Lifelogging", Philadelphia, United States, March 2016, https://hal.inria.fr/hal-01328473.

[41] G. GREFENSTETTE, K. RAFES.*Transforming Wikipedia into an Ontology-based Information Retrieval Search Engine for Local Experts using a Third-Party Taxonomy*, in "Joint Second Workshop on Language and Ontology & Terminology and Knowledge Structures (LangOnto2 + TermiKS) LO2TKS", Portoroz, Slovenia, May 2016, https://hal.inria.fr/hal-01224114.

[42] I. GUYON, I. CHAABANE, H. J. ESCALANTE, S. ESCALERA, D. JAJETIC, J. R. LLOYD, N. MACIA, B. RAY, L. ROMASZKO, M. SEBAG, A. STATNIKOV, S. TREGUER, E. VIEGAS.*A brief Review of the ChaLearn AutoML Challenge: Any-time Any-dataset Learning without Human Intervention*, in "International Conference in Machine Learning (ICML 2016) Workshops", New-York, United States, ICML 2016 AutoML Workshop, JMLR, 2016, vol. 1, p. 1-8, https://hal.archives-ouvertes.fr/hal-01381145.

[43] A. O. KAZAKÇI, C. MEHDI, B. KÉGL.*Digits that are not: Generating new types through deep neural nets*, in "International Conference on Computational Creativity", Paris, France, June 2016, https://hal.archives-ouvertes.fr/hal-01427556.

[44] O. KRAUSE, T. GLASMACHERS, N. HANSEN, C. IGEL.*Unbounded Population MO-CMA-ES for the Bi-Objective BBOB Test Suite*, in "GECCO'16 Companion of Proceedings of the 2016 Genetic and Evolutionary Computation Conference", Denver, United States, ACM, July 2016, p. 1177-1184 [*DOI :* 10.1145/2908961.2931699], https://hal.inria.fr/hal-01381653.

[45] E. MAGGIORI, Y. TARABALKA, G. CHARPIAT, P. ALLIEZ.*Fully Convolutional Neural Networks For Remote Sensing Image Classification*, in "IEEE International Geoscience and Remote Sensing Symposium", Beijing, China, Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS), IEEE, July 2016, p. 5071-5074, https://hal.inria.fr/hal-01350706.

[46] L. MARTI, A. FANSI-TCHANGO, L. NAVARRO, M. SCHOENAUER.*Anomaly Detection with the Voronoi Diagram Evolutionary Algorithm*, in "Parallel Problem Solving from Nature – PPSN XIV", Edinburgh, United Kingdom, J. HANDL, E. HART, P. LEWIS, M. LÓPEZ-IBÁÑEZ, G. OCHOA, B. PAECHTER (editors), LNCS, Springer Verlag, September 2016, vol. 9921, p. 697-706 [*DOI :* 10.1007/978-3-319-45823-6_65], https://hal.inria.fr/hal-01387621.

[47] V. PONCE-LOPEZ, B. CHEN, M. OLIU, C. CORNEARU, A. CLAPES, I. GUYON, X. BARO, H. J. ESCALANTE, S. ESCALERA.*ChaLearn LAP 2016: First Round Challenge on First Impressions - Dataset and Results*, in "European Conference on Computer Vision (ECCV 2016) Workshops", Amsterdam, Netherlands, 2016 [*DOI :* 10.1007/978-3-319-49409-8_32], https://hal.archives-ouvertes.fr/hal-01381149.

[48] F. POPESCU, S. AYACHE, S. ESCALERA, X. BARÓ SOLÉ, C. CAPPONI, P. PANCIATICI, I. GUYON.*From geospatial observations of ocean currents to causal predictors of spatio-economic activity using computer vision and machine learning*, in "EGU General Assembly Conference Abstracts", 2016, https://hal.archives-ouvertes.fr/hal-01381147.

[49] T. SCHMITT, P. CAILLOU, M. SEBAG.*Matching Jobs and Resumes: a Deep Collaborative Filtering Task*, in "GCAI 2016 - 2nd Global Conference on Artificial Intelligence", Berlin, Germany, GCAI 2016. 2nd Global Conference on Artificial Intelligence, September 2016, vol. 41, https://hal.inria.fr/hal-01378589.

[50] P. TAILLANDIER, M. BOURGAIS, P. CAILLOU, C. ADAM, B. GAUDOU.*A BDI agent architecture for the GAMA modeling and simulation platform*, in "MABS 2016 Multi-Agent-Based Simulation", Singapore, Singapore, May 2016, https://hal.inria.fr/hal-01391002.

[51] F. TEYTAUD, O. TEYTAUD.*QR mutations improve many evolution strategies -a lot on highly multimodal problems*, in "ACM-GECCO'16", Denver, United States, T. FRIEDRICH, F. NEUMANN (editors), Poster in GECCO'16 Companion, July 2016, p. 35-36 [*DOI :* 10.1145/1235], https://hal.inria.fr/hal-01406727.

[52] J. WAN, Y. ZHAO, S. ZHOU, I. GUYON, S. ESCALERA, S. Z. LI.*ChaLearn Looking at People RGB-D Isolated and Continuous Datasets for Gesture Recognition*, in "The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops", 2016, https://hal.archives-ouvertes.fr/hal-01381151.

Conferences without Proceedings

[53] M.-L. CAUWET, O. TEYTAUD.*Multivariate bias reduction in capacity expansion planning*, in "19th Power Systems Computation Conference", Gênes, Italy, June 2016, https://hal.archives-ouvertes.fr/hal-01306643.

[54] M. CHERTI, B. KÉGL, A. O. KAZAKÇI.*Out-of-class novelty generation: an experimental foundation **, in "Neural Information Processing Systems", Barcelona, Spain, December 2016, https://hal.archives-ouvertes.fr/hal-01427570.

[55] G. GREFENSTETTE, L. MUCHEMI.*Determining the Characteristic Vocabulary for a Specialized Dictionary using Word2vec and a Directed Crawler*, in "GLOBALEX 2016: Lexicographic Resources for Human Language Technology", Portoroz, Slovenia, May 2016, https://hal.inria.fr/hal-01323591.

[56] D. ROUSSEAU, P. CALAFIURA, C. GERMAIN, V. INNOCENTE, R. CENCI, M. KAGAN, I. GUYON, D. CLARK, S. FARREL, R. CARNEY, A. SALZBURGER, D. COSTANZO, M. ELSING, T. GOLLING, T. TONG, J.-R. V. VLIMANT.*TrackML: a LHC Tracking Machine Learning Challenge*, in "International Conference on Computing in High Energy and Nuclear Physics", San Francisco, United States, October 2016, https://hal.inria.fr/hal-01422939.

Scientific Books (or Scientific Book chapters)

[57] A. A. CASILLI, P. TUBARO.*Le phénomène "pro ana": Troubles alimentaires et réseaux sociaux*, Presses des Mines, October 2016, https://hal.archives-ouvertes.fr/hal-01404890.

[58] P. TUBARO.*Formalization and mathematical modelling*, in "History of Economic Analysis", H. K. GILBERT FACCARELLO (editor), Edward Elgar, July 2016, vol. III, p. 208–221 [*DOI :* 10.4337/9781785365065.00022], https://hal.archives-ouvertes.fr/hal-01350533.

Books or Proceedings Editing

[59] S. BONNEVAY, P. LEGRAND, N. MONMARCHÉ, E. LUTTON, M. SCHOENAUER (editors). *Artificial Evolution 2015*, LNCS - Lecture Notes in Computer Science, Springer, Lyon, France, 2016, vol. 9554 [*DOI :* 10.1007/978-3-319-31471-6], https://hal.inria.fr/hal-01389072.

[60] A. D'ANGELO, L. RYAN, P. TUBARO (editors). *Peer Reviewed Special Section: Visualization in Mixed-Methods Research on Social NetworksSpecial section of Sociological Research Online, Guest edited by Alessio D'Angelo, Louise Ryan and Paola Tubaro*, Sociological Research Online (special section), Sage, May 2016, vol. 21, n° 2, https://hal.archives-ouvertes.fr/hal-01348492.

Other Publications

[61] R. ALLESIARDO, R. FÉRAUD, O.-A. MAILLARD.*Random Shuffling and Resets for the Non-stationary Stochastic Bandit Problem*, November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01400320.

[62] D. BROCKHOFF, T. TUSAR, D. TUSAR, T. WAGNER, N. HANSEN, A. AUGER.*Biobjective Performance Assessment with the COCO Platform*, May 2016, ArXiv e-prints, arXiv:1605.01746, https://hal.inria.fr/hal-01315317.

[63] A. GOPALAN, O.-A. MAILLARD, M. ZAKI.*Low-rank Bandits with Latent Mixtures*, November 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01400318.

[64] N. HANSEN, A. AUGER, D. BROCKHOFF, D. TUSAR, T. TUSAR.*COCO: Performance Assessment*, May 2016, ArXiv e-prints, arXiv:1605.03560, https://hal.inria.fr/hal-01315318.

[65] N. HANSEN, A. AUGER, O. MERSMANN, T. TUSAR, D. BROCKHOFF.*COCO: A Platform for Comparing Continuous Optimizers in a Black-Box Setting*, July 2016, ArXiv e-prints, arXiv:1603.08785, https://hal.inria.fr/hal-01294124.

[66] N. HANSEN, T. TUSAR, O. MERSMANN, A. AUGER, D. BROCKHOFF.*COCO: The Experimental Procedure*, May 2016, ArXiv e-prints, arXiv:1603.08776, https://hal.inria.fr/hal-01294167.

[67] E. MAGGIORI, G. CHARPIAT, Y. TARABALKA, P. ALLIEZ.*Learning Iterative Processes with Recurrent Neural Networks to Correct Satellite Image Classification Maps*, October 2016, working paper or preprint, https://hal.inria.fr/hal-01388551.

[68] E. MAGGIORI, Y. TARABALKA, G. CHARPIAT, P. ALLIEZ.*High-Resolution Semantic Labeling with Convolutional Neural Networks*, November 2016, working paper or preprint, https://hal.inria.fr/hal-01393279.

[69] O.-A. MAILLARD.*Self-normalization techniques for streaming confident regression*, May 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01349727.

[70] L. MARTI, A. FANSI-TCHANGO, L. NAVARRO, M. SCHOENAUER. , T. FRIEDRICH, F. NEUMANN (editors)*VorAIS: A Multi-Objective Voronoi Diagram-based Artificial Immune System*, Poster at GECCO 2016 (Companion), July 2016, p. 11 - 12, Genetic and Evolutionary Computation Conference, GECCO 2016, Poster [*DOI :* 10.1145/2908961.2909027], https://hal.inria.fr/hal-01400263.

[71] L. MUCHEMI, G. GREFENSTETTE.*Word Embedding and Statistical Based Methods for Rapid Induction of Multiple Taxonomies*, June 2016, working paper or preprint, https://hal.inria.fr/hal-01334236.

[72] D. ROUSSEAU, C. GERMAIN, V. INNOCENTE, R. CENCI, M. KAGAN, I. GUYON, D. CLARK, S. FARREL, R. CARNEY, A. SALZBURGER, D. COSTANZO, M. ELSING, T. GOLLING, T. TONG, J.-R. V. VLIMANT, S. WENZEL.*The Tracking Machine Learning Challenge*, December 2016, NIPS workshop: Challenges in Machine Learning (CiML), Poster, https://hal.inria.fr/hal-01422941.

[73] S. TFAILI, D. BUI THI, K. RAFES, A. TFAYLI, A. BAILLET-GUFFROY, C. GERMAIN, P. CHAMINADE.*Data acquisition for analytical platforms: Automating scientific workflows and building an open database platform for chemical anlysis metadata*, January 2016, Chimiométrie XVII, Poster, https://hal.inria.fr/hal-01423371.

[74] T. TUSAR, D. BROCKHOFF, N. HANSEN, A. AUGER.*COCO: The Bi-objective Black Box Optimization Benchmarking (bbob-biobj) Test Suite*, May 2016, ArXiv e-prints, arXiv:1604.00359, https://hal.inria.fr/hal-01296987.

Project-Team TOCCATA

Formally Verified Programs, Certified Tools, Numerical Computations

IN COLLABORATION WITH: Laboratoire de recherche en informatique (LRI)

IN PARTNERSHIP WITH:
Université Paris-Sud (Paris 11)

RESEARCH CENTER
Saclay - Île-de-France

THEME
Proofs and Verification

Table of contents

Project-Team TOCCATA

Creation of the Team: 2012 September 01, updated into Project-Team: 2014 July 01

Keywords:

Computer Science and Digital Science:

2.1.3. - Functional programming
2.1.6. - Concurrent programming
2.1.11. - Proof languages
2.4.2. - Model-checking
2.4.3. - Proofs
7.4. - Logic in Computer Science
7.12. - Computer arithmetic

Other Research Topics and Application Domains:

5.2.2. - Railway
5.2.3. - Aviation
5.2.4. - Aerospace

1. Members

Research Scientists

Claude Marché [Inria, Senior Researcher, Team leader, HDR]
Sylvie Boldo [Inria, Researcher, Team co-leader, HDR]
Arthur Charguéraud [Inria, Researcher, until Sep. 2016]
Jean-Christophe Filliâtre [CNRS, Senior Researcher, HDR]
Guillaume Melquiond [Inria, Researcher]

Faculty Members

Sylvain Conchon [Univ. Paris-Sud, Professor, HDR]
Andrei Paskevich [Univ. Paris-Sud, Associate Professor]
Christine Paulin-Mohring [Univ. Paris-sud, Professor, until Jun. 2016, HDR]

Technical Staff

Clément Fumex [Inria]
Sylvain Dailler [Inria, from Jun. 2016]

PhD Students

Cláudio Belo Da Silva Lourenço [Universidade do Minho, Portugal, visiting from Jun. to Dec. 2016]
Ran Chen [Institute of Software, Chinese Academy of Sciences, Beijing, China, visiting from May 2016 to Feb. 2017]
Martin Clochard [Univ. Paris-Sud (ENS grant)]
Albin Coquereau [ENSTA]
David Declerck [Univ. Paris-Sud]
Florian Faissole [Inria, from Oct 2016]
Léon Gondelman [Inria]
Mario José Parreira Pereira [Grant Portuguese government]
Mattias Roux [Univ. Paris-Sud]

Administrative Assistants

Régine Bricquet [Inria, until Jan 2016]

Katia Evrat [Inria, from Feb 2016]
Others
Thibaut Balabonski [Univ. Paris-Sud, associate member]
Chantal Keller [Univ. Paris-Sud, associate member]
Raphaël Rieu-Helft [Visitor, ENS, from Sep 2016]
Xavier Urbain [ENSIEE & LRI, associate member, until Aug 2016]

2. Overall Objectives

2.1. Overall Objectives

The general objective of the Toccata project is to promote formal specification and computer-assisted proof in the development of software that requires high assurance in terms of safety and correctness with respect to the intended behavior of the software.

2.1.1. *Context*

The importance of software in critical systems increased a lot in the last decade. Critical software appears in various application domains like transportation (e.g., aviation, railway), communication (e.g., smartphones), banking, etc. The number of tasks performed by software is quickly increasing, together with the number of lines of code involved. Given the need of high assurance of safety in the functional behavior of such applications, the need for automated (i.e., computer-assisted) methods and techniques to bring guarantee of safety became a major challenge. In the past and at present, the most widely used approach to check safety of software is to apply heavy test campaigns. These campaigns take a large part of the costs of software development, yet they cannot ensure that all the bugs are caught.

Generally speaking, software verification approaches pursue three goals: (1) verification should be sound, in the sense that no bugs should be missed, (2) verification should not produce false alarms, or as few as possible (3) it should be as automated as possible. Reaching all three goals at the same time is a challenge. A large class of approaches emphasizes goals (2) and (3): testing, run-time verification, symbolic execution, model checking, etc. Static analysis, such as abstract interpretation, emphasizes goals (1) and (3). Deductive verification emphasizes (1) and (2). The Toccata project is mainly interested in exploring the deductive verification approach, although we also consider the others in some cases.

In the past decade, there has been significant progress made in the domain of deductive program verification. They are emphasized by some success stories of application of these techniques on industrial-scale software. For example, the *Atelier B* system was used to develop part of the embedded software of the Paris metro line 14 [45] and other railroad-related systems; a formally proved C compiler was developed using the Coq proof assistant [100]; Microsoft's hypervisor for highly secure virtualization was verified using VCC [79] and the Z3 prover [119]; the L4-verified project developed a formally verified micro-kernel with high security guarantees, using analysis tools on top of the Isabelle/HOL proof assistant [96]. Another sign of recent progress is the emergence of deductive verification competitions (e.g., VerifyThis [3], VScomp [90]).

Finally, recent trends in the industrial practice for development of critical software is to require more and more guarantees of safety, e.g., the upcoming DO-178C standard for developing avionics software adds to the former DO-178B the use of formal models and formal methods. It also emphasizes the need for certification of the analysis tools involved in the process.

2.1.2. *Deductive verification*

There are two main families of approaches for deductive verification. Methods in the first family build on top of mathematical proof assistants (e.g., Coq, Isabelle) in which both the model and the program are encoded; the proof that the program meets its specification is typically conducted in an interactive way using the underlying proof construction engine. Methods from the second family proceed by the design of standalone tools taking as input a program in a particular programming language (e.g., C, Java) specified with a dedicated annotation

language (e.g., ACSL [44], JML [64]) and automatically producing a set of mathematical formulas (the *verification conditions*) which are typically proved using automatic provers (e.g., Z3, Alt-Ergo [46], CVC3 [43], CVC4).

The first family of approaches usually offers a higher level of assurance than the second, but also demands more work to perform the proofs (because of their interactive nature) and makes them less easy to adopt by industry. Moreover, they do not allow to directly analyze a program written in a mainstream programming language like Java or C. The second kind of approaches has benefited in the past years from the tremendous progress made in SAT and SMT solving techniques, allowing more impact on industrial practices, but suffers from a lower level of trust: in all parts of the proof chain (the model of the input programming language, the VC generator, the back-end automatic prover), potential errors may appear, compromising the guarantee offered. Moreover, while these approaches are applied to mainstream languages, they usually support only a subset of their features.

3. Research Program

3.1. Introduction

In the former ProVal project, we have been working on the design of methods and tools for deductive verification of programs. One of our original skills was the ability to conduct proofs by using automatic provers and proof assistants at the same time, depending on the difficulty of the program, and specifically the difficulty of each particular verification condition. We thus believe that we are in a good position to propose a bridge between the two families of approaches of deductive verification presented above. Establishing this bridge is one of the goals of the Toccata project: we want to provide methods and tools for deductive program verification that can offer both a high amount of proof automation and a high guarantee of validity. Toward this objective, a new axis of research was proposed: the development of *certified* analysis tools that are themselves formally proved correct.

The reader should be aware that the word "certified" in this scientific programme means "verified by a formal specification and a formal proof that the program meets this specification". This differs from the standard meaning of "certified" in an industrial context where it means a conformance to some rigorous process and/or norm. We believe this is the right term to use, as it was used for the *Certified Compiler* project [100], the new conference series *Certified Programs and Proofs*, and more generally the important topics of *proof certificates*.

In industrial applications, numerical calculations are very common (e.g. control software in transportation). Typically they involve floating-point numbers. Some of the members of Toccata have an internationally recognized expertise on deductive program verification involving floating-point computations. Our past work includes a new approach for proving behavioral properties of numerical C programs using Frama-C/Jessie [41], various examples of applications of that approach [61], the use of the Gappa solver for proving numerical algorithms [118], an approach to take architectures and compilers into account when dealing with floating-point programs [62], [111]. We also contributed to the Handbook of Floating-Point Arithmetic [110]. A representative case study is the analysis and the proof of both the method error and the rounding error of a numerical analysis program solving the one-dimension acoustic wave equation [4] [54]. Our experience led us to a conclusion that verification of numerical programs can benefit a lot from combining automatic and interactive theorem proving [56], [61]. Certification of numerical programs is the other main axis of Toccata.

Our scientific programme in structured into four objectives:

1. deductive program verification;
2. automated reasoning;
3. formalization and certification of languages, tools and systems;
4. proof of numerical programs.

We detail these objectives below.
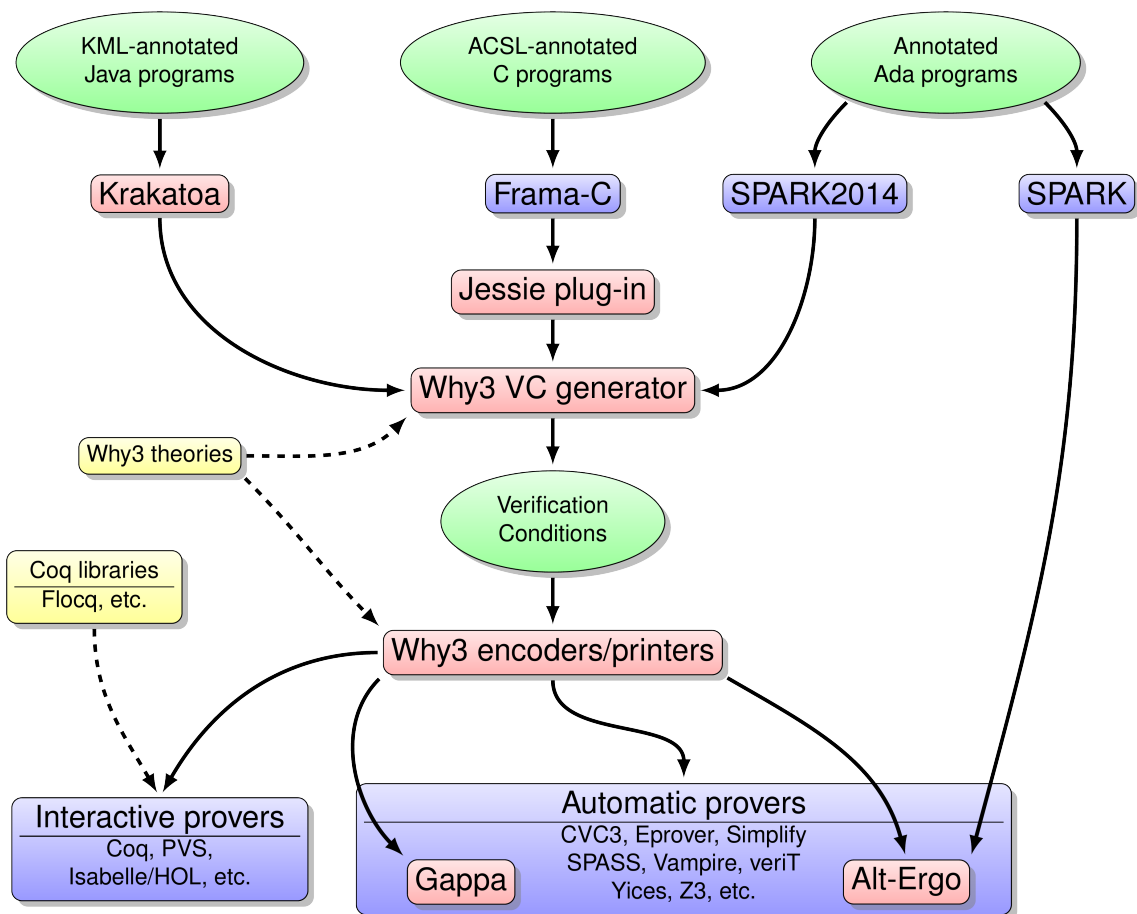
Figure 1. The Why3 ecosystem

3.2. Deductive Program Verification

Permanent researchers: A. Charguéraud, S. Conchon, J.-C. Filliâtre, C. Marché, G. Melquiond, A. Paskevich

3.2.1. *The Why3 Ecosystem*

This ecosystem is central in our work; it is displayed on Figure 1. The boxes in red background correspond to the tools we develop in the Toccata team.

- The initial design of Why3 was presented in 2012 [49], [89]. In the past years, the main improvements concern the specification language (such as support for higher-order logic functions [67]) and the support for provers. Several new interactive provers are now supported: PVS 6 (used at NASA), Isabelle2014 (planned to be used in the context of Ada program via Spark), and Mathematica. We also added support for new automated provers: CVC4, Metitarski, Metis, Beagle, Princess, and Yices2. More technical improvements are the design of a Coq tactic to call provers via Why3 from Coq, and the design of a proof session mechanism [48]. Why3 was presented during several invited talks [88], [87], [84], [85].

- At the level of the C front-end of Why3 (via Frama-C), we have proposed an approach to add a notion of refinement on C programs [117], and an approach to reason about pointer programs with a standard logic, via *separation predicates* [47]

- The Ada front-end of Why3 has mainly been developed during the past three years, leading to the release of SPARK2014 [95] (http://www.spark-2014.org/)

- In collaboration with J. Almeida, M. Barbosa, J. Pinto, and B. Vieira (University do Minho, Braga, Portugal), J.-C. Filliâtre has developed a method for certifying programs involving cryptographic methods. It uses Why as an intermediate language [40].

- With M. Pereira and S. Melo de Sousa (Universidade da Beira Interior, Covilhã, Portugal), J.-C. Filliâtre has developed an environment for proving ARM assembly code. It uses Why3 as an intermediate VC generator. It was presented at the Inforum conference [114] (best student paper).

3.2.2. *Concurrent Programming*

- S. Conchon and A. Mebsout, in collaboration with F. Zaïdi (VALS team, LRI), A. Goel and S. Krstić (Strategic Cad Labs, INTEL) have proposed a new model-checking approach for verifying safety properties of array-based systems. This is a syntactically restricted class of parametrized transition systems with states represented as arrays indexed by an arbitrary number of processes. Cache coherence protocols and mutual exclusion algorithms are typical examples of such systems. It was first presented at CAV 2012 [7] and detailed further [77]. It was applied to the verification of programs with fences [73]. The core algorithm has been extended with a mechanism for inferring invariants. This new algorithm, called BRAB, is able to automatically infer invariants strong enough to prove industrial cache coherence protocols. BRAB computes over-approximations of backward reachable states that are checked to be unreachable in a finite instance of the system. These approximations (candidate invariants) are then model-checked together with the original safety properties. Completeness of the approach is ensured by a mechanism for backtracking on spurious traces introduced by too coarse approximations [74], [106].

- In the context of the ERC DeepSea project [0], A. Charguéraud and his co-authors have developed a unifying semantics for various different paradigms of parallel computing (fork-join, async-finish, and futures), and published a conference paper describing this work [16]. Besides, A. Charguéraud and his co-authors have polished their previous work on granularity control for parallel algorithms using user-provided complexity functions, and produced a journal article [11].

---

[0]Arthur Charguéraud is involved 40% of his time in the ERC DeepSea project, which is hosted at Inria Paris Rocquencourt (team Gallium).

3.2.3. Case Studies

- To provide an easy access to the case studies that we develop using Why3 and its front-ends, we have published a *gallery of verified programs* on our web page http://toccata.lri.fr/gallery/. Part of these examples are the solutions to the competitions VerifyThis 2011 [63], VerifyThis 2012 [3], and the competition VScomp 2011 [90].

- Other case studies that led to publications are the design of a library of data-structures based on AVLs [66], and the verification a two-lines C program (solving the N-queens puzzle) using Why3 [86].

- A. Charguéraud, with F. Pottier (Inria Paris), extended their formalization of the correctness and asymptotic complexity of the classic Union Find data structure, which features the bound expressed in terms of the inverse Ackermann function [39]. The proof, conducted using CFML extended with time credits, was refined using a slightly more complex potential function, allowing to derive a simpler and richer interface for the data structure. A journal article is in preparation.

For other case studies, see also sections of numerical programs and formalization of languages and tools.

3.2.4. Project-team Positioning

Several research groups in the world develop their own approaches, techniques, and tools for deductive verification. With respect to all these related approaches and tools, our originality is our will to use more sophisticated specification languages (with inductive definitions, higher-order features and such) and the ability to use a large set of various theorem provers, including the use of interactive theorem proving to deal with complex functional properties.

- The RiSE team [0] at Microsoft Research Redmond, USA, partly in collaboration with team "programming methodology" team [0] at ETH Zurich develop tools that are closely related to ours: Boogie and Dafny are direct competitors of Why3, VCC is a direct competitor of Frama-C/Jessie.

- The KeY project [0] (several teams, mainly at Karlsruhe and Darmstadt, Germany, and Göteborg, Sweden) develops the KeY tool for Java program verification [38], based on dynamic logic, and has several industrial users. They use a specific modal logic (dynamic logic) for modeling programs, whereas we use standard logic, so as to be able to use off-the-shelf automated provers.

- The "software engineering" group at Augsburg, Germany, develops the KIV system [0], which was created more than 20 years ago (1992) and is still well maintained and efficient. It provides a semi-interactive proof environment based on algebraic-style specifications, and is able to deal with several kinds of imperative style programs. They have a significant industrial impact.

- The VeriFast system [0] aims at verifying C programs specified in Separation Logic. It is developed at the Katholic University at Leuven, Belgium. We do not usually use separation logic (so as to use off-the-shelf provers) but alternative approaches (e.g. static memory separation analysis).

- The Mobius Program Verification Environment [0] is a joint effort for the verification of Java source annotated with JML, combining static analysis and runtime checking. The tool ESC/Java2 [0] is a VC generator similar to Krakatoa, that builds on top of Boogie. It is developed by a community leaded by University of Copenhagen, Denmark. Again, our specificity with respect to them is the consideration of more complex specification languages and interactive theorem proving.

- The Lab for Automated Reasoning and Analysis [0] at EPFL, develop methods and tools for verification of Java (Jahob) and Scala (Leon) programs. They share with us the will and the ability to use several provers at the same time.

- The TLA environment [0], developed by Microsoft Research and the Inria team Veridis, aims at

---

[0] http://research.microsoft.com/en-us/groups/rise/default.aspx
[0] http://www.pm.inf.ethz.ch/
[0] http://www.key-project.org/
[0] http://www.isse.uni-augsburg.de/en/software/kiv/
[0] http://people.cs.kuleuven.be/~bart.jacobs/verifast/
[0] http://kindsoftware.com/products/opensource/Mobius/
[0] http://kindsoftware.com/products/opensource/ESCJava2/
[0] http://lara.epfl.ch/w/

the verification of concurrent programs using mathematical specifications, model checking, and interactive or automated theorem proving.

- The F* project [0], developed by Microsoft Research and the Inria Prosecco team, aims at providing a rich environment for developing programs and proving them.

The KeY and KIV environments mentioned above are partly based on interactive theorem provers. There are other approaches on top of general-purpose proof assistants for proving programs that are not purely functional:

- The Ynot project [0] is a Coq library for writing imperative programs specified in separation logic. It was developed at Harvard University, until the end of the project in 2010. Ynot had similar goals as CFML, although Ynot requires programs to be written in monadic style inside Coq, whereas CFML applies directly on programs written in OCaml syntax, translating them into logical formulae.
- Front-ends to Isabelle were developed to deal with simple sequential imperative programs [116] or C programs [113]. The L4-verified project [96] is built on top of Isabelle.

3.3. Automated Reasoning

Permanent researchers: S. Conchon, G. Melquiond, A. Paskevich

3.3.1. *Generalities on Automated Reasoning*

- J. C. Blanchette and A. Paskevich have 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) [1]. This extension, named TFF1, has been incorporated in the TPTP standard.
- S. Conchon defended his *habilitation à diriger des recherches* in December 2012. The memoir [70] provides a useful survey of the scientific work of the past 10 years, around the SMT solving techniques, that led to the tools Alt-Ergo and Cubicle as they are nowadays.

3.3.2. *Quantifiers and Triggers*

- C. Dross, J. Kanig, S. Conchon, and A. Paskevich have proposed a generic framework for adding a decision procedure for a theory or a combination of theories to an SMT prover. This mechanism is based on the notion of instantiation patterns, or *triggers*, which restrict instantiation of universal premises and can effectively prevent a combinatorial explosion. A user provides an axiomatization with triggers, along with a proof of completeness and termination in the proposed framework, and obtains in return a sound, complete and terminating solver for his theory. A prototype implementation was realized on top of Alt-Ergo. As a case study, a feature-rich axiomatization of doubly-linked lists was proved complete and terminating [82]. C. Dross defended her PhD thesis in April 2014 [83]. The main results of the thesis are: (1) a formal semantics of the notion of *triggers* typically used to control quantifier instantiation in SMT solvers, (2) a general setting to show how a first-order axiomatization with triggers can be proved correct, complete, and terminating, and (3) an extended DPLL(T) algorithm to integrate a first-order axiomatization with triggers as a decision procedure for the theory it defines. Significant case studies were conducted on examples coming from SPARK programs, and on the benchmarks on B set theory constructed within the BWare project.

3.3.3. *Reasoning Modulo Theories*

- S. Conchon, É. Contejean and M. Iguernelala have presented a modular extension of ground AC-completion for deciding formulas in the combination of the theory of equality with user-defined AC symbols, uninterpreted symbols and an arbitrary signature-disjoint Shostak theory X [72]. This work extends the results presented in [71] by showing that a simple preprocessing step allows to get rid of a full AC-compatible reduction ordering, and to simply use a partial multiset extension of a *non-necessarily AC-compatible* ordering.

- S. Conchon, M. Iguernelala, and A. Mebsout have designed a collaborative framework for reasoning modulo simple properties of non-linear arithmetic [76]. This framework has been implemented in the Alt-Ergo SMT solver.
- S. Conchon, G. Melquiond and C. Roux have described a dedicated procedure for a theory of floating-point numbers which allows reasoning on approximation errors. This procedure is based on the approach of the Gappa tool: it performs saturation of consequences of the axioms, in order to refine bounds on expressions. In addition to the original approach, bounds are further refined by a constraint solver for linear arithmetic [78]. This procedure has been implemented in Alt-Ergo.
- In collaboration with A. Mahboubi (Inria project-team Typical), and G. Melquiond, the group involved in the development of Alt-Ergo have implemented and proved the correctness of a novel decision procedure for quantifier-free linear integer arithmetic [2]. This algorithm tries to bridge the gap between projection and branching/cutting methods: it interleaves an exhaustive search for a model with bounds inference. These bounds are computed provided an oracle capable of finding constant positive linear combinations of affine forms. An efficient oracle based on the Simplex procedure has been designed. This algorithm is proved sound, complete, and terminating and is implemented in Alt-Ergo.
- Most of the results above are detailed in M. Iguernelala's PhD thesis [93].

3.3.4. Applications

- We have been quite successful in the application of Alt-Ergo to industrial development: qualification by Airbus France, integration of Alt-Ergo into the Spark Pro toolset.
- In the context of the BWare project, aiming at using Why3 and Alt-Ergo for discharging proof obligations generated by Atelier B, we made progress into several directions. The method of translation of B proof obligations into Why3 goals was first presented at ABZ'2012 [109]. Then, new drivers have been designed for Why3, in order to use new back-end provers Zenon modulo and iProver modulo. A notion of rewrite rule was introduced into Why3, and a transformation for simplifying goals before sending them to back-end provers was designed. Intermediate results obtained so far in the project were presented both at the French conference AFADL [81] and at ABZ'2014 [80].

  On the side of Alt-Ergo, recent developments have been made to efficiently discharge proof obligations generated by Atelier B. This includes a new plugin architecture to facilitate experiments with different SAT engines, new heuristics to handle quantified formulas, and important modifications in its internal data structures to boost performances of core decision procedures. Benchmarks realized on more than 10,000 proof obligations generated from industrial B projects show significant improvements [75].

- Hybrid automatons interleave continuous behaviors (described by differential equations) with discrete transitions. D. Ishii and G. Melquiond have worked on an automated procedure for verifying safety properties (that is, global invariants) of such systems [94].

3.3.5. Project-team Positioning

Automated Theorem Proving is a large community, but several sub-groups can be identified:

- The SMT-LIB community gathers people interested in reasoning modulo theories. In this community, only a minority of participants are interested in supporting first-order quantifiers at the same time as theories. SMT solvers that support quantifiers are Z3 (Microsoft Research Redmond, USA), CVC3 and its successor CVC4 [0].
- The TPTP community gathers people interested in first-order theorem proving.
- Other Inria teams develop provers: veriT by team Veridis, and Psyche by team Parsifal.
- Other groups develop provers dedicated to very specific cases, such as Metitarski [0] at Cambridge, UK, which aims at proving formulas on real numbers, in particular involving special functions such as log or exp. The goal is somewhat similar to our CoqInterval library, *cf* objective 4.

---

[0] http://cvc4.cs.nyu.edu/web/
[0] http://www.cl.cam.ac.uk/~lp15/papers/Arith/

It should be noticed that a large number of provers mentioned above are connected to Why3 as back-ends.

3.4. Formalization and Certification of Languages, Tools and Systems

Permanent researchers: S. Boldo, A. Charguéraud, C. Marché, G. Melquiond, C. Paulin

3.4.1. Real Numbers, Real Analysis, Probabilities

- S. Boldo, C. Lelay, and G. Melquiond have worked on the Coquelicot library, designed to be a user-friendly Coq library about real analysis [59], [60]. An easier way of writing formulas and theorem statements is achieved by relying on total functions in place of dependent types for limits, derivatives, integrals, power series, and so on. To help with the proof process, the library comes with a comprehensive set of theorems and some automation. We have exercised the library on several use cases: on an exam at university entry level [98], for the definitions and properties of Bessel functions [97], and for the solution of the one-dimensional wave equation [99]. We have also conducted a survey on the formalization of real arithmetic and real analysis in various proof systems [12].

- Watermarking techniques are used to help identify copies of publicly released information. They consist in applying a slight and secret modification to the data before its release, in a way that should remain recognizable even in (reasonably) modified copies of the data. Using the Coq ALEA library, which formalizes probability theory and probabilistic programs, D. Baelde together with P. Courtieu, D. Gross-Amblard from Rennes and C. Paulin have established new results about the robustness of watermarking schemes against arbitrary attackers [42]. The technique for proving robustness is adapted from methods commonly used for cryptographic protocols and our work illustrates the strengths and particularities of the ALEA style of reasoning about probabilistic programs.

3.4.2. Formalization of Languages, Semantics

- P. Herms, together with C. Marché and B. Monate (CEA List), has developed a certified VC generator, using Coq. The program for VC calculus and its specifications are both written in Coq, but the code is crafted so that it can be extracted automatically into a stand-alone executable. It is also designed in a way that allows the use of arbitrary first-order theorem provers to discharge the generated obligations [92]. On top of this generic VC generator, P. Herms developed a certified VC generator for C source code annotated using ACSL. This work is the main result of his PhD thesis [91].

- A. Tafat and C. Marché have developed a certified VC generator using Why3 [102], [103]. 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 or 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 significantly higher amount of proof automation compared to Coq.

- A. Charguéraud, together with Alan Schmitt (Inria Rennes) and Thomas Wood (Imperial College), has developed an interactive debugger for JavaScript. The interface, accessible as a webpage in a browser, allows to execute a given JavaScript program, following step by step the formal specification of JavaScript developed in prior work on *JsCert* [50]. Concretely, the tool acts as a double-debugger: one can visualize both the state of the interpreted program and the state of the interpreter program. This tool is intended for the JavaScript committee, VM developpers, and other experts in JavaScript semantics.

- M. Clochard, C. Marché, and A. Paskevich have 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 [69].

- M. Clochard, J.-C. Filliâtre, C. Marché, and A. Paskevich have developed a case study on the formalization of semantics of programming languages using Why3 [67]. This case study aims at illustrating recent improvements of Why3 regarding the support for higher-order logic features in the input logic of Why3, and how these are encoded into first-order logic, so that goals can be discharged by automated provers. This case study also illustrates how reasoning by induction can be done without need for interactive proofs, via the use of *lemma functions*.

- M. Clochard and L. Gondelman have developed a formalization of a simple compiler in Why3 [68]. It compiles a simple imperative language into assembler instructions for a stack machine. This case study was inspired by a similar example developed using Coq and interactive theorem proving. The aim is to improve significantly the degree of automation in the proofs. This is achieved by the formalization of a Hoare logic and a Weakest Precondition Calculus on assembly programs, so that the correctness of compilation is seen as a formal specification of the assembly instructions generated.

3.4.3. *Project-team Positioning*

The objective of formalizing languages and algorithms is very general, and it is pursued by several Inria teams. One common trait is the use of the Coq proof assistant for this purpose: Pi.r2 (development of Coq itself and its meta-theory), Gallium (semantics and compilers of programming languages), Marelle (formalization of mathematics), SpecFun (real arithmetic), Celtique (formalization of static analyzers).

Other environments for the formalization of languages include

- ACL2 system [0]: an environment for writing programs with formal specifications in first-order logic based on a Lisp engine. The proofs are conducted using a prover based on the Boyer-Moore approach. It is a rather old system but still actively maintained and powerful, developed at University of Texas at Austin. It has a strong industrial impact.

- Isabelle environment [0]: both a proof assistant and an environment for developing pure applicative programs. It is developed jointly at University of Cambridge, UK, Technische Universität München, Germany, and to some extent by the VALS team at LRI, Université Paris-Sud. It features highly automated tactics based on ATP systems (the Sledgehammer tool).

- The team "Trustworthy Systems" at NICTA in Australia [0] aims at developing highly trustable software applications. They developed a formally verified micro-kernel called seL4 [96], using a home-made layer to deal with C programs on top of the Isabelle prover.

- The PVS system [0] is an environment for both programming and proving (purely applicative) programs. It is developed at the Computer Science Laboratory of SRI international, California, USA. A major user of PVS is the team LFM [0] at NASA Langley, USA, for the certification of programs related to air traffic control.

In the Toccata team, we do not see these alternative environments as competitors, even though, for historical reasons, we are mainly using Coq. Indeed both Isabelle and PVS are available as back-ends of Why3.

3.5. Proof of Numerical Programs

Permanent researchers: S. Boldo, C. Marché, G. Melquiond

- Linked with objective 1 (Deductive Program Verification), the methodology for proving numerical C programs has been presented by S. Boldo in her habilitation [52] and as invited speaker [53]. An application is the formal verification of a numerical analysis program. S. Boldo, J.-C. Filliâtre, and G. Melquiond, with F. Clément and P. Weis (POMDAPI team, Inria Paris - Rocquencourt), and M. Mayero (LIPN), completed the formal proof of the second-order centered finite-difference scheme for the one-dimensional acoustic wave [55][4].

---

[0]http://www.cs.utexas.edu/~moore/acl2/
[0]http://isabelle.in.tum.de/
[0]http://ssrg.nicta.com.au/projects/TS/
[0]http://pvs.csl.sri.com/
[0]http://shemesh.larc.nasa.gov/fm/fm-main-team.html

- Several challenging floating-point algorithms have been studied and proved. This includes an algorithm by Kahan for computing the area of a triangle: S. Boldo proved an improvement of its error bound and new investigations in case of underflow [51]. This includes investigations about quaternions. They should be of norm 1, but due to the round-off errors, a drift of this norm is observed over time. C. Marché determined a bound on this drift and formally proved it correct [9]. P. Roux formally verified an algorithm for checking that a matrix is semi-definite positive [115]. The challenge here is that testing semi-definiteness involves algebraic number computations, yet it needs to be implemented using only approximate floating-point operations.

- Because of compiler optimizations (or bugs), the floating-point semantics of a program might change once compiled, thus invalidating any property proved on the source code. We have investigated two ways to circumvent this issue, depending on whether the compiler is a black box. When it is, T. Nguyen has proposed to analyze the assembly code it generates and to verify it is correct [112]. On the contrary, S. Boldo and G. Melquiond (in collaboration with J.-H. Jourdan and X. Leroy) have added support for floating-point arithmetic to the CompCert compiler and formally proved that none of the transformations the compiler applies modify the floating-point semantics of the program [58], [57].

- Linked with objectives 2 (Automated Reasoning) and 3 (Formalization and Certification of Languages, Tools and Systems), G. Melquiond has implemented an efficient Coq library for floating-point arithmetic and proved its correctness in terms of operations on real numbers [107]. It serves as a basis for an interval arithmetic on which Taylor models have been formalized. É. Martin-Dorel and G. Melquiond have integrated these models into CoqInterval [15]. This Coq library is dedicated to automatically proving the approximation properties that occur when formally verifying the implementation of mathematical libraries (libm).

- Double rounding occurs when the target precision of a floating-point computation is narrower than the working precision. In some situations, this phenomenon incurs a loss of accuracy. P. Roux has formally studied when it is innocuous for basic arithmetic operations [115]. É. Martin-Dorel and G. Melquiond (in collaboration with J.-M. Muller) have formally studied how it impacts algorithms used for error-free transformations [105]. These works were based on the Flocq formalization of floating-point arithmetic for Coq.

- By combining multi-precision arithmetic, interval arithmetic, and massively-parallel computations, G. Melquiond (in collaboration with G. Nowak and P. Zimmermann) has computed enough digits of the Masser-Gramain constant to invalidate a 30-year old conjecture about its closed form [108].

3.5.1. *Project-team Positioning*

This objective deals both with formal verification and floating-point arithmetic, which is quite uncommon. Therefore our competitors/peers are few. We may only cite the works by J. Duracz and M. Konečný, Aston University in Birmingham, UK.

The Inria team AriC (Grenoble - Rhône-Alpes) is closer to our research interests, but they are lacking manpower on the formal proof side; we have numerous collaborations with them. The Inria team Caramel (Nancy - Grand Est) also shares some research interests with us, though fewer; again, they do not work on the formal aspect of the verification; we have some occasional collaborations with them.

There are many formalization efforts from chip manufacturers, such as AMD (using the ACL2 proof assistant) and Intel (using the Forte proof assistants) but the algorithms they consider are quite different from the ones we study. The works on the topic of floating-point arithmetic from J. Harrison at Intel using HOL Light are really close to our research interests, but they seem to be discontinued.

A few deductive program verification teams are willing to extend their tools toward floating-point programs. This includes the KeY project and SPARK. We have an ongoing collaboration with the latter, in the context of the ProofInUSe project.

Deductive verification is not the only way to prove programs. Abstract interpretation is widely used, and several teams are interested in floating-point arithmetic. This includes the Inria team Antique (Paris - Rocquencourt) and a CEA List team, who have respectively developed the Astrée and Fluctuat tools. This approach targets a different class of numerical algorithms than the ones we are interested in.

Other people, especially from the SMT community (*cf* objective 2), are also interested in automatically proving formulas about floating-point numbers, notably at Oxford University. They are mainly focusing on pure floating-point arithmetic though and do not consider them as approximation of real numbers.

Finally, it can be noted that numerous teams are working on the verification of numerical programs, but assuming the computations are real rather than floating-point ones. This is out of the scope of this objective.

4. Application Domains

4.1. Safety-Critical Software

The application domains we target involve safety-critical software, that is where a high-level guarantee of soundness of functional execution of the software is wanted. Currently our industrial collaborations mainly belong to the domain of transportation, including aeronautics, railroad, space flight, automotive.

Verification of C programs, Alt-Ergo at Airbus   Transportation is the domain considered in the context of the ANR U3CAT project, led by CEA, in partnership with Airbus France, Dassault Aviation, Sagem Défense et Sécurité. It included proof of C programs via Frama-C/Jessie/Why, proof of floating-point programs [104], the use of the Alt-Ergo prover via CAVEAT tool (CEA) or Frama-C/WP. Within this context, we contributed to a qualification process of Alt-Ergo with Airbus industry: the technical documents (functional specifications and benchmark suite) have been accepted by Airbus, and these documents were submitted by Airbus to the certification authorities (DO-178B standard) in 2012. This action is continued in the new project Soprano.

Certified compilation, certified static analyzers   Aeronautics is the main target of the Verasco project, led by Verimag, on the development of certified static analyzers, in partnership with Airbus. This is a follow-up of the transfer of the CompCert certified compiler (Inria team Gallium) to which we contributed to the support of floating-point computations [58].

Transfer to the community of Ada development   The former FUI project Hi-Lite, led by Adacore company, introduced the use of Why3 and Alt-Ergo as back-end to SPARK2014, an environment for verification of Ada programs. This is applied to the domain of aerospace (Thales, EADS Astrium). At the very beginning of that project, Alt-Ergo was added in the Spark Pro toolset (predecessor of SPARK2014), developed by Altran-Praxis: Alt-Ergo can be used by customers as an alternate prover for automatically proving verification conditions. Its usage is described in the new edition of the Spark book (Chapter "Advanced proof tools"). This action is continued in the new joint laboratory ProofInUse. A recent paper [65] provides an extensive list of applications of SPARK, a major one being the British air control management *iFacts*.

Transfer to the community of Atelier B   In the current ANR project BWare, we investigate the use of Why3 and Alt-Ergo as an alternative back-end for checking proof obligations generated by *Atelier B*, whose main applications are railroad-related software [0], a collaboration with Mitsubishi Electric R&D Centre Europe (Rennes) (joint publication [109]) and ClearSy (Aix-en-Provence).

SMT-based Model-Checking: Cubicle   S. Conchon (with A. Mebsout and F. Zaidi from VALS team at LRI) has a long-term collaboration with S. Krstic and A. Goel (Intel Strategic Cad Labs in Hillsboro, OR, USA) that aims in the development of the SMT-based model checker Cubicle (http://cubicle.lri. fr/) based on Alt-Ergo [106][7]. It is particularly targeted to the verification of concurrent programs and protocols.

---

[0]http://www.methode-b.com/en/links/

Apart from transportation, energy is naturally an application in particular with our long-term partner CEA, in the context of U3CAT and Soprano projects. We also indirectly target communications and data, in particular in contexts with a particular need for security or confidentiality: smart phones, Web applications, health records, electronic voting, etc. These are part of the applications of SPARK [65], including verification of security-related properties, including cryptographic algorithms. Also, our new AJACS project addresses issues related to security and privacy in web applications written in Javascript, also including correctness properties.

5. Highlights of the Year

5.1. Highlights of the Year

- S. Conchon: co-organizes POPL'2017 (January, Paris, http://conf.researchr.org/home/POPL-2017)
    - Major Int. Conference on Foundations of Programming Language, Semantics, Type Systems, Formal Proof Techniques

5.1.1. Awards

- [April 2016] Martin Clochard, Léon Gondelman, Mário Pereira: jointly receive the "Best student team" award of the *VerifyThis@ETAPS2016 verification competition*
- [July 2016] S. Boldo: Best Talk Award at workshop NSV *Computing a correct and tight rounding error bound using rounding-to-nearest*

6. New Software and Platforms

6.1. Alt-Ergo

Automated theorem prover for software verification
KEYWORDS: Software Verification - Automated theorem proving
FUNCTIONAL DESCRIPTION

Alt-Ergo is an automatic solver of formulas based on SMT technology. It is especially designed to prove mathematical formulas generated by program verification tools, such as Frama-C for C programs, or SPARK for Ada code. Initially developed in Toccata research team, Alt-Ergo's distribution and support are provided by OCamlPro since September 2013.

- Participants: Sylvain Conchon, Evelyne Contejean, Mohamed Iguernelala, Stephane Lescuyer and Alain Mebsout
- Partner: OCamlPro
- Contact: Sylvain Conchon
- URL: http://alt-ergo.lri.fr

6.2. CFML

Interactive program verification using characteristic formulae
KEYWORDS: Coq - Software Verification - Deductive program verification - Separation Logic
FUNCTIONAL DESCRIPTION

The CFML tool supports the verification of OCaml programs through interactive Coq proofs. CFML proofs establish the full functional correctness of the code with respect to a specification. They may also be used to formally establish bounds on the asymptotic complexity of the code. The tool is made of two parts: on the one hand, a characteristic formula generator implemented as an OCaml program that parses OCaml code and produces Coq formulae, and, on the other hand, a Coq library that provides notation and tactics for manipulating characteristic formulae interactively in Coq.

- Contact: Arthur Chargueraud
- URL: http://www.chargueraud.org/softs/cfml/

6.3. Coq

KEYWORDS: Proof - Certification - Formalisation
FUNCTIONAL DESCRIPTION

Coq provides both a dependently-typed functional programming language and a logical formalism, which, altogether, support the formalisation of mathematical theories and the specification and certification of properties of programs. Coq also provides a large and extensible set of automatic or semi-automatic proof methods. Coq's programs are extractible to OCaml, Haskell, Scheme, ...

- Participants: Benjamin Gregoire, Enrico Tassi, Bruno Barras, Yves Bertot, Pierre Boutillier, Xavier Clerc, Pierre Courtieu, Maxime Denes, Stephane Glondu, Vincent Gross, Hugo Herbelin, Pierre Letouzey, Assia Mahboubi, Julien Narboux, Jean-Marc Notin, Christine Paulin-Mohring, Pierre-Marie Pedrot, Loic Pottier, Matthias Puech, Yann Regis-Gianas, François Ripault, Matthieu Sozeau, Arnaud Spiwack, Pierre-Yves Strub, Benjamin Werner, Guillaume Melquiond and Jean-Christophe Filliatre
- Partners: CNRS - ENS Lyon - Université Paris-Diderot - Université Paris-Sud
- Contact: Hugo Herbelin
- URL: http://coq.inria.fr/

6.4. CoqInterval

Interval package for Coq
KEYWORDS: Interval arithmetic - Coq
FUNCTIONAL DESCRIPTION

CoqInterval is a library for the proof assistant Coq. CoqInterval provides a method for proving automatically the inequality of two expression of real values.

The Interval package provides several tactics for helping a Coq user to prove theorems on enclosures of real-valued expressions. The proofs are performed by an interval kernel which relies on a computable formalization of floating-point arithmetic in Coq.

The Marelle team developed a formalization of rigorous polynomial approximation using Taylor models inside the Coq proof assistant, with a special focus on genericity and efficiency for the computations. In 2014, this library has been included in CoqInterval.

- Participants: Guillaume Melquiond, Erik Martin Dorel, Nicolas Brisebarre, Miora Maria Joldes, Micaela Mayero, Jean Michel Muller, Laurence Rideau and Laurent Thery
- Contact: Guillaume Melquiond
- URL: http://coq-interval.gforge.inria.fr/

6.5. Coquelicot

The Coquelicot library for real analysis in Coq
KEYWORDS: Coq - Real analysis

FUNCTIONAL DESCRIPTION

Coquelicot is library aimed for supporting real analysis in the Coq proof assistant. It is designed with three principles in mind. The first is the user-friendliness, achieved by implementing methods of automation, but also by avoiding dependent types in order to ease the stating and readability of theorems. This latter part was achieved by defining total function for basic operators, such as limits or integrals. The second principle is the comprehensiveness of the library. By experimenting on several applications, we ensured that the available theorems are enough to cover most cases. We also wanted to be able to extend our library towards more generic settings, such as complex analysis or Euclidean spaces. The third principle is for the Coquelicot library to be a conservative extension of the Coq standard library, so that it can be easily combined with existing developments based on the standard library.

- Participants: Sylvie Boldo, Catherine Lelay and Guillaume Melquiond
- Contact: Sylvie Boldo
- URL: http://coquelicot.saclay.inria.fr/

6.6. Cubicle

The Cubicle model checker modulo theories
KEYWORDS: Model Checking - Software Verification
FUNCTIONAL DESCRIPTION

Cubicle is an open source model checker for verifying safety properties of array-based systems, which corresponds to a syntactically restricted class of parametrized transition systems with states represented as arrays indexed by an arbitrary number of processes. Cache coherence protocols and mutual exclusion algorithms are typical examples of such systems.

- Participants: Sylvain Conchon and Alain Mebsout
- Contact: Sylvain Conchon
- URL: http://cubicle.lri.fr/

6.7. Flocq

The Flocq library for formalizing floating-point arithmetic in Coq
KEYWORDS: Floating-point - Arithmetic code - Coq
FUNCTIONAL DESCRIPTION

The Flocq library for the Coq proof assistant is a comprehensive formalization of floating-point arithmetic: core definitions, axiomatic and computational rounding operations, high-level properties. It provides a framework for developers to formally certify numerical applications.

Flocq is currently used by the CompCert certified compiler for its support of floating-point computations.

- Participants: Guillaume Melquiond and Sylvie Boldo
- Contact: Sylvie Boldo
- URL: http://flocq.gforge.inria.fr/

6.8. Gappa

The Gappa tool for automated proofs of arithmetic properties
KEYWORDS: Floating-point - Arithmetic code - Software Verification - Constraint solving
FUNCTIONAL DESCRIPTION

Gappa is a tool intended to help verifying and formally proving properties on numerical programs dealing with floating-point or fixed-point arithmetic. It has been used to write robust floating-point filters for CGAL and it is used to certify elementary functions in CRlibm. While Gappa is intended to be used directly, it can also act as a backend prover for the Why3 software verification plateform or as an automatic tactic for the Coq proof assistant.

- Contact: Guillaume Melquiond
- URL: http://gappa.gforge.inria.fr/

6.9. Why3

The Why3 environment for deductive verification
KEYWORDS: Formal methods - Trusted software - Software Verification - Deductive program verification
FUNCTIONAL DESCRIPTION

Why3 is an environment for deductive program verification. It provides a rich language for specification and programming, called WhyML, and relies on external theorem provers, both automated and interactive, to discharge verification conditions. Why3 comes with a standard library of logical theories (integer and real arithmetic, Boolean operations, sets and maps, etc.) and basic programming data structures (arrays, queues, hash tables, etc.). A user can write WhyML programs directly and get correct-by-construction OCaml programs through an automated extraction mechanism. WhyML is also used as an intermediate language for the verification of C, Java, or Ada programs.

- Participants: Jean-Christophe Filliatre, Claude Marche, Guillaume Melquiond, Andriy Paskevych, François Bobot, Martin Clochard and Levs Gondelmans
- Partners: CNRS - Université Paris-Sud
- Contact: Claude Marche
- URL: http://why3.lri.fr/

7. New Results

7.1. Deductive Verification

A bit-vector library for deductive verification. C. Fumex and C. Marché developed a new library for bit-vectors, in Why3 and SPARK. This library is rich enough for the formal specification of functional behavior of programs that operate at the level of bits. It is also designed to exploit efficiently the support for bit-vectors built-in in some SMT solvers. This work is done in the context of the ProofInUse joint laboratory. The SPARK front-end of Why3, for the verification of Ada programs, is extended to exploit this new bit-vector theory. Several cases studies are conducted: efficient search for rightmost bit of a bit-vector, efficient computation of the number of bits set to 1, efficient solving of the n-queens problem. At the level of SPARK, a program inspired from some industrial code (originally developed in C par J. Gerlach, Fraunhofer FOKUS Institute, Germany and partially proved with Frama-C and Coq) is specified in SPARK and proved with automatic solvers only. A paper on that library together with the way it is connected with the built-in support for bitvectors in SMT solver was presented at the NASA Formal methods Conference [24]. The support for bit-vectors is distributed with SPARK since 2015, and SPARK users already reported that several verification conditions, that couldn't be proved earlier, are now proved automatically.

Counterexamples from proof failures.   D. Hauzar and C. Marché worked on counterexample generation from failed proof attempts. They designed a new approach for generating potential counterexamples in the deductive verification setting, and implemented in Why3. When the logic goal generated for a given verification condition is not shown unsatisfiable by an SMT solvers, some solver can propose a model. By carefully reverting the transformation chain (from an input program through the VC generator and the various translation steps to solvers), this model is turned into a potential counterexample that the user can exploit to analyze why its original code is not proved. The approach is implemented in the chain from Ada programs through SPARK, Why3, and SMT solvers CVC4 and Z3. This work is described in a research report [35] and a paper was rpesented at the SEFM Conference [25]. The work on the implementation was continued by S. Dailler. It was considered robust enough to be distributed in the release Pro 16 of SPARK.

Static versus dynamic verification.   C. Marché, together with Y. Moy from AdaCore, J. Signoles and N. Kosmatov from CEA-LIST, wrote a survey paper about the design of the specification languages of Why3 and its front-ends Frama-C and SPARK. The choices made when designing these specification languages differ significantly, in particular with respect to the executability of specifications. The paper reviews these differences and the issues that result from these choices. The paper also emphasizes two aspects where static and dynamic aspects of the specifications play an important role: the specific feature of *ghost code*, and the techniques that help users understand why static verification fails. This paper was presented at the Isola Symposium [26].

Higher-Order Representation Predicates.   A. Charguéraud investigated how to formalize in Separation Logic representation predicates for describing mutable container data structures that store mutable elements that are themselves described using representation predicates. (In Separation Logic, representation predicates are used to describe mutable data structures, by establishing a relationship between the entry point of the structure, the piece of heap over which this structure spans, and the logical model associated with the structure.) The solution proposed, based on "higher-order representation predicates", allows for concise specifications of such containers. A. Charguéraud has published a paper presenting, through a collection of practical examples, solutions to the challenges associated with verification proofs based on higher-order representation predicates [19].

Temporary Read-Only Permissions for Separation Logic   A. Charguéraud and François Pottier (Inria Paris) have developed an extension of Separation Logic with temporary read-only permissions. This mechanism allows to temporarily convert any assertion (or "permission") to a read-only form. Unlike with fractional permissions, no accounting is required: the proposed read-only permissions can be freely duplicated and discarded. Where mutable data structures are temporarily accessed only for reading, the proposed read-only permissions enable more concise specifications and proofs. All the metatheory is verified in Coq. An article has been submitted to a conference [20].

Reasoning About Iteration.   J.-C. Filliâtre and M. Pereira proposed a new approach to the problem of specifying iteration, verifying iterators (such as cursors or higher-order functions), and using iterators. The idea is to characterize the sequence of elements enumerated so far, and only those. The proof methodology is modular, iterator implementations and clients being verified independently of each other. The proposed method is validated experimentally in Why3. This work has been published first at JFLA 2016 [33] and then at NFM 2016 [22]. A journal version of this work is under submission.

Defunctionalization for proving higher-order programs.   J.-C. Filliâtre and M. Pereira proposed a new approach to the verification of higher-order programs, using the technique of defunctionalization, that is, the translation of first-class functions into first-order values. This is an early experimental work, conducted on examples only within the Why3 system. This work has been published at JFLA 2017 [30].

A Type System for Deductive Verification.   J.-C. Filliâtre, L. Gondelman, and A. Paskevich proposed a practical method to track pointer aliases statically in a large family of computer programs. Their approach relies on a special type system with singleton regions and effects which both can be

inferred automatically, without requiring additional user annotations. This kind of static analysis is important for deductive program verification, since it allows us to construct verification conditions using the traditional rules in the spirit of Hoare and Dijkstra, without recurring to more sophisticated solutions (memory models, separation logic) which incur additional complexity both for a user and a verification tool. The proposed method is implemented in Why3 and described in a technical report [37].

Ghost Code. J.-C. Filliâtre, L. Gondelman, and A. Paskevich published a paper on a general approach to the concept of ghost code in the journal of *Formal Methods in System Design* [14]. Ghost code is a subset of program code that serves the purposes of specification and verification: it can be erased from the program without affecting its result. This work forms the basis of the support for ghost code in Why3. This work is an extended version of the paper presented at the 26th International Conference on Computer Aided Verification (CAV) in 2014.

7.2. Automated Reasoning

Decision Procedures via Axiomatizations with Triggers. C. Dross, A. Paskevich, J. Kanig and S. Conchon published a paper in the *Journal of Automated Reasoning* [13] about integration of first-order axiomatizations with triggers as decision procedures in an SMT solver. This work extends a part of C. Dross PhD thesis [83]. A formal semantics of the notion of trigger is presented, with a general setting to show how a first-order axiomatization with triggers can be proved correct, complete, and terminating. An extended DPLL(T) algorithm can then integrate such an axiomatization with triggers, as a decision procedure for the theory it defines.

Lightweight Approach for Declarative Proofs. M. Clochard designed an extension of first-order logic, for describing reasoning steps needed to discharge a proof obligation. The extension is under the form of two new connectives, called proof indications, that allow the user to encode reasoning steps inside a logic formula. This extension makes possible to use the syntax of formulas as a proof language. The approach was presented at the JFLA conference [29] and implemented in Why3. It brings a lightweight mechanism for declarative proofs in an environment like Why3 where provers are used as black boxes. Moreover, this mechanism restricts the scope of auxiliary lemmas, reducing the size of proof obligations sent to external provers.

7.3. Certification of Algorithms, Languages, Tools and Systems

Case study: Matrix Multiplication. M. Clochard, L. Gondelman and M. Pereira wrote a paper describing a complete solution for the first challenge of the VerifyThis 2016 competition held at the 18th ETAPS Forum, where they obtain the award for the best student team. Two variants for the multiplication of matrices are presented and proved: a naive version using three nested loops and Strassen's algorithm. To formally specify the two multiplication algorithms, they developed a new Why3 theory of matrices, and they applied a reflection methodology to condut some of the proofs. This work was presented at the VSTTE Conference [21]. An extended version that considers arbitrary rectangular martices instead of square ones is in preparation. The development is available at http://toccata.lri.fr/gallery/verifythis_2016_matrix_multiplication.en.html.

Case study: Koda-Ruskey's algorithm for generating ideals of a forest. J.-C. Filliâtre and M. Pereira presented the first formal proof of an implementation of Koda and Ruskey's algorithm (an algorithm for generating all ideals of a forest poset as a Gray code) at VSTTE 2016 [23]. The proof is conducted within the Why3 system and is mostly automatic.

The Lax-Milgram Theorem. S. Boldo, F. Clément, F. Faissole, V. Martin, and M. Mayero have worked on a Coq formal proof of the Lax–Milgram theorem. The Finite Element Method is a widely-used method to solve numerical problems coming for instance from physics or biology. To obtain the highest confidence on the correction of numerical simulation programs implementing the Finite Element Method, one has to formalize the mathematical notions and results that allow to establish the sound-ness of the method. The Lax–Milgram theorem may be seen as one of those

theoretical cornerstones: under some completeness and coercivity assumptions, it states existence and uniqueness of the solution to the weak formulation of some boundary value problems. This article presents the full formal proof of the Lax–Milgram theorem in Coq. It requires many results from linear algebra, geometry, functional analysis , and Hilbert spaces. This has been published at the 6th ACM SIGPLAN Conference on Certified Programs and Proofs (CPP 2017) [18].

ALEA library extended with continuous datatypes   The ALEA library uses a monadic construction to formalize discrete measure theory. F. Faissole and B. Spitters proposed to extend it to continuous datatypes. They used both synthetic topology and homotopy type theory to achieve the formalization. This work is presented at the Workshop on Coq for Programming Languages [32].

Case study: Strongly Connected Components of a Graph   R. Chen and J.-J. Lévy designed a formal proof of Tarjan's algorithm for computing the strongly connected component of a directed graph. The proof is conducted using Why3. This work is presented at the JFLA conference [28]. This case study is part of a larger set of case studies on algorithms on graphs http://pauillac.inria.fr/~levy/why3/.

Case study: Unix Pathname Resolution   R. Chen, M. Clochard and C.-Marché designed a formal proof of an algorithm for resolving a pathname in Unix file systems. The proof is conducted using Why3 [34]. This case study is part of the CoLiS project.

7.4. Floating-Point and Numerical Programs

-

Interval arithmetic and Taylor models.   É. Martin-Dorel and G. Melquiond have worked on integrating the CoqInterval and CoqApprox libraries into a single package. The CoqApprox library is dedicated to computing verified Taylor models of univariate functions so as to compute approximation errors. The CoqInterval library reuses this work to automatically prove bounds on real-valued expressions. A large formalization effort took place during this work, so as to get rid of all the holes remaining in the formal proofs of CoqInterval. It was also the chance to perform a comparison between numerous decision procedures dedicated to proving nonlinear inequalities involving elementary functions. This work has been published in the *Journal of Automated Reasoning* [15].

Interval arithmetic and univariate integrals.   A. Mahboubi, G. Melquiond, and T. Sibut-Pinote have extended the CoqInterval library with support for definite univariate integrals. The library is now able to automatically and formally verify bounds on the value of integrals by computing rigorous polynomial approximations of integrands. This work has been presented at the 7th International Conference on Interactive Theorem Proving [27].

Robustness of 2Sum and Fast2Sum.   S. Boldo, S.Graillat, and J.-M. Muller have worked on the 2Sum and Fast2Sum algorithms, that are important building blocks in numerical computing. They are used (implicitely or explicitely) in many compensated algorithms or for manipulating floating-point expansions. They showed that these algorithms are much more robust than it is usually believed: the returned result makes sense even when the rounding function is not round-to-nearest, and they are almost immune to overflow. This work has been submitted [36].

Computing error bounds without changing the rouding mode.   S. Boldo has created an algorithm to compute a correct and tight rounding error bound for a floating-point computation. The rounding error can be bounded by folklore formulas, such as $\varepsilon |x|$ or $\varepsilon | \circ (x)|$. This gets more complicated when underflow is taken into account. To compute this error bound in practice, a directed rounding is usually used. This work describes an algorithm that computes a correct bound using only rounding to nearest, therefore without requiring a costly change of the rounding mode. This is formally proved using the Coq formal proof assistant to increase the trust in this algorithm. This has been published at the 9th International Workshop on Numerical Software Verification [17].

Floating-Point Computations and Iterators.   S. Boldo has worked on the formal verification of a floating-point case study where the common iterators `fold_left` and `fold_right` have not the wanted behaviors. She then had to define other iterators, which are very similar in most cases, but that do behave well in our case study. This has been published at the 1st Workshop on High-Consequence Control Verification [31].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. *ProofInUse Joint Laboratory*

Participants: Claude Marché [contact], Jean-Christophe Filliâtre, Andrei Paskevich.

ProofInUse is a joint project between the Toccata team and the SME AdaCore. It was selected and funded by the ANR programme "Laboratoires communs", starting from April 2014, for 3 years http://www.spark-2014.org/proofinuse.

The SME AdaCore is a software publisher specializing in providing software development tools for critical systems. A previous successful collaboration between Toccata and AdaCore enabled *Why3* technology to be put into the heart of the AdaCore-developed SPARK technology.

The goal is now to promote and transfer the use of deduction-based verification tools to industry users, who develop critical software using the programming language Ada. The proof tools are aimed at replacing or complementing the existing test activities, whilst reducing costs.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. *ELFIC*

Participants: Sylvie Boldo [contact], Claude Marché, Guillaume Melquiond.

ELFIC is a working group of the Digicosme Labex. S. Boldo is the principal investigator. It began in 2014 for one year and was extended for one year. https://digicosme.lri.fr/GT+ELFIC

The ELFIC project focuses on proving the correctness of the FELiScE (Finite Elements for Life Sciences and Engineering) C++ library which implements the finite element method for approximating solutions to partial differential equations. Finite elements are at the core of numerous simulation programs used in industry. The formal verification of this library will greatly increase confidence in all the programs that rely on it. Verification methods developed in this project will be a breakthrough for the finite element method, but more generally for the reliability of critical software relying on intricate numerical algorithms.

Partners: Inria team Pomdapi; Ecole Polytechnique, LIX; CEA LIST; Université Paris 13, LIPN; UTC, LMAC (Compiègne).

9.1.2. *ELEFFAN*

Participant: Sylvie Boldo [contact].

ELEFFAN is a Digicosme project funding the PhD of F. Faissole. S. Boldo is the principal investigator. It began in 2016 for three years. https://project.inria.fr/eleffan/

The ELEFFAN project aims at formally proving rounding error bounds of numerical schemes.

Partners: ENSTA Paristech (A. Chapoutot)

9.2. National Initiatives

9.2.1. *ANR CoLiS*

Participants: Claude Marché [contact], Andrei Paskevich.

The CoLiS research project is funded by the programme "Société de l'information et de la communication" of the ANR, for a period of 48 months, starting on October 1st, 2015. http://colis.irif.univ-paris-diderot.fr/

The project aims at developing formal analysis and verification techniques and tools for scripts. These scripts are written in the POSIX or bash shell language. Our objective is to produce, at the end of the project, formal methods and tools allowing to analyze, test, and validate scripts. For this, the project will develop techniques and tools based on deductive verification and tree transducers stemming from the domain of XML documents.

Partners: Université Paris-Diderot, IRIF laboratory (formerly PPS & LIAFA), coordinator ; Inria Lille, team LINKS

9.2.2. *ANR Vocal*

Participants: Jean-Christophe Filliâtre [contact], Andrei Paskevich.

The Vocal research project is funded by the programme "Société de l'information et de la communication" of the ANR, for a period of 48 months, starting on October 1st, 2015. https://vocal.lri.fr/

The goal of the Vocal project is to develop the first formally verified library of efficient general-purpose data structures and algorithms. It targets the OCaml programming language, which allows for fairly efficient code and offers a simple programming model that eases reasoning about programs. The library will be readily available to implementers of safety-critical OCaml programs, such as Coq, Astrée, or Frama-C. It will provide the essential building blocks needed to significantly decrease the cost of developing safe software. The project intends to combine the strengths of three verification tools, namely Coq, Why3, and CFML. It will use Coq to obtain a common mathematical foundation for program specifications, as well as to verify purely functional components. It will use Why3 to verify a broad range of imperative programs with a high degree of proof automation. Finally, it will use CFML for formal reasoning about effectful higher-order functions and data structures making use of pointers and sharing.

Partners: team Gallium (Inria Paris-Rocquencourt), team DCS (Verimag), TrustInSoft, and OCamlPro.

9.2.3. *ANR Ajacs*

Participant: Arthur Charguéraud [contact].

The AJACS research project is funded by the programme "Société de l'information et de la communication" of the ANR, for a period of 42 months, starting on October 1st, 2014. http://ajacs.inria.fr/

The goal of the AJACS project is to provide strong security and privacy guarantees on the client side for web application scripts implemented in JavaScript, the most widely used language for the Web. The proposal is to prove correct analyses for JavaScript programs, in particular information flow analyses that guarantee no secret information is leaked to malicious parties. The definition of sub-languages of JavaScript, with certified compilation techniques targeting them, will allow deriving more precise analyses. Another aspect of the proposal is the design and certification of security and privacy enforcement mechanisms for web applications, including the APIs used to program real-world applications. On the Toccata side, the focus will be on the formalization of secure subsets of JavaScript, and on the mechanization of proofs of translations from high-level languages into JavaScript.

Partners: team Celtique (Inria Rennes - Bretagne Atlantique), team Prosecco (Inria Paris - Rocquencourt), team Indes (Inria Sophia Antipolis - Méditerranée), and Imperial College (London).

9.2.4. *ANR FastRelax*

Participants: Sylvie Boldo [contact], Guillaume Melquiond.

This is a research project funded by the programme "Ingénierie Numérique & Sécurité" of the ANR. It is funded for a period of 48 months and it has started on October 1st, 2014. http://fastrelax.gforge.inria.fr/

Our aim is to develop computer-aided proofs of numerical values, with certified and reasonably tight error bounds, without sacrificing efficiency. Applications to zero-finding, numerical quadrature or global optimization can all benefit from using our results as building blocks. We expect our work to initiate a "fast and reliable" trend in the symbolic-numeric community. This will be achieved by developing interactions between our fields, designing and implementing prototype libraries and applying our results to concrete problems originating in optimal control theory.

Partners: team ARIC (Inria Grenoble Rhône-Alpes), team MARELLE (Inria Sophia Antipolis - Méditerranée), team SPECFUN (Inria Saclay - Île-de-France), Université Paris 6, and LAAS (Toulouse).

9.2.5. *ANR Soprano*

Participants: Sylvain Conchon [contact], Guillaume Melquiond.

The Soprano research project is funded by the programme "Sciences et technologies logicielles" of the ANR, for a period of 42 months, starting on October 1st, 2014. http://soprano-project.fr/

The SOPRANO project aims at preparing the next generation of verification-oriented solvers by gathering experts from academia and industry. We will design a new framework for the cooperation of solvers, focused on model generation and borrowing principles from SMT (current standard) and CP (well-known in optimization). Our main scientific and technical objectives are the following. The first objective is to design a new collaboration framework for solvers, centered around synthesis rather than satisfiability and allowing cooperation beyond that of Nelson-Oppen while still providing minimal interfaces with theoretical guarantees. The second objective is to design new decision procedures for industry-relevant and hard-to-solve theories. The third objective is to implement these results in a new open-source platform. The fourth objective is to ensure industrial-adequacy of the techniques and tools developed through periodical evaluations from the industrial partners.

Partners: team DIVERSE (Inria Rennes - Bretagne Atlantique), Adacore, CEA List, Université Paris-Sud, and OCamlPro.

9.2.6. *ANR CAFEIN*

Participant: Sylvain Conchon [contact].

The CAFEIN research project is funded by the programme "Ingénierie Numérique & Sécurité" of the ANR, for a period of 3 years, starting on February 1st, 2013. https://cavale.enseeiht.fr/CAFEIN/.

This project addresses the formal verification of functional properties at specification level, for safety critical reactive systems. In particular, we focus on command and control systems interacting with a physical environment, specified using the synchronous language Lustre.

A first goal of the project is to improve the level of automation of formal verification, by adapting and combining existing verification techniques such as SMT-based temporal induction, and abstract interpretation for invariant discovery. A second goal is to study how knowledge of the mathematical theory of hybrid command and control systems can help the analysis at the controller's specification level. Third, the project addresses the issue of implementing real valued specifications in Lustre using floating-point arithmetic.

Partners: ONERA, CEA List, ENSTA, teams Maxplus (Inria Saclay - Île-de-France), team Parkas (Inria Paris - Rocquencourt), Perpignan University, Prover Technology, Rockwell Collins.

9.2.7. *ANR BWare*

Participants: Sylvain Conchon [contact], Jean-Christophe Filliâtre, Andrei Paskevich, Claude Marché.

The BWare research project is funded by the programme "Ingénierie Numérique & Sécurité" of the ANR, a period of 4 years, starting on September 1st, 2012. http://bware.lri.fr.

BWare is an industrial research project that aims to provide a mechanized framework to support the automated verification of proof obligations coming from the development of industrial applications using the B method and requiring high guarantee of confidence. The methodology used in this project consists of building a generic platform of verification relying on different theorem provers, such as first-order provers and SMT solvers. The variety of these theorem provers aims at allowing a wide panel of proof obligations to be automatically verified by the platform. The major part of the verification tools used in BWare have already been involved in some experiments, which have consisted in verifying proof obligations or proof rules coming from industrial applications [109]. This therefore should be a driving factor to reduce the risks of the project, which can then focus on the design of several extensions of the verification tools to deal with a larger amount of proof obligations.

The partners are: Cedric laboratory at CNAM (CPR Team, project leader); teams Gallium and Deducteam (Inria Paris - Rocquencourt) ; Mitsubishi Electric R&D Centre Europe, ClearSy (the company which develops and maintains *Atelier B*), and the start-up OCamlPro.

9.2.8. *ANR Verasco*

Participants: Guillaume Melquiond [contact], Sylvie Boldo, Arthur Charguéraud, Claude Marché.

The Verasco research project is funded by the programme "Ingénierie Numérique & Sécurité" of the ANR, for a period of 4 years and a half, starting on January 1st, 2012. Project website: http://verasco.imag.fr.

The main goal of the project is to investigate the formal verification of static analyzers and of compilers, two families of tools that play a crucial role in the development and validation of critical embedded software. More precisely, the project aims at developing a generic static analyzer based on abstract interpretation for the C language, along with a number of advanced abstract domains and domain combination operators, and prove the soundness of this analyzer using the *Coq* proof assistant. Likewise, the project keeps working on the CompCert C formally-verified compiler, the first realistic C compiler that has been mechanically proved to be free of miscompilation, and carry it to the point where it could be used in the critical software industry.

Partners: teams Gallium and Abstraction (Inria Paris - Rocquencourt), Airbus avionics and simulation (Toulouse), IRISA (Rennes), Verimag (Grenoble).

9.2.9. *FUI LCHIP*

Participant: Sylvain Conchon [contact].

LCHIP (Low Cost High Integrity Platform) is aimed at easing the development of safety critical applications (up to SIL4) by providing: (i) a complete IDE able to automatically generate and prove bounded complexity software (ii) a low cost, safe execution platform. The full support of DSLs and third party code generators will enable a seamless deployment into existing development cycles. LCHIP gathers scientific results obtained during the last 20 years in formal methods, proof, refinement, code generation, etc. as well as a unique return of experience on safety critical systems design. http://www.clearsy.com/en/2016/10/4260/

Partners: 2 technology providers (ClearSy, OcamlPro), in charge of building the architecture of the platform ; 3 labs (IFSTTAR, LIP6, LRI), to improve LCHIP IDE features ; 2 large companies (SNCF, RATP), representing public ordering parties, to check compliance with standard and industrial railway use-case.

The project lead by ClearSy has started in April 2016 and lasts 3 years. It is funded by BpiFrance as well as French regions.

9.2.10. *ANR PARDI*

Participant: Sylvain Conchon [contact].

Verification of parameterized distributed systems, 2016-2021.

Partners: Université Paris VI - Université Paris XI - Inria NANCY

9.3. European Initiatives

9.3.1. *FP7 & H2020 Projects*

> Project acronym: ERC Deepsea
> Project title: Parallel dynamic computations
> Duration: Jun. 2013 - Jun. 2018
> Coordinator: Umut A. Acar
> Other partners: Carnegie Mellon University

Abstract:

The objective of this project is to develop abstractions, algorithms and languages for parallelism and dynamic parallelism with applications to problems on large data sets. Umut A. Acar (affiliated to Carnegie Mellon University and Inria Paris - Rocquencourt) is the principal investigator of this ERC-funded project. The other main researchers involved are Mike Rainey (Inria, Gallium team), who is full-time on the project, and Arthur Charguéraud (Inria, Toccata team), who works 40% of his time to the project. Project website: http://deepsea.inria.fr/.

9.3.2. Collaborations in European Programs, Except FP7 & H2020

Program: COST (European Cooperation in Science and Technology).

Project acronym: EUTypes https://eutypes.cs.ru.nl/

Project title: The European research network on types for programming and verification

Duration: 2015-2019

Coordinator: Herman Geuvers, Radboud University Nijmegen, The Netherlands

Other partners: 36 members countries, see http://www.cost.eu/COST_Actions/ca/CA15123?parties

Abstract: Types are pervasive in programming and information technology. A type defines a formal interface between software components, allowing the automatic verification of their connections, and greatly enhancing the robustness and reliability of computations and communications. In rich dependent type theories, the full functional specification of a program can be expressed as a type. Type systems have rapidly evolved over the past years, becoming more sophisticated, capturing new aspects of the behaviour of programs and the dynamics of their execution.

This COST Action will give a strong impetus to research on type theory and its many applications in computer science, by promoting (1) the synergy between theoretical computer scientists, logicians and mathematicians to develop new foundations for type theory, for example as based on the recent development of "homotopy type theory", (2) the joint development of type theoretic tools as proof assistants and integrated programming environments, (3) the study of dependent types for programming and its deployment in software development, (4) the study of dependent types for verification and its deployment in software analysis and verification. The action will also tie together these different areas and promote cross-fertilisation.

9.3.3. Collaborations with Major European Organizations

Imperial College London (UK)

Certification of JavaScript, AJACS project

9.4. International Research Visitors

9.4.1. Visits of International Scientists

- Ran Chen is a PhD student from Institute of Software (Chinese Academy of Sciences, Beijing, China) visiting the team for 10 months under the supervision of C. Marché and J.-J. Lévy (PiR2 team, Inria Paris). She is working on the formal verification of graphs algorithms, and also in the context of the CoLiS project on verification of some aspects of the Unix file system and shell scripts. [34]

- Cláudio Belo Lourenço is a PhD student from Universidade do Minho, Portugal. He studies deductive verification of imperative programs and the behaviour of different kinds of verification condition generators [101]. The goal of his visit is to use Why3 as a platform for prototyping and experimental evaluation of these generators.

9.4.2. Visits to International Teams

9.4.2.1. Research Stays Abroad

- F. Faissole has spent two months visiting B. Spitters at Aarhus University (Denmark) They proposed an extension of ALEA library to continuous datatypes.[32].

- M. Roux has spent three months with D. Jovanovic and B. Dutertre at SRI (California, USA). They worked on extending Sally, the new model checker of SRI based on SAL, to add the verification of parameterized cache coherence protocols. The software can be found on https://github.com/SRI-CSL/sally.

- S. Conchon has been invited a month at SRI by D. Jovanovic. During this visit, he has collaborated with CSL researchers to compare the design and implementation choices between the model checkers Sally and Cubicle.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- S. Boldo, vice-president of the 28th "Journées Francophones des Langages Applicatifs" (JFLA 2017)

- J.-C. Filliâtre, organizer of EJCP (École Jeunes Chercheurs en Programmation du GDR GPL) at Lille on June 27–July 1, 2016. 42 participants. http://ejcp2016.univ-lille1.fr/

- A. Paskevich, program chair of the 9th Working Conference on Verified Software: Theories, Tools, and Experiments (VSTTE 2017), in collaboration with Thomas Wies (NYU).

10.1.1.2. Member of the Organizing Committees

- S. Conchon, local chair for the 44th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL 2017), held in Paris, France in January 2017.

- G. Melquiond, web chair for the 23rd IEEE Symposium on Computer Arithmetic (Arith 23), held in Silicon Valley, USA in July 2016.

- S. Boldo, member of the organization committee of the Inria Scientific Days in Rennes (June 2016).

10.1.2. Scientific Events Selection

10.1.2.1. Member of the Conference Program Committees

- S. Boldo, PC of the 23rd IEEE Symposium on Computer Arithmetic (ARITH 2016).

- S. Boldo, PC of the 9th International Workshop on Numerical Software Verification (NSV 2016).

- S. Boldo, PC of the 1st Workshop on High-Consequence Control Verification (HCCV 2016).

- S. Boldo, PC of the 27th "Journées Francophones des Langages Applicatifs" (JFLA 2016).

- A. Charguéraud, PC of the International Workshop on Hammers for Type Theories (HaTT 2016).

- A. Charguéraud, PC of the Workshop on ML (ML 2016).

- G. Melquiond, PC of the 3rd International Workshop on Coq for Programming Languages (CoqPL 2017).

10.1.2.2. Reviewer

The members of the Toccata team have reviewed papers for numerous international conferences.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- G. Melquiond is a member of the editorial board of *Reliable Computing*.
- S. Boldo is member of the editorial board of Binaire http://binaire.blog.lemonde.fr, the blog of the French Computer Science Society.
- J.-C. Filliâtre is member of the editorial board of the *Journal of Functional Programming*.
- C. Paulin is member of the editorial board of the *Journal of Formalized Reasoning*.

10.1.3.2. Reviewer - Reviewing Activities

The members of the Toccata team have reviewed numerous papers for numerous international journals.

10.1.4. Invited Talks

- S. Boldo gave an invited lecture at Effective Analysis: Foundations, Implementations, Certification in January 2016 at Marseille, France.
- S. Boldo gave a keynote talk in Cambridge at a local workshop about Testing and Verification in Computational Science.
- A. Charguéraud gave an invited talk at the Royal Society specialist meeting "Verified trustworthy software systems", presenting an interactive interpreter for the semantics of JavaScript, in London, on April 7th.
- S. Conchon gave an invited lecture "Model Checking Modulo Theories with Cubicle" at the 2016 edition of the School for Junior Researchers in Programming (École Jeunes Chercheurs en Programmation, EJCP 2016, http://ejcp2016.univ-lille1.fr/) held in Lille, France.
- A. Paskevich gave an invited lecture "Deductive Program Verification using Why3" at the 2016 edition of the School for Junior Researchers in Programming (École Jeunes Chercheurs en Programmation, EJCP 2016, http://ejcp2016.univ-lille1.fr/) held in Lille, France.

10.1.5. Leadership within the Scientific Community

- C. Paulin, scientific leader of Labex DigiCosme http://labex-digicosme.fr (Digital Worlds: distributed data, programs and architectures), until June 2016. It is a project launched by the French Ministry of research and higher education as part of the program "Investissements d'avenir", it involves the 14 research units in computer science and communications from the "Paris-Saclay" cluster.
- C. Paulin, dean of the Faculty of Sciences of Université Paris-Sud, since July 2016.

10.1.6. Scientific Expertise

- S. Boldo, member of the reviewing board for the ANR (first step in 2016).
- S. Boldo, member of the 2016 committee for the Gilles Kahn PhD award of the French Computer Science Society.
- S. Conchon and A. Paskevich, members of the *"commission consultative de spécialistes de l'université"*, Section 27, University Paris-Sud since December 2014.
- C. Marché, president of the evaluation committee of the joint Digiteo-DigiCosme call for projects https://digicosme.lri.fr/AAPDigiteoDigiCosme2016. The committee selected 10 thesis projects for funding, among 52 submissions. The committee also selected to support 10 scientific events in the Île-de-France region.
- C. Marché, member of the scientific commission of Inria-Saclay, in charge of selecting candidates for PhD grants, Post-doc grants, temporary leaves from universities ("délégations")
- C. Marché, member of the "Bureau du Comité des Projets" of Inria-Saclay, in charge of examining proposals for creation of new Inria project-teams.

- C. Marché, member of a hiring committee for an associate professor position in computer science at CNAM, Paris, France. (sep-oct 2016)

10.1.7. Research Administration

- S. Boldo, member of the CCD, *commission consultative des doctorants*.
- S. Boldo, member of the CLFP, *comité local de formation permanente*.
- S. Boldo, scientific head for Saclay for the MECSI group for networking about computer science popularization inside Inria.
- A. Charguéraud is vice-president of *France-ioi*, a non-profit organization in charge of the selection and the training of the French team to the International Olympiads in Informatics (IOI). France-ioi also provides online exercises in programming and algorithmics—in average, over 100,000 such exercises are solved every month on the website.
- A. Charguéraud is a board member of the non-profit organization *Animath*, which aims at developing interest in mathematics among young students.
- A. Charguéraud and G. Melquiond are members of the committee for the monitoring of PhD students (*"commission de suivi des doctorants"*).
- J.-C. Filliâtre is *correcteur au concours d'entrée à l'École Polytechnique et aux ENS* (computer science examiner for the entrance exam at École Polytechnique and Écoles Normales Supérieures) since 2008.
- C. Marché, director of the ProofInUse Joint Laboratory between Inria and AdaCore, http://www.spark-2014.org/proofinuse
- C. Paulin, member of the *"commission consultative de spécialistes de l'université"*, Section 27, University Paris-Sud since April 2010. C. Paulin is the president of this committee since December 2014.
- C. Paulin, chaired the hiring committee for a professor position in computer science at Université Paris-Sud.
- J.-C. Filliâtre, member of the board of GDR GPL, since January 2016.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master Parisien de Recherche en Informatique (MPRI) https://wikimpri.dptinfo.ens-cachan.fr/doku.php: "Proofs of Programs" http://www.lri.fr/~marche/MPRI-2-36-1/ (M2), C. Marché (12h), A. Charguéraud (12h), Université Paris-Diderot, France.

Master: Fondements de l'informatique et ingénierie du logiciel (FIIL) https://www.lri.fr/~conchon/parcours_fiil/: "Software Model Checking" (M2), S. Conchon (9h), "Programmation C++11 avancée" (M2), G. Melquiond (12h), "Vérification déductive de programmes" (M2), A. Paskevich (10.5h), Université Paris-Sud, France.

DUT (Diplôme Universitaire de Technologie): M1101 "Introduction aux systèmes informatiques", A. Paskevich (36h), M3101 "Principes des systèmes d'exploitation", A. Paskevich (58.5h), IUT d'Orsay, Université Paris-Sud, France.

Licence: "Langages de programmation et compilation" (L3), J.-C. Filliâtre (26h), École Normale Supérieure, France.

Licence: "INF411: Les bases de l'algorithmique et de la programmation" (L3), J.-C. Filliâtre (16h), École Polytechnique, France.

Licence: "Programmation fonctionnelle avancée" (L3), S. Conchon (45h), Université Paris-Sud, France.

Licence: "Introduction à la programmation fonctionnelle" (L2), S. Conchon (25h), Université Paris-Sud, France.

10.2.2. Internships

- Raphaël Rieu-Helft (ENS, Paris) is a pre-PhD student doing an internship for 6 months under supervision of C. Marché, G. Melquiond and A. Paskevich. He is working on the design and the formal verification of a library for unbounded integer arithmetic. Why3 is used for formally verifying the functional behaviour of the library operations. Raphaël is also implementing in Why3 a mechanism for extracting code to the C language, in order to obtain a certified code that runs very efficiently.

- Lucas Baudin (ENS, Paris) is a master 1 intern under the supervision of J.-C. Filliâtre between September 2016 and January 2017. He is working on the inference of loop invariants by abstract interpretation in the tool Why3.

- F. Faissole was a master 2 trainee under the supervision of S. Boldo between March and August 2016. He worked on the formal proof of the Lax-Milgram theorem.

10.2.3. Supervision

PhD: L. Gondelmans, "Obtention de programmes corrects par raffinement dans un langage de haut niveau", Université Paris-Saclay & Université Paris-Sud, December 13, 2016, supervised by J.-C. Filliâtre and A. Paskevich.

PhD in progress: M. Clochard, "A unique language for developing programs and prove them at the same time", since Oct. 2013, supervised by C. Marché and A. Paskevich.

PhD in progress: D. Declerck, "Vérification par des techniques de test et model checking de programmes C11", since Sep. 2014, supervised by F. Zaïdi (LRI) and S. Conchon.

PhD in progress: M. Roux, "Model Checking de systèmes paramétrés et temporisés", since Sep. 2015, supervised by Sylvain Conchon.

PhD in progress: M. Pereira, "A Verified Graph Library. Tools and techniques for the verification of modular higher-order programs, with extraction", since May 2015, supervised by J.-C. Filliâtre.

PhD in progress: A. Coquereau, "[ErgoFast] Amélioration de performances pour le solveur SMT Alt-Ergo : conception d'outils d'analyse, optimisations et structures de données efficaces pour OCaml", since Sep. 2015, supervised by Sylvain Conchon, Fabrice Le Fessant et Michel Mauny.

PhD in progress: F. Faissole, "Stabilité(s): liens entre l'arithmétique flottante et l'analyse numérique", since Oct 2016, supervised by S. Boldo and A. Chapoutot.

10.2.4. Juries

J.-C. Filliâtre: president of the PhD committee of A. Djoudi, "Analyse statique au niveau binaire", Université Paris Saclay, France, December 2016.

S. Conchon: examiner of the PhD of M. Morterol, "Méthodes avancées de raisonnement en logique propositionnelle : application aux réseaux métaboliques", Université Paris-Saclay, December 2016.

S. Conchon: reviewer of the PhD of A. Blanchard, "Aide à la vérification de programmes concurrents par transformation de code et de spécifications", Université d'Orléans, December 2016.

S. Conchon: reviewer of the HDR of X. Thirioux, "Verifying Embedded Systems", Institut National Polytechnique de Toulouse, September 2016.

S. Conchon : examiner of the PhD of L. Cabaret, "Algorithmes d'étiquetage en composantes connexes efficaces pour architectures hautes performances", September 2016.

10.3. Popularization

- A. Charguéraud is one of the three organizers of the *Concours Castor informatique*http://castor-informatique.fr/. The purpose of the Concours Castor in to introduce pupils (from *CM1* to *Terminale*) to computer sciences. 475,000 teenagers played with the interactive exercises in November 2016.
- S. Boldo is a speaker for a MOOC for computer science teachers. She was also invited to Poitiers in November 2016 to discuss with teachers and present this MOOC.
- S. Boldo was invited to a panel about teaching computer science before university in Besançon in June 2016 during the GDR GPL days.
- During the "Fête de la science" on October 14th to 16th, S. Boldo gave several talks about computer arithmetic to teenagers and F. Faissole run a stand about an introduction to programming with robots.
- S. Boldo and F. Voisin did an introduction to computer science with an activity on computer hardware as a 1-hour extracurricular activity in schools for pupils in *CM1-CM2* on October 4th.
- S. Boldo gave a talk during a girls & maths weekend on November 22nd. See http://www.animath.fr/spip.php?article2897&lang=fr.

11. Bibliography

Major publications by the team in recent years

[1] J. C. BLANCHETTE, A. PASKEVICH.*TFF1: The TPTP typed first-order form with rank-1 polymorphism*, in "24th International Conference on Automated Deduction (CADE-24)", Lake Placid, USA, Lecture Notes in Artificial Intelligence, Springer, June 2013, vol. 7898, http://hal.inria.fr/hal-00825086.

[2] F. BOBOT, S. CONCHON, É. CONTEJEAN, M. IGUERNELALA, A. MAHBOUBI, A. MEBSOUT, G. MELQUIOND.*A Simplex-Based Extension of Fourier-Motzkin for Solving Linear Integer Arithmetic*, in "IJCAR 2012: Proceedings of the 6th International Joint Conference on Automated Reasoning", Manchester, UK, B. GRAMLICH, D. MILLER, U. SATTLER (editors), Lecture Notes in Computer Science, Springer, June 2012, vol. 7364, p. 67–81, http://hal.inria.fr/hal-00687640.

[3] F. BOBOT, J.-C. FILLIÂTRE, C. MARCHÉ, A. PASKEVICH.*Let's Verify This with Why3*, in "International Journal on Software Tools for Technology Transfer (STTT)", 2015, vol. 17, n° 6, p. 709–727, http://hal.inria.fr/hal-00967132/en.

[4] S. BOLDO, F. CLÉMENT, J.-C. FILLIÂTRE, M. MAYERO, G. MELQUIOND, P. WEIS.*Wave Equation Numerical Resolution: a Comprehensive Mechanized Proof of a C Program*, in "Journal of Automated Reasoning", April 2013, vol. 50, n° 4, p. 423–456, http://hal.inria.fr/hal-00649240/en/.

[5] S. BOLDO, G. MELQUIOND.*Flocq: A Unified Library for Proving Floating-point Algorithms in Coq*, in "Proceedings of the 20th IEEE Symposium on Computer Arithmetic", Tübingen, Germany, E. ANTELO, D. HOUGH, P. IENNE (editors), 2011, p. 243–252, http://hal.archives-ouvertes.fr/inria-00534854/.

[6] A. CHARGUÉRAUD.*Pretty-Big-Step Semantics*, in "Proceedings of the 22nd European Symposium on Programming", M. FELLEISEN, P. GARDNER (editors), Lecture Notes in Computer Science, Springer, March 2013, vol. 7792, p. 41–60, http://hal.inria.fr/hal-00798227.

[7] S. CONCHON, A. GOEL, S. KRSTIĆ, A. MEBSOUT, F. ZAÏDI.*Cubicle: A Parallel SMT-based Model Checker for Parameterized Systems*, in "CAV 2012: Proceedings of the 24th International Conference on Computer Aided Verification", Berkeley, California, USA, M. PARTHASARATHY, S. A. SESHIA (editors), Lecture Notes in Computer Science, Springer, July 2012, vol. 7358, http://hal.archives-ouvertes.fr/hal-00799272.

[8] J.-C. FILLIâTRE, L. GONDELMAN, A. PASKEVICH.*The Spirit of Ghost Code*, in "26th International Conference on Computer Aided Verification", Vienna, Austria, A. BIERE, R. BLOEM (editors), Lecture Notes in Computer Science, Springer, July 2014, vol. 8859, p. 1–16, http://hal.archives-ouvertes.fr/hal-00873187/en/.

[9] C. MARCHÉ.*Verification of the Functional Behavior of a Floating-Point Program: an Industrial Case Study*, in "Science of Computer Programming", March 2014, vol. 96, n° 3, p. 279–296, http://hal.inria.fr/hal-00967124/en.

[10] G. MELQUIOND.*Proving bounds on real-valued functions with computations*, in "Proceedings of the 4th International Joint Conference on Automated Reasoning", Sydney, Australia, A. ARMANDO, P. BAUMGARTNER, G. DOWEK (editors), Lecture Notes in Artificial Intelligence, 2008, vol. 5195, p. 2–17.

Publications of the year

Articles in International Peer-Reviewed Journal

[11] U. A. ACAR, A. CHARGUÉRAUD, M. RAINEY.*Oracle-Guided Scheduling for Controlling Granularity in Implicitly Parallel Languages*, in "Journal of Functional Programming", November 2016, vol. 26 [*DOI :* 10.1017/S0956796816000101], https://hal.inria.fr/hal-01409069.

[12] S. BOLDO, C. LELAY, G. MELQUIOND.*Formalization of Real Analysis: A Survey of Proof Assistants and Libraries*, in "Mathematical Structures in Computer Science", October 2016, vol. 26, n° 7, p. 1196-1233 [*DOI :* 10.1017/S0960129514000437], https://hal.inria.fr/hal-00806920.

[13] C. DROSS, S. CONCHON, J. KANIG, A. PASKEVICH.*Adding Decision Procedures to SMT Solvers using Axioms with Triggers*, in "Journal of Automated Reasoning", 2016, vol. 56, n° 4, p. 387-457 [*DOI :* 10.1007/s10817-015-9352-2], https://hal.archives-ouvertes.fr/hal-01221066.

[14] J.-C. FILLIâTRE, L. GONDELMAN, A. PASKEVICH.*The Spirit of Ghost Code*, in "Formal Methods in System Design", June 2016, vol. 48, n° 3, p. 152-174, Extended version of https://hal.inria.fr/hal-00873187 [*DOI :* 10.1007/s10703-016-0243-x], https://hal.archives-ouvertes.fr/hal-01396864.

[15] É. MARTIN-DOREL, G. MELQUIOND.*Proving Tight Bounds on Univariate Expressions with Elementary Functions in Coq*, in "Journal of Automated Reasoning", October 2016, vol. 57, n° 3, p. 187-217 [*DOI :* 10.1007/s10817-015-9350-4], https://hal.inria.fr/hal-01086460.

International Conferences with Proceedings

[16] U. A. ACAR, A. CHARGUÉRAUD, M. RAINEY, F. SIECZKOWSKI.*Dag-calculus: a calculus for parallel computation*, in "Proceedings of the 21st ACM SIGPLAN International Conference on Functional Programming (ICFP)", Nara, Japan, September 2016, p. 18 - 32 [*DOI :* 10.1145/2951913.2951946], https://hal.inria.fr/hal-01409022.

[17] S. BOLDO.*Computing a correct and tight rounding error bound using rounding-to-nearest*, in "9th International Workshop on Numerical Software Verification", Toronto, Canada, July 2016, https://hal.inria.fr/hal-01377152.

[18] S. BOLDO, F. CLÉMENT, F. FAISSOLE, V. MARTIN, M. MAYERO.*A Coq Formal Proof of the Lax–Milgram theorem*, in "6th ACM SIGPLAN Conference on Certified Programs and Proofs", Paris, France, January 2017, https://hal.inria.fr/hal-01391578.

[19] A. CHARGUÉRAUD.*Higher-order representation predicates in separation logic*, in "Proceedings of the 5th ACM SIGPLAN Conference on Certified Programs and Proofs (CPP)", St. Petersburg, FL, United States, ACM, January 2016, p. 3 - 14 [*DOI :* 10.1145/2854065.2854068], https://hal.inria.fr/hal-01408670.

[20] A. CHARGUÉRAUD, F. POTTIER.*Temporary Read-Only Permissions for Separation Logic*, in "Proceedings of the 26th European Symposium on Programming (ESOP 2017)", Uppsala, Sweden, April 2017, https://hal.inria.fr/hal-01408657.

[21] M. CLOCHARD, L. GONDELMAN, M. PEREIRA.*The Matrix Reproved (Verification Pearl)*, in "VSTTE 2016", Toronto, Canada, July 2016, https://hal.inria.fr/hal-01316902.

[22] J.-C. FILLIÂTRE, M. PEREIRA.*A Modular Way to Reason About Iteration*, in "8th NASA Formal Methods Symposium", Minneapolis, United States, June 2016, https://hal.inria.fr/hal-01281759.

[23] J.-C. FILLIÂTRE, M. PEREIRA.*Producing All Ideals of a Forest, Formally (Verification Pearl)*, in "VSTTE 2016", Toronto, Canada, July 2016, p. 46 - 55 [*DOI :* 10.1007/978-3-319-48869-1_4], https://hal.inria.fr/hal-01316859.

[24] C. FUMEX, C. DROSS, J. GERLACH, C. MARCHÉ.*Specification and Proof of High-Level Functional Properties of Bit-Level Programs*, in "NASA Formal methods", Minneapolis, United States, NASA Formal methods, Springer, June 2016, https://hal.inria.fr/hal-01314876.

[25] D. HAUZAR, C. MARCHÉ, Y. MOY.*Counterexamples from Proof Failures in SPARK*, in "Software Engineering and Formal Methods", Vienna, Austria, Software Engineering and Formal Methods, Springer, July 2016, https://hal.inria.fr/hal-01314885.

[26] N. KOSMATOV, C. MARCHÉ, Y. MOY, J. SIGNOLES.*Static versus Dynamic Verification in Why3, Frama-C and SPARK 2014*, in "7th International Symposium on Leveraging Applications", Corfu, Greece, 7th International Symposium on Leveraging Applications, Springer, October 2016, 16, https://hal.inria.fr/hal-01344110.

[27] A. MAHBOUBI, G. MELQUIOND, T. SIBUT-PINOTE.*Formally Verified Approximations of Definite Integrals*, in "Interactive Theorem Proving", Nancy, France, J. C. BLANCHETTE, S. MERZ (editors), Lecture Notes in Computer Science, August 2016, vol. 9807 [*DOI :* 10.1007/978-3-319-43144-4_17], https://hal.inria.fr/hal-01289616.

National Conferences with Proceeding

[28] R. CHEN, J.-J. LÉVY.*Une preuve formelle de l'algorithme de Tarjan-1972 pour trouver les composantes fortement connexes dans un graphe*, in "JFLA 2017 - Vingt-huitièmes Journées Francophones des Langages Applicatifs", Gourette, France, Vingt-huitièmes Journées Francophones des Langages Applicatifs, January 2017, https://hal.inria.fr/hal-01422215.

[29] M. CLOCHARD.*Preuves taillées en biseau*, in "vingt-huitièmes Journées Francophones des Langages Applicatifs (JFLA)", Gourette, France, January 2017, https://hal.inria.fr/hal-01404935.

[30] M. PEREIRA.*Défonctionnaliser pour prouver*, in "JFLA 2017", Gourette, France, January 2017, https://hal.inria.fr/hal-01378068.

Conferences without Proceedings

[31] S. BOLDO.*Iterators: where folds fail*, in "Workshop on High-Consequence Control Verification", Toronto, Canada, July 2016, https://hal.inria.fr/hal-01377155.

[32] F. FAISSOLE, B. SPITTERS.*Synthetic topology in HoTT for probabilistic programming*, in "The Third International Workshop on Coq for Programming Languages (CoqPL 2017)", Paris, France, January 2017, https://hal.inria.fr/hal-01405762.

[33] J.-C. FILLIÂTRE, M. PEREIRA.*Itérer avec confiance*, in "Journées Francophones des Langages Applicatifs", Saint-Malo, France, January 2016, https://hal.inria.fr/hal-01240891.

Research Reports

[34] R. CHEN, M. CLOCHARD, C. MARCHÉ.*A Formal Proof of a Unix Path Resolution Algorithm*, Inria, December 2016, n^o RR-8987, 27, https://hal.inria.fr/hal-01406848.

[35] D. HAUZAR, C. MARCHÉ, Y. MOY.*Counterexamples from proof failures in the SPARK program verifier*, Inria, February 2016, n^o RR-8854, 22, https://hal.inria.fr/hal-01271174.

Other Publications

[36] S. BOLDO, S. GRAILLAT, J.-M. MULLER.*On the robustness of the 2Sum and Fast2Sum algorithms*, May 2016, working paper or preprint, https://hal-ens-lyon.archives-ouvertes.fr/ensl-01310023.

[37] J.-C. FILLIÂTRE, L. GONDELMAN, A. PASKEVICH.*A Pragmatic Type System for Deductive Verification*, February 2016, working paper or preprint, https://hal.inria.fr/hal-01256434.

References in notes

[38] B. BECKERT, R. HÄHNLE, P. H. SCHMITT (editors). *Verification of Object-Oriented Software: The KeY Approach*, Lecture Notes in Computer Science, Springer, 2007, vol. 4334.

[39] U. A. ACAR, A. CHARGUÉRAUD, M. RAINEY.*Theory and Practice of Chunked Sequences*, in "European Symposium on Algorithms", Wroclaw, Poland, A. SCHULZ, D. WAGNER (editors), Springer, September 2014, vol. Lecture Notes in Computer Science, n° 8737, p. 25–36, https://hal.inria.fr/hal-01087245.

[40] J. B. ALMEIDA, M. BARBOSA, J.-C. FILLIÂTRE, J. S. PINTO, B. VIEIRA.*CAOVerif: An Open-Source Deductive Verification Platform for Cryptographic Software Implementations*, in "Science of Computer Programming", October 2012.

[41] A. AYAD, C. MARCHÉ.*Multi-Prover Verification of Floating-Point Programs*, in "Fifth International Joint Conference on Automated Reasoning", Edinburgh, Scotland, J. GIESL, R. HÄHNLE (editors), Lecture Notes in Artificial Intelligence, Springer, July 2010, vol. 6173, p. 127–141, http://hal.inria.fr/inria-00534333.

[42] D. BAELDE, P. COURTIEU, D. GROSS-AMBLARD, C. PAULIN-MOHRING.*Towards Provably Robust Watermarking*, in "ITP 2012", Lecture Notes in Computer Science, August 2012, vol. 7406, http://hal.inria.fr/hal-00682398.

[43] C. BARRETT, C. TINELLI.*CVC3*, in "19th International Conference on Computer Aided Verification", Berlin, Germany, W. DAMM, H. HERMANNS (editors), Lecture Notes in Computer Science, Springer, July 2007, vol. 4590, p. 298–302.

[44] P. BAUDIN, J.-C. FILLIÂTRE, C. MARCHÉ, B. MONATE, Y. MOY, V. PREVOSTO.*ACSL: ANSI/ISO C Specification Language, version 1.4*, 2009.

[45] P. BEHM, P. BENOIT, A. FAIVRE, J.-M. MEYNADIER.*METEOR : A successful application of B in a large project*, in "Proceedings of FM'99: World Congress on Formal Methods", J. M. WING, J. WOODCOCK, J. DAVIES (editors), Lecture Notes in Computer Science (Springer-Verlag), Springer Verlag, September 1999, p. 369–387.

[46] F. BOBOT, S. CONCHON, É. CONTEJEAN, M. IGUERNELALA, S. LESCUYER, A. MEBSOUT.*The Alt-Ergo Automated Theorem Prover*, 2008.

[47] F. BOBOT, J.-C. FILLIÂTRE.*Separation Predicates: a Taste of Separation Logic in First-Order Logic*, in "14th International Conference on Formal Ingineering Methods (ICFEM)", Kyoto, Japan, Lecture Notes in Computer Science, Springer, November 2012, vol. 7635, http://hal.inria.fr/hal-00825088.

[48] F. BOBOT, J.-C. FILLIÂTRE, C. MARCHÉ, G. MELQUIOND, A. PASKEVICH.*Preserving User Proofs Across Specification Changes*, in "Verified Software: Theories, Tools, Experiments (5th International Conference VSTTE)", Atherton, USA, E. COHEN, A. RYBALCHENKO (editors), Lecture Notes in Computer Science, Springer, May 2013, vol. 8164, p. 191–201, http://hal.inria.fr/hal-00875395.

[49] F. BOBOT, J.-C. FILLIÂTRE, C. MARCHÉ, G. MELQUIOND, A. PASKEVICH.*The Why3 platform, version 0.81*, version 0.81, LRI, CNRS & Univ. Paris-Sud & Inria Saclay, March 2013, http://hal.inria.fr/hal-00822856/.

[50] M. BODIN, A. CHARGUÉRAUD, D. FILARETTI, P. GARDNER, S. MAFFEIS, D. NAUDZIUNIENE, A. SCHMITT, G. SMITH.*A Trusted Mechanised JavaScript Specification*, in "Proceedings of the 41st ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages", San Diego, USA, ACM Press, January 2014, http://hal.inria.fr/hal-00910135.

[51] S. BOLDO.*How to Compute the Area of a Triangle: a Formal Revisit*, in "Proceedings of the 21th IEEE Symposium on Computer Arithmetic", Austin, Texas, USA, 2013, http://hal.inria.fr/hal-00790071.

[52] S. BOLDO.*Deductive Formal Verification: How To Make Your Floating-Point Programs Behave*, Université Paris-Sud, October 2014, Thèse d'habilitation, https://hal.inria.fr/tel-01089643.

[53] S. BOLDO.*Formal verification of tricky numerical computations*, in "16th GAMM-IMACS International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics", Würzburg, Germany, September 2014, https://hal.inria.fr/hal-01088692.

[54] S. BOLDO, F. CLÉMENT, J.-C. FILLIÂTRE, M. MAYERO, G. MELQUIOND, P. WEIS.*Formal Proof of a Wave Equation Resolution Scheme: the Method Error*, in "Proceedings of the First Interactive Theorem Proving Conference", Edinburgh, Scotland, M. KAUFMANN, L. C. PAULSON (editors), LNCS, Springer, July 2010, vol. 6172, p. 147–162, http://hal.inria.fr/inria-00450789/.

[55] S. BOLDO, F. CLÉMENT, J.-C. FILLIÂTRE, M. MAYERO, G. MELQUIOND, P. WEIS.*Trusting Computations: a Mechanized Proof from Partial Differential Equations to Actual Program*, in "Computers and Mathematics with Applications", 2014, vol. 68, nᴼ 3, p. 325–352, http://hal.inria.fr/hal-00769201.

[56] S. BOLDO, J.-C. FILLIÂTRE, G. MELQUIOND.*Combining Coq and Gappa for Certifying Floating-Point Programs*, in "16th Symposium on the Integration of Symbolic Computation and Mechanised Reasoning", Grand Bend, Canada, Lecture Notes in Artificial Intelligence, Springer, July 2009, vol. 5625, p. 59–74.

[57] S. BOLDO, J.-H. JOURDAN, X. LEROY, G. MELQUIOND.*A Formally-Verified C Compiler Supporting Floating-Point Arithmetic*, in "Proceedings of the 21th IEEE Symposium on Computer Arithmetic", Austin, Texas, USA, 2013, http://hal.inria.fr/hal-00743090.

[58] S. BOLDO, J.-H. JOURDAN, X. LEROY, G. MELQUIOND.*Verified Compilation of Floating-Point Computations*, in "Journal of Automated Reasoning", February 2015, vol. 54, nᴼ 2, p. 135-163, https://hal.inria.fr/hal-00862689.

[59] S. BOLDO, C. LELAY, G. MELQUIOND.*Improving Real Analysis in Coq: a User-Friendly Approach to Integrals and Derivatives*, in "Proceedings of the Second International Conference on Certified Programs and Proofs", Kyoto, Japan, C. HAWBLITZEL, D. MILLER (editors), Lecture Notes in Computer Science, December 2012, vol. 7679, p. 289–304, http://hal.inria.fr/hal-00712938.

[60] S. BOLDO, C. LELAY, G. MELQUIOND.*Coquelicot: A User-Friendly Library of Real Analysis for Coq*, in "Mathematics in Computer Science", June 2015, vol. 9, nᴼ 1, p. 41-62, http://hal.inria.fr/hal-00860648.

[61] S. BOLDO, C. MARCHÉ.*Formal verification of numerical programs: from C annotated programs to mechanical proofs*, in "Mathematics in Computer Science", 2011, vol. 5, p. 377–393, http://hal.inria.fr/hal-00777605.

[62] S. BOLDO, T. M. T. NGUYEN.*Proofs of numerical programs when the compiler optimizes*, in "Innovations in Systems and Software Engineering", 2011, vol. 7, p. 151–160, http://hal.inria.fr/hal-00777639.

[63] T. BORMER, M. BROCKSCHMIDT, D. DISTEFANO, G. ERNST, J.-C. FILLIÂTRE, R. GRIGORE, M. HUISMAN, V. KLEBANOV, C. MARCHÉ, R. MONAHAN, W. MOSTOWSKI, N. POLIKARPOVA, C. SCHEBEN, G. SCHELLHORN, B. TOFAN, J. TSCHANNEN, M. ULBRICH.*The COST IC0701 Verification Competition 2011*, in "Formal Verification of Object-Oriented Software, Revised Selected Papers Presented at the International Conference, FoVeOOS 2011", B. BECKERT, F. DAMIANI, D. GUROV (editors), Lecture Notes in Computer Science, Springer, 2012, vol. 7421, http://hal.inria.fr/hal-00789525.

[64] L. BURDY, Y. CHEON, D. R. COK, M. D. ERNST, J. R. KINIRY, G. T. LEAVENS, K. R. M. LEINO, E. POLL.*An overview of JML tools and applications*, in "International Journal on Software Tools for Technology Transfer (STTT)", June 2005, vol. 7, nᴼ 3, p. 212–232.

[65] R. CHAPMAN, F. SCHANDA.*Are We There Yet? 20 Years of Industrial Theorem Proving with SPARK*, in "Interactive Theorem Proving - 5th International Conference, ITP 2014, Held as Part of the Vienna Summer of Logic, VSL 2014, Vienna, Austria, July 14-17, 2014. Proceedings", G. KLEIN, R. GAMBOA (editors), Lecture Notes in Computer Science, Springer, 2014, vol. 8558, p. 17–26.

[66] M. CLOCHARD.*Automatically verified implementation of data structures based on AVL trees*, in "6th Working Conference on Verified Software: Theories, Tools and Experiments (VSTTE)", Vienna, Austria, D. GIAN-

NAKOPOULOU, D. KROENING (editors), Lecture Notes in Computer Science, Springer, July 2014, vol. 8471, p. 167–180, http://hal.inria.fr/hal-01067217.

[67] M. CLOCHARD, J.-C. FILLIÂTRE, C. MARCHÉ, A. PASKEVICH.*Formalizing Semantics with an Automatic Program Verifier*, in "6th Working Conference on Verified Software: Theories, Tools and Experiments (VSTTE)", Vienna, Austria, D. GIANNAKOPOULOU, D. KROENING (editors), Lecture Notes in Computer Science, Springer, July 2014, vol. 8471, p. 37–51, http://hal.inria.fr/hal-01067197.

[68] M. CLOCHARD, L. GONDELMAN.*Double WP: vers une preuve automatique d'un compilateur*, in "Vingt-sixièmes Journées Francophones des Langages Applicatifs", Val d'Ajol, France, January 2015, https://hal.inria.fr/hal-01094488.

[69] M. CLOCHARD, C. MARCHÉ, A. PASKEVICH.*Verified Programs with Binders*, in "Programming Languages meets Program Verification (PLPV)", ACM Press, 2014, http://hal.inria.fr/hal-00913431.

[70] S. CONCHON.*SMT Techniques and their Applications: from Alt-Ergo to Cubicle*, Université Paris-Sud, December 2012, In English, http://www.lri.fr/~conchon/publis/conchonHDR.pdf, Thèse d'habilitation.

[71] S. CONCHON, É. CONTEJEAN, M. IGUERNELALA.*Canonized Rewriting and Ground AC Completion Modulo Shostak Theories*, in "Tools and Algorithms for the Construction and Analysis of Systems", Saarbrücken, Germany, P. A. ABDULLA, K. R. M. LEINO (editors), Lecture Notes in Computer Science, Springer, April 2011, vol. 6605, p. 45-59, http://hal.inria.fr/hal-00777663.

[72] S. CONCHON, É. CONTEJEAN, M. IGUERNELALA.*Canonized Rewriting and Ground AC Completion Modulo Shostak Theories : Design and Implementation*, in "Logical Methods in Computer Science", September 2012, vol. 8, n° 3, p. 1–29, hal.inria.fr/hal-00798082.

[73] S. CONCHON, D. DECLERCK, L. MARANGET, A. MEBSOUT.*Vérification de programmes C concurrents avec Cubicle : Enforcer les barrières*, in "Vingt-cinquièmes Journées Francophones des Langages Applicatifs", Fréjus, France, January 2014, https://hal.inria.fr/hal-01088655.

[74] S. CONCHON, A. GOEL, S. KRSTIĆ, A. MEBSOUT, F. ZAÏDI.*Invariants for Finite Instances and Beyond*, in "FMCAD", Portland, Oregon, États-Unis, October 2013, p. 61–68, http://hal.archives-ouvertes.fr/hal-00924640.

[75] S. CONCHON, M. IGUERNELALA.*Tuning the Alt-Ergo SMT Solver for B Proof Obligations*, in "Abstract State Machines, Alloy, B, VDM, and Z (ABZ)", Toulouse, France, Lecture Notes in Computer Science, Springer, June 2014, vol. 8477, p. 294–297, https://hal.inria.fr/hal-01093000.

[76] S. CONCHON, M. IGUERNELALA, A. MEBSOUT.*A Collaborative Framework for Non-Linear Integer Arithmetic Reasoning in Alt-Ergo*, 2013, https://hal.archives-ouvertes.fr/hal-00924646.

[77] S. CONCHON, A. MEBSOUT, F. ZAÏDI.*Vérification de systèmes paramétrés avec Cubicle*, in "Vingt-quatrièmes Journées Francophones des Langages Applicatifs", Aussois, France, February 2013, http://hal.inria.fr/hal-00778832.

[78] S. CONCHON, G. MELQUIOND, C. ROUX, M. IGUERNELALA.*Built-in Treatment of an Axiomatic Floating-Point Theory for SMT Solvers*, in "SMT workshop", Manchester, UK, P. FONTAINE, A. GOEL (editors), LORIA, 2012, p. 12–21.

[79] M. DAHLWEID, M. MOSKAL, T. SANTEN, S. TOBIES, W. SCHULTE.*VCC: Contract-based modular verification of concurrent C*, in "31st International Conference on Software Engineering, ICSE 2009, May 16-24, 2009, Vancouver, Canada, Companion Volume", IEEE Comp. Soc. Press, 2009, p. 429-430.

[80] D. DELAHAYE, C. DUBOIS, C. MARCHÉ, D. MENTRÉ.*The BWare Project: Building a Proof Platform for the Automated Verification of B Proof Obligations*, in "Abstract State Machines, Alloy, B, VDM, and Z (ABZ)", Toulouse, France, Lecture Notes in Computer Science, Springer, June 2014, vol. 8477, p. 290–293, http://hal.inria.fr/hal-00998092/en/.

[81] D. DELAHAYE, C. MARCHÉ, D. MENTRÉ.*Le projet BWare : une plate-forme pour la vérification automatique d'obligations de preuve B*, in "Approches Formelles dans l'Assistance au Développement de Logiciels (AFADL)", Paris, France, EasyChair, June 2014, http://hal.inria.fr/hal-00998094/en/.

[82] C. DROSS, S. CONCHON, J. KANIG, A. PASKEVICH.*Reasoning with Triggers*, in "SMT workshop", Manchester, UK, P. FONTAINE, A. GOEL (editors), LORIA, 2012.

[83] C. DROSS.*Generic Decision Procedures for Axiomatic First-Order Theories*, Université Paris-Sud, April 2014, http://tel.archives-ouvertes.fr/tel-01002190.

[84] J.-C. FILLIÂTRE.*Combining Interactive and Automated Theorem Proving in Why3 (invited talk)*, in "Automation in Proof Assistants 2012", Tallinn, Estonia, K. HELJANKO, H. HERBELIN (editors), April 2012.

[85] J.-C. FILLIÂTRE.*Combining Interactive and Automated Theorem Proving using Why3 (invited tutorial)*, in "Second International Workshop on Intermediate Verification Languages (BOOGIE 2012)", Berkeley, California, USA, Z. RAKAMARIĆ (editor), July 2012.

[86] J.-C. FILLIÂTRE.*Verifying Two Lines of C with Why3: an Exercise in Program Verification*, in "Verified Software: Theories, Tools, Experiments (4th International Conference VSTTE)", Philadelphia, USA, R. JOSHI, P. MÜLLER, A. PODELSKI (editors), Lecture Notes in Computer Science, Springer, January 2012, vol. 7152, p. 83–97.

[87] J.-C. FILLIÂTRE.*Deductive Program Verification*, in "Programming Languages Mentoring Workshop (PLMW 2013)", Rome, Italy, N. FOSTER, P. GARDNER, A. SCHMITT, G. SMITH, P. THIEMAN, T. WRIGSTAD (editors), January 2013, http://hal.inria.fr/hal-00799190.

[88] J.-C. FILLIÂTRE.*One Logic To Use Them All*, in "24th International Conference on Automated Deduction (CADE-24)", Lake Placid, USA, Lecture Notes in Artificial Intelligence, Springer, June 2013, vol. 7898, p. 1–20, http://hal.inria.fr/hal-00809651/en/.

[89] J.-C. FILLIÂTRE, A. PASKEVICH.*Why3 — Where Programs Meet Provers*, in "Proceedings of the 22nd European Symposium on Programming", M. FELLEISEN, P. GARDNER (editors), Lecture Notes in Computer Science, Springer, March 2013, vol. 7792, p. 125–128, http://hal.inria.fr/hal-00789533.

[90] J.-C. FILLIÂTRE, A. PASKEVICH, A. STUMP.*The 2nd Verified Software Competition: Experience Report*, in "COMPARE2012: 1st International Workshop on Comparative Empirical Evaluation of Reasoning Systems", Manchester, UK, V. KLEBANOV, S. GREBING (editors), EasyChair, June 2012, http://hal.inria.fr/hal-00798777.

[91] P. HERMS.*Certification of a Tool Chain for Deductive Program Verification*, Université Paris-Sud, January 2013, http://tel.archives-ouvertes.fr/tel-00789543.

[92] P. HERMS, C. MARCHÉ, B. MONATE.*A Certified Multi-prover Verification Condition Generator*, in "Verified Software: Theories, Tools, Experiments (4th International Conference VSTTE)", Philadelphia, USA, R. JOSHI, P. MÜLLER, A. PODELSKI (editors), Lecture Notes in Computer Science, Springer, January 2012, vol. 7152, p. 2–17, http://hal.inria.fr/hal-00639977.

[93] M. IGUERNELALA.*Strengthening the Heart of an SMT-Solver: Design and Implementation of Efficient Decision Procedures*, Université Paris-Sud, June 2013, http://tel.archives-ouvertes.fr/tel-00842555.

[94] D. ISHII, G. MELQUIOND, S. NAKAJIMA.*Inductive Verification of Hybrid Automata with Strongest Post-condition Calculus*, in "Proceedings of the 10th Conference on Integrated Formal Methods", Turku, Finland, E. B. JOHNSEN, L. PETRE (editors), Lecture Notes in Computer Science, 2013, vol. 7940, p. 139–153, http://hal.inria.fr/hal-00806701.

[95] J. KANIG, E. SCHONBERG, C. DROSS.*Hi-Lite: the convergence of compiler technology and program verification*, in "Proceedings of the 2012 ACM Conference on High Integrity Language Technology, HILT '12", Boston, USA, B. BROSGOL, J. BOLENG, S. T. TAFT (editors), ACM Press, 2012, p. 27–34.

[96] G. KLEIN, J. ANDRONICK, K. ELPHINSTONE, G. HEISER, D. COCK, P. DERRIN, D. ELKADUWE, K. ENGELHARDT, R. KOLANSKI, M. NORRISH, T. SEWELL, H. TUCH, S. WINWOOD.*seL4: Formal verification of an OS kernel*, in "Communications of the ACM", June 2010, vol. 53, n° 6, p. 107–115.

[97] C. LELAY.*A New Formalization of Power Series in Coq*, in "5th Coq Workshop", Rennes, France, July 2013, p. 1–2, http://hal.inria.fr/hal-00880212.

[98] C. LELAY.*Coq passe le bac*, in "JFLA - Journées francophones des langages applicatifs", Fréjus, France, January 2014.

[99] C. LELAY, G. MELQUIOND.*Différentiabilité et intégrabilité en Coq. Application à la formule de d'Alembert*, in "Vingt-troisièmes Journées Francophones des Langages Applicatifs", Carnac, France, February 2012, http://hal.inria.fr/hal-00642206/fr/.

[100] X. LEROY.*A formally verified compiler back-end*, in "Journal of Automated Reasoning", 2009, vol. 43, n° 4, p. 363–446, http://hal.inria.fr/inria-00360768/en/.

[101] C. B. LOURENÇO, S. LAMRAOUI, S. NAKAJIMA, J. S. PINTO.*Studying Verification Conditions for Imperative Programs*, in "Electronic Communication of the European Association of Software Science and Technology", 2015, vol. 72.

[102] C. MARCHÉ, A. TAFAT.*Weakest Precondition Calculus, revisited using Why3*, Inria, December 2012, n^o RR-8185, http://hal.inria.fr/hal-00766171.

[103] C. MARCHÉ, A. TAFAT.*Calcul de plus faible précondition, revisité en Why3*, in "Vingt-quatrièmes Journées Francophones des Langages Applicatifs", Aussois, France, February 2013, http://hal.inria.fr/hal-00778791.

[104] C. MARCHÉ.*Verification of the Functional Behavior of a Floating-Point Program: an Industrial Case Study*, in "Science of Computer Programming", March 2014, vol. 96, nᵒ 3, p. 279–296, http://hal.inria.fr/hal-00967124/en.

[105] É. MARTIN-DOREL, G. MELQUIOND, J.-M. MULLER.*Some Issues related to Double Roundings*, in "BIT Numerical Mathematics", 2013, vol. 53, nᵒ 4, p. 897–924, http://hal-ens-lyon.archives-ouvertes.fr/ensl-00644408.

[106] A. MEBSOUT.*Invariants inference for model checking of parameterized systems*, Université Paris-Sud, September 2014, https://tel.archives-ouvertes.fr/tel-01073980.

[107] G. MELQUIOND.*Floating-point arithmetic in the Coq system*, in "Information and Computation", 2012, vol. 216, p. 14–23, http://hal.inria.fr/hal-00797913.

[108] G. MELQUIOND, W. G. NOWAK, P. ZIMMERMANN.*Numerical Approximation of the Masser-Gramain Constant to Four Decimal Digits: delta=1.819...*, in "Mathematics of Computation", 2013, vol. 82, p. 1235–1246, http://hal.inria.fr/hal-00644166/en/.

[109] D. MENTRÉ, C. MARCHÉ, J.-C. FILLIÂTRE, M. ASUKA.*Discharging Proof Obligations from Atelier B using Multiple Automated Provers*, in "ABZ'2012 - 3rd International Conference on Abstract State Machines, Alloy, B and Z", Pisa, Italy, S. REEVES, E. RICCOBENE (editors), Lecture Notes in Computer Science, Springer, June 2012, vol. 7316, p. 238–251, http://hal.inria.fr/hal-00681781/en/.

[110] J.-M. MULLER, N. BRISEBARRE, F. DE DINECHIN, C.-P. JEANNEROD, V. LEFÈVRE, G. MELQUIOND, N. REVOL, D. STEHLÉ, S. TORRES.*Handbook of Floating-Point Arithmetic*, Birkhäuser, 2010.

[111] T. M. T. NGUYEN, C. MARCHÉ.*Hardware-Dependent Proofs of Numerical Programs*, in "Certified Programs and Proofs", J.-P. JOUANNAUD, Z. SHAO (editors), Lecture Notes in Computer Science, Springer, December 2011, p. 314–329, http://hal.inria.fr/hal-00772508.

[112] T. M. T. NGUYEN.*Taking architecture and compiler into account in formal proofs of numerical programs*, Université Paris-Sud, June 2012, http://tel.archives-ouvertes.fr/tel-00710193.

[113] M. NORRISH.*C Formalised in HOL*, University of Cambridge, November 1998.

[114] M. PEREIRA, J.-C. FILLIÂTRE, S. M. DE SOUSA.*ARMY: a Deductive Verification Platform for ARM Programs Using Why3*, in "INForum 2012", September 2012.

[115] P. ROUX.*Formal Proofs of Rounding Error Bounds*, in "Journal of Automated Reasoning", 2015, https://hal.archives-ouvertes.fr/hal-01091189.

[116] N. SCHIRMER.*Verification of Sequential Imperative Programs in Isabelle/HOL*, Technische Universität München, 2006.

[117] A. TAFAT.*Preuves par raffinement de programmes avec pointeurs*, Université Paris-Sud, September 2013, http://tel.archives-ouvertes.fr/tel-00874679.

[118] F. DE DINECHIN, C. LAUTER, G. MELQUIOND.*Certifying the floating-point implementation of an elementary function using Gappa*, in "IEEE Transactions on Computers", 2011, vol. 60, n° 2, p. 242–253, http://hal.inria.fr/inria-00533968/en/.

[119] L. DE MOURA, N. BJØRNER.*Z3, An Efficient SMT Solver*, in "TACAS", Lecture Notes in Computer Science, Springer, 2008, vol. 4963, p. 337–340.

Team TROPICAL

Tropical methods: structures, algorithms and interactions

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

Table of contents

Team TROPICAL

Creation of the Team: 2016 January 01

Keywords:

Computer Science and Digital Science:

2.4. - Verification, reliability, certification
6.2.5. - Numerical Linear Algebra
6.2.6. - Optimization
6.4.1. - Deterministic control
6.4.2. - Stochastic control
7.2. - Discrete mathematics, combinatorics
7.3. - Optimization
7.5. - Geometry, Topology
7.11. - Performance evaluation
7.14. - Game Theory

Other Research Topics and Application Domains:

1.1.10. - Mathematical biology
4.4. - Energy delivery
4.4.1. - Smart grids
9.9. - Risk management

1. Members

Research Scientists

Stéphane Gaubert [Team leader, Inria, Senior Researcher]
Marianne Akian [Inria, Senior Researcher, HDR]
Xavier Allamigeon [Corps des Mines, under secondment, Inria, Researcher]
Jean-Pierre Quadrat [Inria, Emiritus Senior Researcher, until Jun 2016, HDR]
Cormac Walsh [Inria, Researcher]

PhD Students

Vianney Boeuf [Ingénieur du corps des Ponts, ENPC]
Jean-Bernard Eytard [Inria]
Eric Fodjo [I-Fihn Consulting, Consultant]
Antoine Hochart [Hadamard, Ecole polytechnique, until Sep 2016]
Paulin Jacquot [EDF, from Oct 2016]
Mateusz Skomra [EDX, Ecole polytechnique]
Nikolas Stott [Inria]

Post-Doctoral Fellows

Marie Maccaig [Fondation Mathématique Jacques Hadamard, Ecole polytechnique]
Adi Niv [Inria, until Sep 2016]

Visiting Scientist

Ricardo Katz [CIFASIS, CONICET, May-June 2016]

Administrative Assistants

Jessica Gameiro [Inria, until Sep 2016]

Corinne Petitot [Inria, from Sep 2016]
Other
Shahrzad Heydarshahi [Inria, Research internship, June - Sep 2016]

2. Overall Objectives

2.1. Introduction

The project develops tropical methods motivated by applications arising in decision theory (deterministic and stochastic optimal control, game theory, optimization and operations research), in the analysis or control of classes of dynamical systems (including timed discrete event systems and positive systems), in the verification of programs and systems, and in the development of numerical algorithms. Tropical algebra tools are used in interaction with various methods, coming from convex analysis, Hamilton–Jacobi partial differential equations, metric geometry, Perron-Frobenius and nonlinear fixed-point theories, combinatorics or algorithmic complexity. The emphasis of the project is on mathematical modelling and computational aspects.

The subtitle of the *Tropical* project, namely, "structures, algorithms, and interactions", refers to the spirit of our research, including a methodological component, computational aspects, and finally interactions with other scientific fields or real world applications, in particular through mathematical modelling.

2.2. Scientific context

Tropical algebra, geometry, and analysis have enjoyed spectacular development in recent years. Tropical structures initially arose to solve problems in performance evaluation of discrete event systems [66], combinatorial optimization [68], or automata theory [106]. They also arose in mathematical physics and asymptotic analysis [97], [94]. More recently, these structures have appeared in several areas of pure mathematics, in particular in the study of combinatorial aspects of algebraic geometry [87], [115], [108], [91], in algebraic combinatorics [82], and in arithmetics [72]. Also, further applications of tropical methods have appeared, including optimal control [100], program invariant computation [63] and timed systems verification [96], and zero-sum games [2].

The term 'tropical' generally refers to algebraic structures in which the laws originate from optimization processes. The prototypical tropical structure is the max-plus semifield, consisting of the real numbers, equipped with the maximum, thought of as an additive law, and the addition, thought of as a multiplicative law. Tropical objects appear as limits of classical objects along certain deformations ("log-limits sets" of Bergman, "Maslov dequantization", or "Viro deformation"). For this reason, the introduction of tropical tools often yields new insights into old familiar problems, leading either to counterexamples or to new methods and results; see for instance [115], [102]. In some applications, like optimal control, discrete event systems, or static analysis of programs, tropical objects do not appear through a limit procedure, but more directly as a modelling or computation/analysis tool; see for instance [112], [66], [89], [69].

Tropical methods are linked to the fields of positive systems and of metric geometry [105], [11]. Indeed, tropically linear maps are monotone (a.k.a. order-preserving). They are also nonexpansive in certain natural metrics (sup-norm, Hopf oscillation, Hilbert's projective metric, ...). In this way, tropical dynamical systems appear to be special cases of nonexpansive, positive, or monotone dynamical systems, which are studied as part of linear and non-linear Perron-Frobenius theory [95], [14]. Such dynamical systems are of fundamental importance in the study of repeated games [101]. Monotonicity properties are also essential in the understanding of the fixed points problems which determine program invariants by abstract interpretation [73]. The latter problems are actually somehow similar to the ones arising in the study of zero-sum games; see [7]. Moreover, positivity or monotonicity methods are useful in population dynamics, either in a discrete space setting [113] or in a PDE setting [67]. In such cases, solving tropical problems often leads to solutions or combinatorial insights on classical problems involving positivity conditions (e.g., finding equilibria of dynamical systems with nonnegative coordinates, understanding the qualitative and quantitative behavior of growth rates / Floquet eigenvalues [9], etc). Other applications of Perron-Frobenius theory originate from quantum information and control [107], [111].

3. Research Program

3.1. Optimal control and zero-sum games

The dynamic programming approach allows one to analyze one or two-player dynamic decision problems by means of operators, or partial differential equations (Hamilton–Jacobi or Isaacs PDEs), describing the time evolution of the value function, i.e., of the optimal reward of one player, thought of as a function of the initial state and of the horizon. We work especially with problems having long or infinite horizon, modelled by stopping problems, or ergodic problems in which one optimizes a mean payoff per time unit. The determination of optimal strategies reduces to solving nonlinear fixed point equations, which are obtained either directly from discrete models, or after a discretization of a PDE.

The geometry of solutions of optimal control and game problems Basic questions include, especially for stationary or ergodic problems, the understanding of existence and uniqueness conditions for the solutions of dynamic programming equations, for instance in terms of controllability or ergodicity properties, and more generally the understanding of the structure of the full set of solutions of stationary Hamilton–Jacobi PDEs and of the set of optimal strategies. These issues are already challenging in the one-player deterministic case, which is an application of choice of tropical methods, since the Lax-Oleinik semigroup, i.e., the evolution semigroup of the Hamilton-Jacobi PDE, is a linear operator in the tropical sense. Recent progress in the deterministic case has been made by combining dynamical systems and PDE techniques (weak KAM theory [79]), and also using metric geometry ideas (abstract boundaries can be used to represent the sets of solutions [90], [4]). The two player case is challenging, owing to the lack of compactness of the analogue of the Lax-Oleinik semigroup and to a richer geometry. The conditions of solvability of ergodic problems for games (for instance, solvability of ergodic Isaacs PDEs), and the representation of solutions are only understood in special cases, for instance in the finite state space case, through tropical geometry and non-linear Perron-Frobenius methods [54],[47], [14].

Algorithmic aspects: from combinatorial algorithms to the attenuation of the curse of dimensionality Our general goal is to push the limits of solvable models by means of fast algorithms adapted to large scale instances. Such instances arise from discrete problems, in which the state space may so large that it is only accessible through local oracles (for instance, in some web ranking applications, the number of states may be the number of web pages) [80]. They also arise from the discretization of PDEs, in which the number of states grows exponentially with the number of degrees of freedom, according to the "curse of dimensionality". A first line of research is the development of *new approximation methods for the value function*. So far, classical approximations by linear combinations have been used, as well as approximation by suprema of linear or quadratic forms, which have been introduced in the setting of dual dynamic programming and of the so called "max-plus basis methods" [81]. We believe that more concise or more accurate approximations may be obtained by unifying these methods. Also, some max-plus basis methods have been shown to *attenuate the curse of dimensionality* for very special problems (for instance involving switching) [98], [84]. This suggests that the complexity of control or games problems may be measured by more subtle quantities that the mere number of states, for instance, by some forms of metric entropy (for example, certain large scale problems have a low complexity owing to the presence of decomposition properties, "highway hierarchies", etc.). A second line of of our research is the development of *combinatorial algorithms*, to solve large scale zero-sum two-player problems with discrete state space. This is related to current open problems in algorithmic game theory. In particular, the existence of polynomial-time algorithms for games with ergodic payment is an open question. See e.g. [5] for a polynomial time average complexity result derived by tropical methods. The two lines of research are related, as the understanding of the geometry of solutions allows to develop better approximation or combinatorial algorithms.

3.2. Non-linear Perron-Frobenius theory, nonexpansive mappings and metric geometry

Several applications (including population dynamics [9] and discrete event systems [66], [71], [62]) lead to studying classes of dynamical systems with remarkable properties: preserving a cone, preserving an order, or being nonexpansive in a metric. These can be studied by techniques of non-linear Perron-Frobenius theory [14] or metric geometry [10]. Basic issues concern the existence and computation of the "escape rate" (which determines the throughput, the growth rate of the population), the characterizations of stationary regimes (non-linear fixed points), or the study of the dynamical properties (convergence to periodic orbits). Nonexpansive mappings also play a key role in the "operator approach" to zero-sum games, since the one-day operators of games are nonexpansive in several metrics, see [8].

3.3. Tropical algebra and convex geometry

The different applications mentioned in the other sections lead us to develop some basic research on tropical algebraic structures and in convex and discrete geometry, looking at objects or problems with a "piecewise-linear" structure. These include the geometry and algorithmics of tropical convex sets [64], [56], tropical semialgebraic sets [49], the study of semi-modules (analogues of vector spaces when the base field is replaced by a semi-field), the study of systems of equations linear in the tropical sense, investigating for instance the analogues of the notions of rank, the analogue of the eigenproblems [15], and more generally of systems of tropical polynomial equations. Our research also builds on, and concern, classical convex and discrete geometry methods.

3.4. Tropical methods applied to optimization, perturbation theory and matrix analysis

Tropical algebraic objects appear as a deformation of classical objects thought various asymptotic procedures. A familiar example is the rule of asymptotic calculus,

$$e^{-a/\epsilon} + e^{-b/\epsilon} \asymp e^{-\min(a,b)/\epsilon} \ , \qquad e^{-a/\epsilon} \times e^{-b/\epsilon} = e^{-(a+b)/\epsilon} \ , \tag{16}$$

when $\epsilon \to 0^+$. Deformations of this kind have been studied in different contexts: large deviations, zero-temperature limits, Maslov's "dequantization method" [97], non-archimedean valuations, log-limit sets and Viro's patchworking method [116], etc.

This entails a relation between classical algorithmic problems and tropical algorithmic problems, one may first solve the $\epsilon = 0$ case (non-archimedean problem), which is sometimes easier, and then use the information gotten in this way to solve the $\epsilon = 1$ (archimedean) case.

In particular, tropicalization establishes a connection between polynomial systems and piecewise affine systems that are somehow similar to the ones arising in game problems. It allows one to transfer results from the world of combinatorics to "classical" equations solving. We investigate the consequences of this correspondence on complexity and numerical issues. For instance, combinatorial problems can be solved in a robust way. Hence, situations in which the tropicalization is faithful lead to improved algorithms for classical problems. In particular, scalings for the polynomial eigenproblems based on tropical preprocessings have started to be used in matrix analysis [85], [88].

Moreover, the tropical approach has been recently applied to construct examples of linear programs in which the central path has an unexpectedly high total curvature [61], and it has also led to positive polynomial-time average case results concerning the complexity of mean payoff games. Similarly, we are studying semidefinite programming over non-archimedean fields [49], [29], with the goal to better understand complexity issues in classical semidefinite and semi-algebraic programming.

4. Application Domains

4.1. Discrete event systems (manufacturing systems, networks)

One important class of applications of max-plus algebra comes from discrete event dynamical systems [66]. In particular, modelling timed systems subject to synchronization and concurrency phenomena leads to studying dynamical systems that are non-smooth, but which have remarkable structural properties (nonexpansiveness in certain metrics , monotonicity) or combinatorial properties. Algebraic methods allow one to obtain analytical expressions for performance measures (throughput, waiting time, etc). A recent application, to emergency call centers, can be found in [62].

4.2. Optimal control and games

Optimal control and game theory have numerous well established applications fields: mathematical economy and finance, stock optimization, optimization of networks, decision making, etc. In most of these applications, one needs either to derive analytical or qualitative properties of solutions, or design exact or approximation algorithms adapted to large scale problems.

4.3. Operations Research

We develop, or have developed, several aspects of operations research, including the application of stochastic control to optimal pricing, optimal measurement in networks [109]. Applications of tropical methods arise in particular from discrete optimization [68], [70], scheduling problems with and-or constraints [103], or product mix auctions [114].

4.4. Computing program and dynamical systems invariants

A number of programs and systems verification questions, in which safety considerations are involved, reduce to computing invariant subsets of dynamical systems. This approach appears in various guises in computer science, for instance in static analysis of program by abstract interpretation, along the lines of P. and R. Cousot [73], but also in control (eg, computing safety regions by solving Isaacs PDEs). These invariant sets are often sought in some tractable effective class: ellipsoids, polyhedra, parametric classes of polyhedra with a controlled complexity (the so called "templates" introduced by Sankaranarayanan, Sipma and Manna [110]), shadows of sets represented by linear matrix inequalities, disjunctive constraints represented by tropical polyhedra [63], etc. The computation of invariants boils down to solving large scale fixed point problems. The latter are of the same nature as the ones encountered in the theory of zero-sum games, and so, the techniques developed in the previous research directions (especially methods of monotonicity, nonexpansiveness, discretization of PDEs, etc) apply to the present setting, see e.g. [83], [86] for the application of policy iteration type algorithms, or for the application for fixed point problems over the space of quadratic forms [7]. The problem of computation of invariants is indeed a key issue needing the methods of several fields: convex and nonconvex programming, semidefinite programming and symbolic computation (to handle semialgebraic invariants), nonlinear fixed point theory, approximation theory, tropical methods (to handle disjunctions), and formal proof (to certify numerical invariants or inequalities).

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. *Awards*

- The Gaspard Monge Programme for Optimization and Operations Research (PGMO), a corporate sponsorship of EDF operated by Fondation Mathématique Jacques Hadamard, coordinated by Stéphane Gaubert, received the "Grand Prix AEF – meilleures initiatives partagées Universités Entreprises".
- Mateusz Skomra received the Dodu prize (distinction for the best talk of a young researcher) at the conférence SMAI-MODE 2016.

6. New Software and Platforms

6.1. New Software

6.1.1. Coq-Polyhedra

Coq-Polyhedra is a library which aims at formalizing convex polyhedra in Coq. A description of the associated contributions can be found in Section 7.3.1. Coq-Polyhedra is distributed under the CeCILL-B licence, and can be found at https://github.com/nhojem/Coq-Polyhedra.

7. New Results

7.1. Optimal control and zero-sum games

7.1.1. *Fixed points of order preserving homogeneous maps and zero-sum games*

Participants: Marianne Akian, Stéphane Gaubert, Antoine Hochart.

The PhD work of Antoine Hochart [12] deals with the applications of methods of non-linear fixed point theory to zero-sum games.

A highlight of his PhD is the characterization of the property of ergodicity for zero-sum games. In the special "zero-player" case, i.e., for a Markov chain equipped with an additive functional (payment) of the trajectory, the ergodicity condition entails that the mean payoff is independent of the initial state, for any choice of the payment. In the case of finite Markov chains, ergodicity admits several characterizations, including a combinatorial one (the uniqueness of the final class). This carries over to the two player case: ergodicity is now characterized by the absence of certain pairs of conjugate invariant sets (dominions), and it can be checked using directed hypergraphs algorithms. This leads to an explicit combinatorial sufficient condition for the solvability of the "ergodic equation", which is the main tool in the numerical approach of the mean payoff problem. These results appeared in [59], [58], [60]. A more general approach was developed in [30], in which zero-sum games are now studied abstractly in terms of accretive operators. This allows one to show that the bias vector (the solution of the ergodic equation) is unique for a generic perturbation of the payments.

Another series of results of the thesis concern the finite action space, showing that the set of payments for which the bias vector is not unique coincides with the union of lower dimensional cells of a polyhedral complex, which an application to perturbation schemes in policy iteration [47].

A last result of the thesis is a representation theorem for "payment free" Shapley operators, showing that these are characterized by monotonicity and homogeneity axioms [48]. This extends to the two-player case known representation theorems for risk measures.

7.1.2. *Probabilistic and max-plus approximation of Hamilton-Jacobi-Bellman equations*

Participants: Marianne Akian, Eric Fodjo.

The PhD thesis of Eric Fodjo concerns stochastic control problems obtained in particular in the modelisation of portfolio selection with transaction costs. The dynamic programming method leads to a Hamilton-Jacobi-Bellman partial differential equation, on a space with a dimension at least equal to the number of risky assets. The curse of dimensionality does not allow one to solve numerically these equations for a large dimension (greater to 5). We propose to tackle these problems with numerical methods combining policy iterations, probabilistic discretisations, max-plus discretisations, in order to increase the possible dimension. Another solution is to replace policy iterations by an approximation with optimal switching problems.

In [27], [26] (also presented in [35], [23]), we consider fully nonlinear Hamilton-Jacobi-Bellman equations associated to diffusion control problems with finite horizon involving a finite set-valued (or switching) control and possibly a continuum-valued control. We construct a lower complexity probabilistic numerical algorithm by combining the idempotent expansion properties obtained by McEneaney, Kaise and Han [93], [99] for solving such problems with a numerical probabilistic method such as the one proposed by Fahim, Touzi and Warin [78] for solving some fully nonlinear parabolic partial differential equations, when the volatility does not oscillate too much. Numerical tests on a small example of pricing and hedging an option are presented. Moreover, more recently, we improved the method of Fahim, Touzi and Warin to allow one to solve fully nonlinear parabolic partial differential equations with general volatilities.

7.2. Non-linear Perron-Frobenius theory, nonexpansive mappings and metric geometry

7.2.1. *Isometries of the Hilbert geometry*
Participant: Cormac Walsh.

In a collaboration with Bas Lemmens (Kent University, UK), we have been studying the Hilbert geometry in finite dimensions, especially its horofunction boundary and isometry group. The book chapter [117] contains a survey of this work. However, the infinite dimensional case is also interesting, and has been used as a tool for many years in non-linear analysis. Despite this, very little is known about the geometry of these spaces when the dimension is infinite.

An example of a problem in which we are interested is the following. In finite dimension it is known that a Hilbert geometry is isometric to a normed space if and only if it is a simplex. We have shown [118] that, more generally, a Hilbert geometry is isometric to a Banach space if and only if it is the cross-section of a positive cone, that is, the cone of positive continuous functions on some compact topological space. To solve this problem we found it useful to study the horofunction boundary in the infinite-dimensional case.

We are continuing to study similar problems in relation to this topic in collaboration with Bas Lemmens of the University of Kent.

7.2.2. *Volume growth in the Hilbert geometry*
Participant: Cormac Walsh.

In a collaboration with Constantin Vernicos of Université Montpellier 2, we are investigating how the volume of a ball in a Hilbert geometry grows as its radius increases. Specifically, we are studying the volume entropy

$$\lim_{r \to \infty} \frac{\log \operatorname{Vol} B(x,r)}{r}, \tag{17}$$

where $B(x,r)$ is the ball with center x and radius r, and Vol denotes some notion of volume, for example, the Holmes–Thompson or Busemann definitions. Note that the entropy does not depend on the particular choice of x, nor on the choice of the volume. It is known that the hyperbolic space, or indeed any Hilbert geometry with a C^2-smooth boundary of strictly positive curvature, has entropy $n-1$, where n is the dimension, and it has recently been proved that this is the maximal entropy possible for Hilbert geometries of the given dimension.

Constantin Vernicos has shown that, in dimension 2 and 3, the volume entropy of a Hilbert geometry on a convex body is equal to exactly twice the *approximability* of the body, that is, the power of $1/\epsilon$ governing the growth of the number of vertices needed to approximate the body by a polytope within ϵ, as ϵ decreases.

Studying polytopal Hilbert geometries, we have demonstrated [53] a close relation between the volume and the number of *flags* of the polytope, more precisely, that the volume of large balls is asymptotically proportional to the number of flags. This suggested to us defining a new notion of approximability using flags rather than vertices. We have shown [53] that the volume entropy of a Hilbert geometry on a convex body is equal to exactly twice this *flag-approximability* in all dimensions. This implies in particular that the volume entropy of a convex body is equal to that of its dual.

7.2.3. *The set of minimal upper bounds of two matrices in the Loewner order*
Participant: Nikolas Stott.

A classical theorem of Kadison shows that the space of symmetric matrices equipped with the Loewner order is an anti-lattice, meaning that two matrices have a least upper bound if and only if they are comparable. In [52], we refined this theorem by characterizing the set of minimal upper bounds: we showed that it is homeomorphic to the quotient space $O(p) \setminus O(p,q)/O(q)$, where $O(p,q)$ denotes the orthogonal group associated to the quadratic form with signature (p,q), and $O(p)$ denotes the standard pth orthogonal group.

7.2.4. *Checking the strict positivity of Kraus maps is NP-hard*
Participant: Stéphane Gaubert.

In collaboration with Zheng Qu (now with HKU, Hong Kong), I studied several decision problems arising from the spectral theory of Kraus maps (trace preserving completely positive maps), acting on the cone of positive semidefinite matrices. The latter appear in quantum information. We showed that checking the irreducibility (absence of non-trivial invariant face of the cone) and primitivity properties (requiring the iterates of the map to send the cone to its interior) can be checked in polynomial time, whereas checking positivity (whether the map sends the cone to its interior) is NP-hard. In [17], we studied complexity issues related to Kraus maps, and showed in particular that checking whether a Kraus map sends the cone to its interior is NP-hard.

7.3. Tropical algebra and convex geometry

7.3.1. *Formalizing convex polyhedra in Coq*
Participants: Xavier Allamigeon, Ricardo Katz [Conicet, Argentine].

We formalize a certain fragment of the theory of convex polyhedra and their combinatorial properties. Our motivation is that convex polyhedra are involved in a wide range of analysis techniques such as in formal verification, and that their combinatorial properties are used to establish more fundamental results, especially in tropical geometry.

This formalization has been conducted in Coq using the Mathematical Components library. We have implemented a full formalization of the simplex algorithm, which allows to make several key properties of convex polyhedra (feasibility, unboundedness, etc) decidable. From this, we have deduced a formal proof of strong duality theorem in linear programming, and of Farkas lemma. We also have a formal implementation of Motzkin's double description method, which provides a constructive way to prove Minkowski theorem for polyhedra.

7.3.2. *Tropical totally positive matrices*
Participants: Stéphane Gaubert, Adi Niv.

In [50], we investigate the tropical analogues of totally positive and totally non-negative matrices, i.e, the images by the valuation of the corresponding classes of matrices over a non-archimedean field. We show that tropical totally positive matrices essentially coincide with the Monge matrices (defined by the positivity of 2×2 tropical minors), arising in optimal transport. These results have been presented in [41], [40].

7.3.3. *Tropical compound matrix identities*
Participants: Marianne Akian, Stéphane Gaubert, Adi Niv.

In [55], [57], we proved some identities on matrices using a weak and a strong transfer principles. In the present work, we prove identities on compound matrices in extended tropical semirings. Such identities include analogues to properties of conjugate matrices, powers of matrices and $\mathrm{adj}\,(A)\det\,(A)^{-1}$, all of which have implications on the eigenvalues of the corresponding matrices. A tropical Sylvester-Franke identity is provided as well. Even though part of these identities hold over any commutative ring, they cannot be adjusted to semirings with symmetry using the existing weak and strong transfer principles. Here, we provide the proofs by means of graph theory arguments.

7.3.4. *Supertropical algebra*

Participant: Adi Niv.

Several properties of matrices over the tropical algebra are studied using the supertropical algebra introduced in [92].

The only invertible matrices in tropical algebra are diagonal matrices, permutation matrices and their products. However, the pseudo-inverse A^∇, defined as $\frac{1}{\det(A)} \operatorname{adj}(A)$, with $\det(A)$ being the tropical permanent, inherits some classical algebraic properties and has some surprising new ones. In [104], defining B and B' to be tropically similar if $B' = A^\nabla B A$, we examine the characteristic (max-)polynomials of tropically similar matrices as well as those of pseudo-inverses. Other miscellaneous results include a new proof of the identity for $\det(AB)$ and a connection to stabilization of the powers of definite matrices.

In a joint work with Louis Rowen (Bar Ilan Univ.) [21], we study the pathology that causes tropical eigenspaces of distinct supertropical eigenvalues of a non-singular matrix A, to be dependent. We show that in lower dimensions the eigenvectors of distinct eigenvalues are independent, as desired. The index set that differentiates between subsequent essential monomials of the characteristic polynomial, yields an eigenvalue λ, and corresponds to the columns of the eigenmatrix $A + \lambda I$ from which the eigenvectors are taken. We ascertain the cause for failure in higher dimensions, and prove that independence of the eigenvectors is recovered in case the "difference criterion" holds, defined in terms of disjoint differences between index sets of subsequent coefficients. We conclude by considering the eigenvectors of the matrix $A^\nabla := \frac{1}{\det(A)} \operatorname{adj}(A)$ and the connection of the independence question to generalized eigenvectors.

7.3.5. *Volume and integer points of tropical polytopes*

Participants: Marie Maccaig, Stéphane Gaubert.

We investigated the volume of tropical polytopes, as well as the number of integer points contained in integer polytopes. We proved that even approximating these values for a tropical polytope given by its vertices is hard, with no approximation algorithm with factor $2^{\operatorname{poly}(m,n)}$ existing. We further proved the $\sharp$P-hardness for the analogous problems for tropical polytopes instead defined by inequalities. We also investigated the relation between the set of integer points of a tropical polytope and the image by the valuation of polytopes over the nonarchimedean field of Puiseux series.

7.3.6. *Primal dual pair of max-algebraic integer linear programs (MLP)*

Participant: Marie Maccaig.

There are known weak and strong duality theorems for max-algebraic linear programs. I investigated the integer versions of these problems; considering the impact of requiring integer solutions instead of real solutions. I proved a tight bound on the duality gap for a pair of integer solutions to the primal and dual MLPs, and searched for conditions on when the optimal values of the integer primal and dual MLPs coincide.

7.3.7. *Tropical Jacobi identity*

Participants: Marie Maccaig, Adi Niv.

In a joint work with Sergei Sergeev (Birmingham), we investigated the combinatorial interpretation for the Tropical Jacobi identity. Inspired by Butkovic's paper, "Max-algebra, the algebra of combinatorics?" and many other links between max-algebra and combinatorics, we try to link this tropical quantity to a new type of multiple assignment problem.

7.4. Tropical methods applied to optimization, perturbation theory and matrix analysis

7.4.1. *Majorization inequalities for valuations of eigenvalues using tropical algebra*

Participants: Marianne Akian, Stéphane Gaubert.

We consider a matrix with entries over the field of Puiseux series, equipped with its non-archimedean valuation (the leading exponent). In [13], with Ravindra Bapat (Univ. New Delhi), we establish majorization inequalities relating the sequence of the valuations of the eigenvalues of a matrix with the tropical eigenvalues of its valuation matrix (the latter is obtained by taking the valuation entrywise). We also show that, generically in the leading coefficients of the Puiseux series, the precise asymptotics of eigenvalues, eigenvectors and condition numbers can be determined. For this, we apply diagonal scalings constructed from the dual variables of a parametric optimal assignment constructed from the valuation matrix.

In recent works with Andrea Marchesini and Françoise Tisseur (Manchester University), we use the same technique to establish an archimedean analogue of the above inequalities, which applies to matrix polynomials with coefficients in the field of complex numbers, equipped with the modulus as its valuation. This allows us in particular to improve the accuracy of the numerical computation of the eigenvalues of such matrix polynomials.

In [15], with Meisam Sharify (IPM, Tehran, Iran), we also establish log-majorization inequalities of the eigenvalues of matrix polynomials using the tropical roots of some scalar polynomials depending only on the norms of the matrix coefficients. This extends to the case of matrix polynomials some bounds obtained by Hadamard, Ostrowski and Pólya for the roots of scalar polynomials.

These works have been presented in [22].

7.4.2. *Tropicalization of the central path and application to the complexity of interior point methods*

Participants: Xavier Allamigeon, Stéphane Gaubert.

This work is in collaboration with Pascal Benchimol (now with EDF Labs) and Michael Joswig (TU-Berlin).

In optimization, path-following interior point methods are driven to an optimal solution along a trajectory called the central path. The *central path* of a linear program $\mathrm{LP}(A, b, c) \equiv \min\{c \cdot x \mid Ax \leq b, \ x \geq 0\}$ is defined as the set of the optimal solutions (x^μ, w^μ) of the barrier problems:

$$
\begin{aligned}
\text{minimize} \quad & c \cdot x - \mu \left(\sum_{j=1}^{n} \log x_j + \sum_{i=1}^{m} \log w_i \right) \\
\text{subject to} \quad & Ax + w = b, \ x > 0, \ w > 0
\end{aligned}
$$

While the complexity of interior point methods is known to be polynomial, an important question is to study the number of iterations which are performed by interior point methods, in particular whether it can be bounded by a polynomial in the dimension (mn) of the problem. This is motivated by one of Smale's problems, on the existence of a strongly polynomial complexity algorithm for linear programming. So far, this question has been essentially addressed though the study of the curvature of the central path, which measures how far a path differs from a straight line, see [75], [74], [77], [76]. In particular, by analogy with the classical Hirsch conjecture, Deza, Terlaky and Zinchencko [76] conjectured that $O(m)$ is also an upper bound for the total curvature.

In a work of X. Allamigeon, P. Benchimol, S. Gaubert, and M. Joswig, we study the tropicalization of the central path. The *tropical central path* is defined as the logarithmic limit of the central paths of a parametric family of linear programs $\mathrm{LP}(A(t), b(t), c(t))$, where the entries $A_{ij}(t)$, $b_i(t)$ and $c_j(t)$ are definable functions in an o-minimal structure called the *Hardy field*.

A first contribution is to provide a purely geometric characterization of the tropical central path. We have shown that the tropical analytic center is the greatest element of the tropical feasible set. Moreover, any point of the tropical central path is the greatest element of the tropical feasible set intersected with a sublevel set of the tropical objective function.

Thanks to this characterization, we identify a class of path-following interior-point methods which are not strongly polynomial. This class corresponds to primal-dual interior-point methods which iterates in the so-called "wide" neighborhood of the central path arising from the logarithmic barrier. It includes short step, long step as well as predictor-corrector types of interior-point methods. In more details, we establish a lower bound on the number of iterations of these methods, expressed in terms of the number of tropical segments constituting the tropical central path. In this way, we exhibit a family of linear programs with $3d + 1$ inequalities in dimension $2d$ on which the aforementioned interior point methods require $\Omega(2^d)$ iterations. The same family provides a counterexample to Deza, Terlaky and Zinchenko's conjecture, having a total curvature in $\Omega(2^d)$.

A first part of these results is in the preprint [61], further results been presented in [32].

7.4.3. *Tropical approach to semidefinite programming*

Participants: Xavier Allamigeon, Stéphane Gaubert, Mateusz Skomra.

Semidefinite programming consists in optimizing a linear function over a spectrahedron. The latter is a subset of $\mathbb{R}^n$ defined by linear matrix inequalities, i.e., a set of the form

$$\left\{ x \in \mathbb{R}^n : Q^{(0)} + x_1 Q^{(1)} + \cdots + x_n Q^{(n)} \succeq 0 \right\}$$

where the $Q^{(k)}$ are symmetric matrices of order m, and $\succeq$ denotes the Loewner order on the space of symmetric matrices. By definition, $X \succeq Y$ if and only if $X - Y$ is positive semidefinite.

Semidefinite programming is a fundamental tool in convex optimization. It is used to solve various applications from engineering sciences, and also to obtain approximate solutions or bounds for hard problems arising in combinatorial optimization and semialgebraic optimization.

A general issue in computational optimization is to develop combinatorial algorithms for semidefinite programming. Indeed, semidefinite programs are usually solved via interior point methods. However, the latter provide an approximate solution in a polynomial number of iterations, provided that a strictly feasible initial solution. Semidefinite programming becomes a much harder matter if one requires an exact solution. The feasibility problem belongs to $\mathsf{NP}_\mathbb{R} \cap \mathsf{coNP}_\mathbb{R}$, where the subscript $\mathbb{R}$ refers to the BSS model of computation. It is not known to be in NP in the bit model.

We address semidefinite programming in the case where the field $\mathbb{R}$ is replaced by a nonarchimedean field, like the field of Puiseux series. In this case, methods from tropical geometry can be applied and are expected to allow one, in generic situations, to reduce semialgebraic problems to combinatorial problems, involving only the nonarchimedean valuations (leading exponents) of the coefficients of the input.

To this purpose, we first study tropical spectrahedra, which are defined as the images by the valuation of nonarchimedean spectrahedra. We establish that they are closed semilinear sets, and that, under a genericity condition, they are described by explicit inequalities expressing the nonnegativity of tropical minors of order 1 and 2. These results are gathered in the preprint [49].

Then, we show that the feasibility problem for a generic tropical spectrahedron is equivalent to solving a stochastic mean payoff game (with perfect information). The complexity of these games is a long-standing open problem. They are not known to be polynomial, however they belong to the class $\mathsf{NP} \cap \mathsf{coNP}$, and they can be solved efficiently in practice. This allows to apply stochastic game algorithms to solve nonarchimedean semidefinite feasibility problems. We obtain in this way both theoretical bounds and a practicable method which solves some large scale instances. Part of this latter work has been published in the proceedings of the conference ISSAC 2016 [29].

7.5. Applications

7.5.1. *Geometry of the Loewner order and application to the synthesis of quadratic invariants in static analysis of program*

Participants: Xavier Allamigeon, Stéphane Gaubert, Nikolas Stott.

This work is in collaboration with Éric Goubault and Sylvie Putot (from LIX).

We introduce a new numerical abstract domain based on ellipsoids designed for the formal verification of switched linear systems. The novelty of this domain does not consist in the use of ellipsoids as abstractions, but rather in the fact that we overcome two key difficulties which so far have limited the use of ellipsoids in abstract interpretation. The first issue is that the ordered set of ellipsoids does not constitute a lattice. This implies that there is a priori no canonical choice of the abstraction of the union of two sets, making the analysis less predictable as it relies on the selection of good upper bounds. The second issue is that most recent works using on ellipsoids rely on LMI methods. The latter are efficient on moderate size examples but they are inherently limited by the complexity of interior point algorithms, which, in the case of matrix inequality problems, do not scale as well as for linear programming or second order cone programming problems.

We developed a new approach, in which we reduce the computation of an invariant to the determination of a fixed point, or eigenvector, of a non-linear map that provides a safe upper-approximation of the action induced by the program on the space of quadratic forms. This allows one to obtain invariants of systems of sized inaccessible by LMI methods, at the price of a limited loss of precision. A key ingredient here is the fast computation of least upper bounds in Löwner ordering, by an algebraic algorithm. This relies on the study of the geometry of the space of quadratic forms (Section 7.2.3).

A first part of this work is described in the article [16], which is the extended version of [65] which won the best paper award at the conference EMSOFT 2015. Followup work is dealing with the extension of these results to switched affine systems with guards.

7.5.2. *Performance evaluation of an emergency call center based on tropical polynomial systems*

Participants: Xavier Allamigeon, Vianney Boeuf, Stéphane Gaubert.

This work arose from a question raised by Régis Reboul from Préfecture de Police de Paris (PP), regarding the analysis of the projected evolution of the treatment of emergency calls (17-18-112). This work benefited from the help of LtL Stéphane Raclot, from Brigade de Sapeurs de Pompiers de Paris (BSPP). It is part of the PhD work of Vianney Bœuf, carried out in collaboration with BSPP.

We introduced an algebraic approach which allows to analyze the performance of systems involving priorities and modeled by timed Petri nets. Our results apply to the class of Petri nets in which the places can be partitioned in two categories: the routing in certain places is subject to priority rules, whereas the routing at the other places is free choice.

In [62], we introduced a discrete model, showing that the counter variables, which determine the number of firings of the different transitions as a function of time, are the solutions of a piecewise linear dynamical system. Moreover, we establish that in the fluid approximation of this model, the stationary regimes are precisely the solutions of a set of lexicographic piecewise linear equations, which constitutes a polynomial system over a tropical (min-plus) semifield of germs.

In essence, this result shows that computing stationary regimes reduces to solving tropical polynomial systems. Solving tropical polynomial systems is one of the most basic problems of tropical geometry. The latter provides insights on the nature of solutions, as well as algorithmic tools. In particular, the tropical approach allows one to determine the different congestion phases of the system.

We applied this approach to a case study relative to the project led by Préfecture de Police de Paris, involving BSPP, of a new organization to handle emergency calls to Police (number 17), Firemen (number 18), and untyped emergency calls (number 112), in the Paris area. We initially introduced, in [62], a simplified model of emergency call center, and we concentrated on the analysis of an essential feature of the organization: the two level emergency procedure. Operators at level 1 initially receive the calls, qualify their urgency, handle the non urgent ones, and transfer the urgent cases to specialized level 2 operators who complete the instruction. We solved the associated system of tropical polynomial equations and arrived at an explicit computation of the different congestion phases, depending on the ratio of the numbers of operators of level 2 and 1.

We subsequently developed a more complex model, taking into account the different characteristics of the calls to 17 and 18, and developed a realistic simulation tool to validate the results. Moreover, in [28], we developed an alternative model, relying on fluid Petri nets (dynamical systems with piecewise affine vector fields). We showed that the fluid and discrete models have the same stationary regimes, and that some pathological features of the discrete model (anomalous periodic orbits appearing under certain arithmetical conditions) vanish in the fluid Petri net case.

7.5.3. *Smart Data Pricing*

Participants: Marianne Akian, Jean-Bernard Eytard.

This work is in collaboration with Mustapha Bouhtou (Orange Labs).

The PhD work of Jean-Bernard Eytard concerns the optimal pricing of data trafic in mobile networks. We developed a bilevel programming approach, allowing to an operator to balance the load in the network through price incentives. We showed that a subclass of bilevel programs can be solved in polynomial time, by combining methods of tropical geometry and of discrete convexity. This work has been presented in [31].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Yield management methods applied to the pricing of data trafic in mobile networks. CRE (research contract) with Orange Labs (Orange Labs partner: Mustapha Bouhtou).
- Decentralized mechanisms of operation of power systems: equilibria and efficiency. A collaboration started on this topic at the fall, Nadia Oujdane, Olivier Beaude, and Riadh Zorgati from EDF-labs. This leads to the PhD work of Paulin Jacquot, supervised by Stéphane Gaubert (starting CIFRE PhD).

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. *ANR*

- Participation of Cormac Walsh to the ANR white project FINSLER (Géométrie de Finsler et applications), 2012-2016.
- Projet ANR CAFEIN (Combinaison d'approches formelles pour l'étude d'invariants numériques), responsable P.L. Garoche. Partenaires : ONERA, CEA LIST, ENSTA Paristech, Inria Saclay (Maxplus, Toccata, Parkas), Université de Perpignan, Prover, Rockwell Collins France.
- Projet ANR MALTHY (Méthodes ALgébriques pour la vérification de modèles Temporisés et HYbrides), responsable T. Dang. Partenaires : Verimag, CEA LIST, Inria Rennes, Inria Saclay, VISEO/Object Direct.

- Projet ANR DEMOCRITE ("DEmonstrateur d'un MOteur de Couverture des Risques sur un TErritoire), responsable Emmanuel Lapébie (CEA). Partenaires : CEA-GRAMAT, BSPP, Inria Saclay (Maxplus), Institut PPRIME - UPR3346 (CNRS, Univ. Poitiers, ISAE-ENSMA), IPSIS, SYSTEL, ARMINES-E.M. Alès-ISR, CERDACC (Univ. de Haute-Alsace).

9.1.2. Programme Gaspard Monge pour l'Optimisation

- Projet intitulé "Méthodes tropicales pour l'optimisation", responsable X. Allamigeon, faisant intervenir M. Akian, V. Boeuf, S. Gaubert, A.Hochart, R. Katz, et M. Skomra.

9.1.3. iCODE (Institut pour le Contrôle et la Décision de l'Idex Paris-Saclay)

- White project "Stabilité et stabilisation des systèmes commutés" (Oct 2014-June 2016), including M. Akian, X. Allamigeon, S. Gaubert, and members of EPI Geco, L2S, LIX, LSV (ENS Cachan), UVSQ.

9.2. International Research Visitors

9.2.1. Visits of International Scientists

- Ricardo Katz (Conicet and Cifasis, Argentina), May–June 2016
- Rajendra Bhatia (Indian Statistical Institute, New Delhi), 2 weeks in June 2016.
- Vladimir Gurvich (Rutgers), 2 weeks in Dec 2016.

9.2.2. Visits to International Teams

9.2.2.1. Research Stays Abroad

- S. Gaubert, invitation of one week to HKU, Hong-Kong, collaboration with Zheng Qu.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- S. Gaubert co-organized, jointly with D. Grigoriev (CNRS, Lille), M. Joswig (TU Berlin) and T. Theobald (Frankfurt), a Dagstuhl workshop, on "Effectivity in tropical mathematics, and beyond".

10.1.1.2. Member of the Organizing Committees

- M. Akian co-organized a workshop on hybrid systems, IHP, 2016.
- S. Gaubert co-organizes the "Séminaire Parisien d'Optimisation".
- S. Gaubert, co-organized with S. Charousset (EDF) the PGMO days at EDF labs Paris-Saclay.
- X. Allamigeon co-organized two invited sessions on semidefinite programming and tropical methods at the conference PGMO Days.

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

- S. Gaubert, president of the scientific committee of SMAI-MODE 2016 (Toulouse, March 2016).

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- S. Gaubert is member of the editorial committee of the collection Mathématiques et Applications, SMAI and Springer.

- S. Gaubert is associate editor of Linear and Multilinear Algebra.
- S. Gaubert is associate editor of RAIRO Operations research.

10.1.4. Invited Talks

- M. Akian, Keynote lecture at ETAMM 2016.
- S. Gaubert, invited lecture at the 2016 Conference on Applied Mathematics, Hong Kong University.

10.1.5. Leadership within the Scientific Community

- S. Gaubert coordinates the Gaspard Monge Programme for Optimization and Operations Research (PGMO), a corporate sponsorship of EDF operated by Fondation Mathématique Jacques Hadamard at Paris-Saclay. The goal of the program is to help to develop the research community in these fields, connecting academic and industrial researchers. It includes a research initiative on energy, led by S. Charousset from EDF (IROE, funding focused projects on the optimization of energy), and a subprogram dedicated to basic research (PRMO, funding smaller size projects). Projects are selected after an open call, instructed by the scientific commitee of PGMO. The program organizes advanced invited lectures for PhD students and researchers (in 2016, lectures by Yuri Nesterov from Louvain and Jean-Bernard Lasserre from LAAS), a regular seminar, and an annual conference (PGMO days, 250 participants in 2016). The program is currently being renewed, with an opening to new industrial partners interested by optimization. See https://www.fondation-hadamard.fr/PGMO for more information on PGMO.

10.1.6. Research Administration

- M. Akian :
  - Member of the "comité de liaison SMAI-MODE" since June 2015.
- S. Gaubert :
  - Coordinator of PGMO (Gaspard Monge Program for Optimization and Operations Research, a corporate sponsorship of EDF operated by FMJH).
  - Member of the scientific council of CMAP.
- X. Allamigeon:
  - Member of the scientific committee of Inria Saclay – Ile-de-France.
  - Member of the laboratory council of CMAP.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- M. Akian
  - Course "Markov decision processes: dynamic programming and applications" joint between (3rd year of) ENSTA and M2 "Mathématiques et Applications", U. Paris Saclay, "Optimization", and shared with Jean-Philippe Chancelier (ENPC), 15 hours each.
- X. Allamigeon
  - Petites classes et encadrement d'enseignements d'approfondissement de Recherche Opérationnelle en troisième année à l'École Polytechnique (programme d'approfondissement de Mathématiques Appliquées) (niveau M1).
  - Cours du M2 "Optimisation" de l'Université Paris Saclay, cours partagé avec Céline Gicquel et Dominique Quadri (LRI, Université Paris Sud).
  - Co-responsabilité du programme d'approfondissement en mathématiques appliquées (troisième année) à l'École Polytechnique.
- V. Boeuf

– Petite classe du cours de tronc commun de 1ere année "Introduction à l'optimisation" de l'École des ponts (ENPC), niveau L3.

- S. Gaubert
    – Course "Systèmes à Événements Discrets", option MAREVA, ENSMP.
    – Course "Algèbre max-plus pour le contrôle optimal et les jeux" of "Parcours Optimisation, Jeux et Dynamique" (ODJ) of M2 "Mathématiques et Applications" of Paris 6 University and École Polytechnique.
    – Lecture of Operations Research, third year of École Polytechnique. The lectures notes were published this year as a book [46].

- A. Hochart
    – Cours de niveau L1 et L2 à l'Univ. Paris Diderot (Paris VII), dans le cadre d'un monitorat (34h).

- M. Skomra
    – Petite classe du cours de tronc commun de 1ere année "Introduction à l'optimisation" de l'École des ponts (ENPC), niveau L1.

- N. Stott
    – Cours et TD "Introduction à la programmation graphique en C++", option MAREVA et semaine Athens à l'ENSMP (École des Mines de Paris), niveau M1.

10.2.2. Supervision

- PhD: Antoine Hochart, registered at École Polytechnique, since October 2013, thesis supervisor: Stéphane Gaubert, cosupervision: Marianne Akian, defended on 14 Nov 2016.
- PhD in progress : Eric Fodjo, registered at École Polytechnique, since October 2013, thesis supervisor: Marianne Akian.
- PhD in progress : Nikolas Stott, registered at École Polytechnique, since October 2014, thesis supervisor: Stéphane Gaubert, cosupervision: Xavier Allamigeon, Éric Goubault, Sylvie Putot.
- PhD in progress : Vianney Boeuf, registered at École Polytechnique, since October 2014, thesis supervisor: Stéphane Gaubert, cosupervision: Stéphane Raclot (BSPP), Marianne Akian, Xavier Allamigeon.
- PhD in progress : Mateusz Skomra, registered at Univ. Paris Saclay since October 2015, thesis supervisor: Stéphane Gaubert, cosupervision: Xavier Allamigeon.
- PhD in progress : Jean-Bernard Eytard, registered at Univ. Paris Saclay since October 2015, thesis supervisor: Stéphane Gaubert, cosupervision: Marianne Akian, Mustapha Bouhtou.
- PhD in progress: Paulin Jacquot, registered at Univ. Paris Saclay since November 2016, thesis supervisor: Stéphane Gaubert, cosupervision: Nadia Oujdane, Olivier Beaude (EDF).

10.2.3. Juries

- M. Akian
    – Vice-president of the jury of the 2016 competition for CR2 positions of Inria Saclay–Île-de-France.
    – Member of the jury selecting the 2016 PGMO PhD price.
    – Jury of the PhD thesis of A. Hochart (X, Palaiseau, November 2016).

- S. Gaubert
    – Member of hiring commitee (Professor position) at Paris 6 University.
    – Member of hiring commitee (Assistant Professor position) at Limoges University.
    – Jury of the HdR of A. Auger (Saclay, 2016).

- Jury of the HdR of B. Scherrer (Nancy, reviewer, 2016).
- Jury of the PhD thesis of C. Josz (Paris 6, reviewer, 2016).
- Jury of the PhD thesis of A. Parmentier (ENPC, reviewer, 2016).
- Jury of the PhD thesis of A. Hochart (X, Palaiseau, 2016).

10.3. Popularization

- J.P. Quadrat :
  - Webmaster of the site http://www.maxplus.org, dedicated to max-plus algebra.

10.4. Conferences, Seminars

- M. Akian
  - Conference "Emerging Trends in Applied Mathematics and Mechanics" (ETAMM), Perpignan, May 30 - June 3, 2016, Title of the talk: "Majorization inequalities for valuations of eigenvalues using tropical algebra".
  - MTNS 2016 (22nd International Symposium on Mathematical Theory of Networks and Systems), Minneapolis, July 12-15, 2016, Title of the talk: "Solving Hamilton-Jacobi-Bellman equations by combining a max-plus linear approximation and a probabilistic numerical method".
  - Workshop "Numerical methods for Hamilton-Jacobi equations in optimal control and related fields" at the Radon Institute, Austrian Academy of Sciences, Linz, Austria, Nov. 21 - Nov. 25, 2016. Title of the talk: "Solving Hamilton-Jacobi-Bellman equations by combining a max-plus linear approximation and a probabilistic numerical method".
  - Dagstuhl Seminar "Algorithms and Effectivity in Tropical Mathematics and Beyond", Nov. 28 - Dec. 02, 2016. Title of the talk: "Majorization inequalities for valuations of eigenvalues using tropical algebra".
- X. Allamigeon
  - Journée Equipe McTAO, Inria Sophia Antipolis, January 18, 2016. Title of the talk: "Long and winding central paths".
  - Conference "Emerging Trends in Applied Mathematics and Mechanics" (ETAMM), Perpignan, May 30 - June 3, 2016. Title of the talk: "Long and winding central paths".
  - Groupe de travail combinatoire du Plateau de Saclay, June 8, 2016. Title of the talk: "Long and winding central paths".
  - Dagstuhl Seminar "Algorithms and Effectivity in Tropical Mathematics and Beyond", Nov. 28 - Dec. 02, 2016. Title of the talk: "Log-barrier interior-point methods are not strongly polynomial".
  - Séminaire Parisien d'Optimisation, Paris, December 12, 2016. Title of the talk: "Log-barrier interior-point methods are not strongly polynomial".
- V. Boeuf
  - Congrès annuel de la société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF), Compiègne, February 10-12, 2016. Title of the talk: "Évaluation de performance en réception d'appels d'urgence : débits asymptotiques dans un réseau de Pétri avec priorités.".
  - Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS), Taormina, October 25-28, 2016. Title of the talk: "Stationary solutions of discrete and continuous Petri nets with priorities".
- J.B. Eytard

- – PGMO Days, Nov. 8-9, 2016, Palaiseau. Title of the talk: "Price incentives in mobile networks: a tropical approach".
- E. Fodjo
  - – 9th European Summer School in Financial Mathematics, Aug. 29- Sep. 2, 2016, St Petersbourg, Russia. Title of the talk: "A probabilistic max-plus numerical method for solving stochastic control problems".
  - – SIAM Conference on Financial Mathematics & Engineering, Nov. 17-19, 2016, Austin, Texas, USA. Title of the talk: "A probabilistic max-plus numerical method for solving stochastic control problems".
  - – 55th Conference on Decision and Control (CDC 2016), Dec. 12-14, 2016, Las Vegas, USA. Title of the talk: "A probabilistic max-plus numerical method for solving stochastic control problems".
- S. Gaubert
  - – International Conference on Tensors, Matrices and their Applications, Chern Institute of Mathematics, Nankai University, Tianjin, 20-24 May 2016. Title of the talk: "Ergodicity conditions in non-linear Perron-Frobenius theory and application to nonnegative tensors".
  - – SIAM Conference on Discrete Mathematics, Atlanta, June 6-10, 2016. Title of the talk: "Stochastic mean payoff games are tropical semidefinite programs".
  - – Workshop: Advances in Convex Analysis and Optimization Erice July 6-11, 2016. Title of the talk: "Zero-sum games, non-archimedean convexity and sinuous central paths".
  - – 2016 Conference on Applied Mathematics, The University of Hong Kong Aug 23-26, 2016. Title of the talk: "Zero-sum games, non-archimedean convexity and sinuous central paths".
  - – Seminar of the algebraic geometry group at the University of Hong Kong, Aug 24, 2016. Title of the talk. "Tropical spectrahedra".
  - – Dagstuhl Seminar "Algorithms and Effectivity in Tropical Mathematics and Beyond", Nov. 28 - Dec. 02, 2016. Title of the talk "The equivalence between zero-sum games and tropical convexity".
- A. Hochart
  - – Game Theory PhD Seminar, Paris, February 1, 2016. Title of the talk: "Ergodic problems for zero-sum stochastic games".
  - – Conference SMAI-MODE, Toulouse, March 23-27, 2016. Title of the talk: "Une approche opérateur accrétif pour les jeux stochastiques avec critère ergodique".
  - – MTNS 2016 (22nd International Symposium on Mathematical Theory of Networks and Systems), Minneapolis, July 12-15, 2016, Title of the talk: "An Accretive Operator Approach to Ergodic Problems for Zero-Sum Games".
- M. MacCaig
  - – Birmingham Young Mathematicians Conference 2016. Title: "Tropical algebra: Optimisation, tropical polytopes and integer points".
  - – Student tropical algebraic geometry seminar (STAGS 2016), Yale. Title: "Calculating the volume of tropical polytopes is hard".
  - – ILAS 20th annual meeting, K.U.Leuven, July 11, 2016. Title: "Scaling to integer matrices in max-algebra".
  - – Dagstuhl Seminar on Algorithms and Effectivity in Tropical Mathematics.
- A. Niv
  - – Conference "Recent advances in linear algebra and graph-theory", U.T.Chattanooga, March 5-6, 2016. Title of the talk: "Introduction to tropical total positivity".

– Seminar at Afeka Academic College of Engineering, Tel-Aviv, May 8, 2016. Title of the talk: "Assignment problems via tropical matrices".

– Tropical symposium, ILAS 20th annual meeting, K.U.Leuven, July 11, 2016. Title of the talk: "Total non-negativity via valuations in tropical algebra".

• M. Skomra

– Conference SMAI-MODE, Toulouse, March 23–25, 2016. Title of the talk: "Les spectrahèdres tropicaux et leur relation aux jeux stochastiques".

– Conference "Emerging Trends in Applied Mathematics and Mechanics" (ETAMM), Perpignan, May 30 - June 3, 2016. Title of the talk: "Nonarchimedean semidefinite programming and stochastic games".

– Séminaire des doctorants du CMAP, Palaiseau, June 10, 2016. Title of the talk: "Une relation entre la programmation semi-définie paramétrique et les jeux stochastiques".

– Conference of the International Linear Algebra Society (ILAS), Leuven, July 11-16, 2016. Title of the talk: "Nonarchimedean semidefinite programming and stochastic games".

– International Symposium on Symbolic and Algebraic Computation (ISSAC), Waterloo, Ontario, Canada, July 19-22, 2016. Title of the talk: "Solving Generic Nonarchimedean Semidefinite Programs using Stochastic Game Algorithms".

– Conference PGMO Days, EDF Labs Paris-Saclay, November 8-9, 2016. Title of the talk: "Solving Generic Nonarchimedean Semidefinite Programs using Stochastic Game Algorithms".

– Dagstuhl Seminar "Algorithms and Effectivity in Tropical Mathematics and Beyond", Nov. 28 - Dec. 02, 2016. Title of the talk: "Tropical spectrahedra and stochastic mean payoff games".

• C. Walsh

– Conference "New Methods in Finsler Geometry", Leipzig, July 5-9, 2016. Title of the talk: "Studying isometry groups using the horofunction boundary".

11. Bibliography

Major publications by the team in recent years

[1] M. AKIAN, S. GAUBERT, R. BAPAT.*Non-archimedean valuations of eigenvalues of matrix polynomials*, in "Linear Algebra and its Applications", June 2016, vol. 498, p. 592–627, Also arXiv:1601.00438 [*DOI :* 10.1016/J.LAA.2016.02.036], https://hal.inria.fr/hal-01251803.

[2] M. AKIAN, S. GAUBERT, A. GUTERMAN.*Tropical polyhedra are equivalent to mean payoff games*, in "Internat. J. Algebra Comput.", 2012, vol. 22, n⁰ 1, 1250001, 43, http://dx.doi.org/10.1142/S0218196711006674.

[3] M. AKIAN, S. GAUBERT, R. NUSSBAUM.*Uniqueness of the fixed point of nonexpansive semidifferentiable maps*, in "Transactions of the American Mathematical Society", February 2016, vol. 368, n⁰ 2, Also arXiv:1201.1536 [*DOI :* 10.1090/S0002-9947-2015-06413-7], https://hal.inria.fr/hal-00783682.

[4] M. AKIAN, S. GAUBERT, C. WALSH.*The max-plus Martin boundary*, in "Doc. Math.", 2009, vol. 14, p. 195–240.

[5] X. ALLAMIGEON, P. BENCHIMOL, S. GAUBERT.*The Tropical Shadow-Vertex Algorithm Solves Mean Payoff Games in Polynomial Time on Average*, in "ICALP 2014", Lecture Notes in Computer Science, Springer, 2014, vol. 8572, p. 89–100, http://dx.doi.org/10.1007/978-3-662-43948-7_8.

[6] X. ALLAMIGEON, P. BENCHIMOL, S. GAUBERT, M. JOSWIG.*Combinatorial simplex algorithms can solve mean payoff games*, in "SIAM J. Opt.", 2015, vol. 24, nᵒ 4, p. 2096–2117.

[7] X. ALLAMIGEON, S. GAUBERT, E. GOUBAULT, S. PUTOT, N. STOTT.*A scalable algebraic method to infer quadratic invariants of switched systems*, in "Proceedings of the International Conference on Embedded Software (EMSOFT)", 2015.

[8] J. BOLTE, S. GAUBERT, G. VIGERAL.*Definable zero-sum stochastic games*, in "Mathematics of Operations Research", 2014, vol. 40, nᵒ 1, p. 171–191, Also arXiv:1301.1967, http://dx.doi.org/10.1287/moor.2014.0666.

[9] S. GAUBERT, T. LEPOUTRE.*Discrete limit and monotonicity properties of the Floquet eigenvalue in an age structured cell division cycle model*, in "J. Math. Biol.", 2015, http://dx.doi.org/10.1007/s00285-015-0874-3.

[10] S. GAUBERT, G. VIGERAL.*A maximin characterization of the escape rate of nonexpansive mappings in metrically convex spaces*, in "Math. Proc. of Cambridge Phil. Soc.", 2012, vol. 152, p. 341–363, Also arXiv:1012.4765, http://dx.doi.org/10.1017/S0305004111000673.

[11] C. WALSH.*The horofunction boundary and isometry group of the Hilbert geometry*, in "Handbook of Hilbert Geometry", IRMA Lectures in Mathematics and Theoretical Physics, European Mathematical Society, 2014, vol. 22, https://hal.inria.fr/hal-00782827.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[12] A. HOCHART.*Nonlinear Perron-Frobenius theory and mean-payoff zero-sum stochastic games*, Ecole polytechnique, Palaiseau, France, November 2016, https://pastel.archives-ouvertes.fr/tel-01423953.

Articles in International Peer-Reviewed Journal

[13] M. AKIAN, S. GAUBERT, R. BAPAT.*Non-archimedean valuations of eigenvalues of matrix polynomials*, in "Linear Algebra and its Applications", June 2016, vol. 498, p. 592–627, Also arXiv:1601.00438 [*DOI :* 10.1016/J.LAA.2016.02.036], https://hal.inria.fr/hal-01251803.

[14] M. AKIAN, S. GAUBERT, R. NUSSBAUM.*Uniqueness of the fixed point of nonexpansive semidifferentiable maps*, in "Transactions of the American Mathematical Society", February 2016, vol. 368, nᵒ 2, Also arXiv:1201.1536 [*DOI :* 10.1090/S0002-9947-2015-06413-7], https://hal.inria.fr/hal-00783682.

[15] M. AKIAN, S. GAUBERT, M. SHARIFY.*Log-majorization of the moduli of the eigenvalues of a matrix polynomial by tropical roots*, in "Linear Algebra and its Applications", 2017, Also arXiv:1304.2967 [*DOI :* 10.1016/J.LAA.2016.11.004], https://hal.inria.fr/hal-00881196.

[16] X. ALLAMIGEON, S. GAUBERT, E. GOUBAULT, S. PUTOT, N. STOTT.*A scalable algebraic method to infer quadratic invariants of switched systems*, in "ACM Transactions on Embedded Computing Systems (TECS)", August 2016, vol. 15, nᵒ 4 [*DOI :* 10.1145/2932187], https://hal.inria.fr/hal-01423744.

[17] S. GAUBERT, Z. QU.*Checking the strict positivity of Kraus maps is NP-hard*, in "Information Processing Letters", February 2017, vol. 118, p. 35–43, Also preprint arXiv:1402.1429 [*DOI :* 10.1016/J.IPL.2016.09.008], https://hal.inria.fr/hal-01097942.

[18] S. Gaubert, Z. Qu, S. Sridharan.*Maximizing concave piecewise affine functions on the unitary group*, in "Optimization Letters", April 2016, vol. 10, n$^{\text{o}}$ 4, p. 655-665 [*DOI :* 10.1007/s11590-015-0951-y], https://hal.inria.fr/hal-01248813.

[19] M. MacCaig.*On scaling to an integer matrix and graphs with integer weighted cycles*, in "Linear Algebra and Applications", June 2016, vol. 498, p. 490–520 [*DOI :* 10.1016/j.laa.2016.01.018], https://hal.inria.fr/hal-01423517.

[20] M. MacCaig.*Exploring the complexity of the integer image problem in the max-algebra*, in "Discrete Applied Mathematics", January 2017, vol. 217, n$^{\text{o}}$ 2, p. 261–275 [*DOI :* 10.1016/j.dam.2016.09.016], https://hal.inria.fr/hal-01423520.

[21] A. Niv, L. Rowen.*Dependence of tropical eigenspaces*, in "Communications in Algebra", 2017, vol. 45, n$^{\text{o}}$ 3, p. 924-942 [*DOI :* 10.1080/00927872.2016.1172603], https://hal.inria.fr/hal-01253422.

Invited Conferences

[22] M. Akian.*Majorization inequalities for valuations of eigenvalues using tropical algebra*, in "Emerging Trends in Applied Mathematics and Mechanics (ETAMM)", Perpignan, France, May 2016, https://hal.inria.fr/hal-01425463.

[23] M. Akian.*Solving Hamilton-Jacobi-Bellman equations by combining a max-plus linear approximation and a probabilistic numerical method*, in "Workshop "Numerical methods for Hamilton-Jacobi equations in optimal control and related fields" at the Radon Institute, Austrian Academy of Sciences", Linz, Austria, November 2016, https://hal.inria.fr/hal-01425485.

[24] S. Gaubert.*Zero-sum games, non-archimedean convexity and sinuous central paths*, in "Workshop: Advances in Convex Analysis and Optimization", Erice, Italy, July 2016, https://hal.inria.fr/hal-01428994.

[25] S. Gaubert.*Zero-sum games, non-archimedean convexity and sinuous central paths*, in "2016 Conference on Applied Mathematics", Hong Kong, China, August 2016, https://hal.inria.fr/hal-01428997.

International Conferences with Proceedings

[26] M. Akian, E. Fodjo.*A probabilistic max-plus numerical method for solving stochastic control problems*, in "55th Conference on Decision and Control (CDC 2016)", Las Vegas, United States, December 2016, Also arXiv:1605.02816, https://hal.inria.fr/hal-01425344.

[27] M. Akian, E. Fodjo.*Solving Hamilton-Jacobi-Bellman equations by combining a max-plus linear approximation and a probabilistic numerical method*, in "22nd International Symposium on Mathematical Theory of Networks and Systems (MTNS)", Minneapolis, United States, July 2016, https://hal.inria.fr/hal-01425473.

[28] X. Allamigeon, V. Boeuf, S. Gaubert.*Stationary solutions of discrete and continuous Petri nets with priorities*, in "10th EAI International Conference on Performance Evaluation Methodologies and Tools", Taormina, France, 10th EAI International Conference on Performance Evaluation Methodologies and Tools, October 2016, 13 pages, 3 figures + 1 table. The version appearing in the proceedings of the conference VALUETOOLS 2016 is an extended abstract, https://hal.inria.fr/hal-01421885.

[29] X. ALLAMIGEON, S. GAUBERT, M. SKOMRA.*Solving Generic Nonarchimedean Semidefinite Programs Using Stochastic Game Algorithms*, in "ISSAC '16: International Symposium on Symbolic and Algebraic Computation", Waterloo, France, Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation (ISSAC'16), ACM, July 2016, Also arXiv:1603.06916 [*DOI : 10.1145/2930889.2930935*], https://hal.inria.fr/hal-01422638.

[30] A. HOCHART.*An Accretive Operator Approach to Ergodic Problems for Zero-Sum Games*, in "22nd International Symposium on Mathematical Theory of Networks and Systems (MTNS)", Minneapolis, United States, July 2016, https://hal.inria.fr/hal-01425315.

Conferences without Proceedings

[31] M. AKIAN, M. BOUHTOU, J. B. EYTARD, S. GAUBERT.*Price incentives in mobile networks: a tropical approach*, in "PGMO Days", Palaiseau, France, November 2016, https://hal.inria.fr/hal-01425352.

[32] X. ALLAMIGEON.*Long and winding central paths*, in "Emerging Trends in Applied Mathematics and Mechanics (ETAMM)", Perpignan, France, May 2016, https://hal.inria.fr/hal-01428770.

[33] X. ALLAMIGEON, S. GAUBERT, M. SKOMRA.*Solving Generic Nonarchimedean Semidefinite Programs using Stochastic Game Algorithms*, in "PGMO Days", Palaiseau, France, November 2016, https://hal.inria.fr/hal-01429021.

[34] V. BOEUF.*Évaluation de performance en réception d'appels d'urgence : débits asymptotiques dans un réseau de Pétri avec priorités*, in "Congrès annuel de la société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF)", Compiègne, France, February 2016, https://hal.inria.fr/hal-01428975.

[35] E. FODJO.*A probabilistic max-plus numerical method for solving stochastic control problems*, in "SIAM Conference on Financial Mathematics & Engineering", Austin, United States, November 2016, https://hal.inria.fr/hal-01425335.

[36] S. GAUBERT.*Ergodicity conditions in non-linear Perron-Frobenius theory and application to nonnegative tensors*, in "International Conference on Tensors, Matrices and their Applications", Tianjin, China, May 2016, https://hal.inria.fr/hal-01428984.

[37] S. GAUBERT.*Stochastic mean payoff games are tropical semidefinite programs*, in "SIAM Conference on Discrete Mathematics", Atlanta, United States, June 2016, https://hal.inria.fr/hal-01428990.

[38] A. HOCHART.*Une approche opérateur accrétif pour les jeux stochastiques avec critère ergodique*, in "Conference SMAI-MODE", Toulouse, France, March 2016, https://hal.inria.fr/hal-01429185.

[39] M. MACCAIG.*Scaling to integer matrices in max-algebra*, in "20th Conference of the International Linear Algebra Society (ILAS)", Leuven, Belgium, July 2016, Tropical symposium, https://hal.inria.fr/hal-01428924.

[40] A. NIV.*Introduction to tropical total positivity*, in "Recent advances in linear algebra and graph-theory", Chattanooga, United States, March 2016, https://hal.inria.fr/hal-01425299.

[41] A. NIV.*Total non-negativity via valuations in tropical algebra*, in "20th Conference of the International Linear Algebra Society (ILAS)", Leuven, Belgium, July 2016, Tropical symposium, https://hal.inria.fr/hal-01425308.

[42] M. Skomra.*Les spectrahèdres tropicaux et leur relation aux jeux stochastiques*, in "Conference SMAI-MODE", Toulouse, France, March 2016, https://hal.inria.fr/hal-01429006.

[43] M. Skomra.*Nonarchimedean semidefinite programming and stochastic games*, in "20th Conference of the International Linear Algebra Society (ILAS)", Leuven, Belgium, July 2016, Tropical symposium, https://hal.inria.fr/hal-01428932.

[44] M. Skomra.*Nonarchimedean semidefinite programming and stochastic games*, in "Emerging Trends in Applied Mathematics and Mechanics (ETAMM)", Perpignan, France, May 2016, https://hal.inria.fr/hal-01428789.

[45] C. Walsh.*Studying isometry groups using the horofunction boundary*, in "New Methods in Finsler Geometry", Leipzig, Germany, July 2016, https://hal.inria.fr/hal-01428906.

Scientific Books (or Scientific Book chapters)

[46] F. Bonnans, S. Gaubert.*Recherche opérationnelle. Aspects mathématiques et applications*, Ellipse, March 2016, 391, https://hal.inria.fr/hal-01422645.

Other Publications

[47] M. Akian, S. Gaubert, A. Hochart.*Generic uniqueness of the bias vector of finite stochastic games with perfect information*, October 2016, arXiv:1610.09651, https://hal.inria.fr/hal-01425543.

[48] M. Akian, S. Gaubert, A. Hochart.*Minimax representation of nonexpansive functions and application to zero-sum recursive games*, January 2017, arXiv:1605.04518, https://hal.inria.fr/hal-01425551.

[49] X. Allamigeon, S. Gaubert, M. Skomra.*Tropical spectrahedra*, October 2016, arXiv:1610.06746, https://hal.inria.fr/hal-01422639.

[50] S. Gaubert, A. Niv.*Tropical totally positive matrices*, December 2016, arXiv:1606.00238, https://hal.inria.fr/hal-01423747.

[51] D. Jules, S. Gaubert, M. Joswig.*A tropical isoperimetric inequality*, December 2016, arXiv:1611.04148, https://hal.inria.fr/hal-01422522.

[52] N. Stott.*Maximal lower bounds in the Löwner order*, December 2016, eprint: arXiv:1612.05664 [math.RA], https://hal.inria.fr/hal-01423497.

[53] C. Vernicos, C. Walsh.*Flag-approximability of convex bodies and volume growth of Hilbert geometries*, December 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01423693.

References in notes

[54] M. Akian, S. Gaubert.*Spectral theorem for convex monotone homogeneous maps, and ergodic control*, in "Nonlinear Anal.", 2003, vol. 52, n° 2, p. 637–679, http://dx.doi.org/10.1016/S0362-546X(02)00170-0.

[55] M. Akian, S. Gaubert, A. Guterman.*Linear independence over tropical semirings and beyond*, in "Proceedings of the International Conference on Tropical and Idempotent Mathematics", G. Litvinov, S.

SERGEEV (editors), Contemporary Mathematics, American Mathematical Society, 2009, vol. 495, p. 1-38, http://www.arxiv.org/abs/0812.3496.

[56] M. AKIAN, S. GAUBERT, A. GUTERMAN.*Tropical polyhedra are equivalent to mean payoff games*, in "Internat. J. Algebra Comput.", 2012, vol. 22, n⁰ 1, 1250001, 43 [*DOI :* 10.1142/S0218196711006674], http://arxiv.org/abs/0912.2462.

[57] M. AKIAN, S. GAUBERT, A. GUTERMAN.*Tropical Cramer Determinants Revisited*, in "Tropical and Idempotent Mathematics and Applications", G. LITVINOV, S. SERGEEV (editors), Contemporary Mathematics, AMS, 2014, vol. 616, 45, See also arXiv:1309.6298, https://hal.inria.fr/hal-00881203.

[58] M. AKIAN, S. GAUBERT, A. HOCHART.*Ergodicity Condition for Zero-Sum Games*, in "SIAM Conference on Control and its Applications (SIAM CT'15)", Paris, France, July 2015, https://hal.inria.fr/hal-01252413.

[59] M. AKIAN, S. GAUBERT, A. HOCHART.*Ergodicity conditions for zero-sum games*, in "Discrete and Continuous Dynamical Systems - Series A", September 2015, vol. 35, n⁰ 9, 31, See also arXiv: 1405.4658 [*DOI :* 10.3934/DCDS.2015.35.3901], https://hal.inria.fr/hal-01096206.

[60] M. AKIAN, S. GAUBERT, A. HOCHART.*Hypergraph conditions for the solvability of the ergodic equation for zero-sum games*, in "54th IEEE Conference on Decision and Control (CDC 2015)", Osaka, Japan, Proceedings of the 54th IEEE Annual Conference on Decision and Control (CDC), Osaka, 2015, December 2015, https://hal.inria.fr/hal-01249321.

[61] X. ALLAMIGEON, P. BENCHIMOL, S. GAUBERT, M. JOSWIG.*Long and winding central paths*, May 2014, Preprint arXiv:1405.4161, v2 May 2015, https://hal.inria.fr/hal-01096452.

[62] X. ALLAMIGEON, V. BOEUF, S. GAUBERT.*Performance evaluation of an emergency call center: tropical polynomial systems applied to timed Petri nets*, in "13th International Conference, Formal Modeling and Analysis of Timed Systems (FORMATS 2015)", Madrid, Spain, Formal Modeling and Analysis of Timed Systems, Springer, September 2015, vol. 9268 [*DOI :* 10.1007/978-3-319-22975-1_2], https://hal.inria.fr/hal-01248814.

[63] X. ALLAMIGEON, S. GAUBERT, É. GOUBAULT.*Inferring Min and Max Invariants Using Max-plus Polyhedra*, in "Proceedings of the 15th International Static Analysis Symposium (SAS'08)", Valencia, Spain, LNCS, Springer, 2008, vol. 5079, p. 189–204, http://dx.doi.org/10.1007/978-3-540-69166-2_13.

[64] X. ALLAMIGEON, S. GAUBERT, E. GOUBAULT.*Computing the Vertices of Tropical Polyhedra using Directed Hypergraphs*, in "Discrete Comp. Geom.", 2012, Published on line [*DOI :* 10.1007/S00454-012-9469-6], http://fr.arxiv.org/abs/0904.3436v3.

[65] X. ALLAMIGEON, S. GAUBERT, E. GOUBAULT, S. PUTOT, N. STOTT.*A scalable algebraic method to infer quadratic invariants of switched systems*, in "International Conference on Embedded Software (EMSOFT'2015)", Amsterdam, Netherlands, International Conference on Embedded Software (EMSOFT'2015), Alain Girault, Inria, Grenoble, France and Nan Guan, Northeastern University, China, October 2015 [*DOI :* 10.1109/EMSOFT.2015.7318262], https://hal.inria.fr/hal-01249320.

[66] F. BACCELLI, G. COHEN, G.-J. OLSDER, J.-P. QUADRAT.*Synchronization and linearity: an algebra for discrete event systems*, Wiley, 1992.

[67] G. BARLES, S. MIRRAHIMI, B. PERTHAME.*Concentration in Lotka-Volterra parabolic or integral equations: a general convergence result*, in "Methods Appl. Anal.", 2009, vol. 16, n⁰ 3, p. 321–340, http://dx.doi.org/10.4310/MAA.2009.v16.n3.a4.

[68] P. BUTKOVIČ.*Max-algebra: the linear algebra of combinatorics?*, in "Linear Algebra and its applications", 2003, vol. 367, p. 313–335.

[69] P. BUTKOVIČ.*Max-linear systems: theory and algorithms*, Springer Monographs in Mathematics, Springer-Verlag London, Ltd., London, 2010, xviii+272, http://dx.doi.org/10.1007/978-1-84996-299-5.

[70] J. COCHET-TERRASSON, G. COHEN, S. GAUBERT, M. MC GETTRICK, J.-P. QUADRAT.*Numerical computation of spectral elements in max-plus algebra*, in "Proc. of the IFAC Conference on System Structure and Control", Nantes, July 1998.

[71] G. COHEN, S. GAUBERT, J.-P. QUADRAT.*Max-plus algebra and system theory: where we are and where to go now*, in "Annual Reviews in Control", 1999, vol. 23, p. 207–219.

[72] A. CONNES, C. CONSANI.*Geometry of the arithmetic site*, 2015, arXiv:1502.05580.

[73] P. COUSOT, R. COUSOT.*Abstract Interpretation: A unified lattice model for static analysis of programs by construction of approximations of fixed points*, in "Principles of Programming Languages 4", 1977, p. 238–252.

[74] J. DE LOERA, B. STURMFELS, C. VINZANT.*The central curve in linear programming*, in "Foundations of Computational Mathematics", 2012, vol. 12, n⁰ 4, p. 509–540.

[75] J.-P. DEDIEU, G. MALAJOVICH, M. SHUB.*On the Curvature of the Central Path of Linear Programming Theory*, in "Foundations of Computational Mathematics", 2005, vol. 5, n⁰ 2, p. 145–171.

[76] A. DEZA, T. TERLAKY, Y. ZINCHENKO.*Polytopes and arrangements: diameter and curvature*, in "Operations Research Letters", 2008, vol. 36, n⁰ 2, p. 215–222.

[77] A. DEZA, T. TERLAKY, Y. ZINCHENKO.*Central path curvature and iteration-complexity for redundant Klee-Minty cubes*, in "Advances in applied mathematics and global optimization", New York, Adv. Mech. Math., Springer, 2009, vol. 17, p. 223–256, http://dx.doi.org/10.1007/978-0-387-75714-8_7.

[78] A. FAHIM, N. TOUZI, X. WARIN.*A probabilistic numerical method for fully nonlinear parabolic PDEs*, in "Ann. Appl. Probab.", 2011, vol. 21, n⁰ 4, p. 1322–1364, http://dx.doi.org/10.1214/10-AAP723.

[79] A. FATHI, A. SICONOLFI.*Existence of C^1 critical subsolutions of the Hamilton-Jacobi equation*, in "Invent. Math.", 2004, vol. 155, n⁰ 2, p. 363–388, http://dx.doi.org/10.1007/s00222-003-0323-6.

[80] O. FERCOQ, M. AKIAN, M. BOUHTOU, S. GAUBERT.*Ergodic control and polyhedral approaches to PageRank optimization*, in "IEEE Trans. Automat. Control", 2013, vol. 58, n⁰ 1, p. 134–148, http://dx.doi.org/10.1109/TAC.2012.2226103.

[81] W. FLEMING, W. MCENEANEY.*A max-plus based algorithm for an HJB equation of non-linear filtering*, in "SIAM J. Control and Opt.", 2000, p. 683–710.

[82] S. FOMIN, A. ZELEVINSKY.*Cluster algebras. I. Foundations*, in "J. Amer. Math. Soc.", 2002, vol. 15, n⁰ 2, p. 497–529, http://arxiv.org/abs/math.RT/0104151.

[83] S. GAUBERT, E. GOUBAULT, A. TALY, S. ZENNOU.*Static Analysis by Policy Iteration in Relational Domains*, in "Proceedings of the Proc. of the 16th European Symposium on Programming (ESOP'07)", Braga (Portugal), LNCS, Springer, 2007, vol. 4421, p. 237–252, http://dx.doi.org/10.1007/978-3-540-71316-6_17.

[84] S. GAUBERT, W. MCENEANEY, Z. QU.*Curse of dimensionality reduction in max-plus based approximation methods: theoretical estimates and improved pruning algorithms*, in "Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC 11)", Orlando, FL, USA, December 2011, p. 1054-1061, http://arxiv.org/abs/1109.5241.

[85] S. GAUBERT, M. SHARIFY.*Tropical scaling of polynomial matrices*, in "Positive systems", Berlin, Lecture Notes in Control and Inform. Sci., Springer, 2009, vol. 389, p. 291–303, http://dx.doi.org/10.1007/978-3-642-02894-6_28.

[86] T. M. GAWLITZA, H. SEIDL, A. ADJÉ, S. GAUBERT, E. GOUBAULT.*Abstract interpretation meets convex optimization*, in "J. Symbolic Comput.", 2012, vol. 47, n⁰ 12, p. 1416–1446, Special issue on Invariant generation and reasoning about loops, http://dx.doi.org/10.1016/j.jsc.2011.12.048.

[87] I. M. GELFAND, M. M. KAPRANOV, A. V. ZELEVINSKY.*Discriminants, resultants and multidimensional determinants*, Modern Birkhäuser Classics, Birkhäuser Boston, Inc., Boston, MA, 2008, x+523, Reprint of the 1994 edition.

[88] S. HAMMARLING, C. J. MUNRO, F. TISSEUR.*An algorithm for the complete solution of quadratic eigenvalue problems*, in "ACM Trans. Math. Software", 2013, vol. 39, n⁰ 3, Art. 18, 19, http://dx.doi.org/10.1145/2450153.2450156.

[89] B. HEIDERGOTT, G. J. OLSDER, J. VAN DER WOUDE.*Max Plus at Work: Modeling and Analysis of Synchronized Systems: A Course on Max-Plus Algebra and Its Applications*, Princeton, 2005.

[90] H. ISHII, H. MITAKE.*Representation formulas for solutions of Hamilton-Jacobi equations with convex Hamiltonians*, in "Indiana Univ. Math. J.", 2007, vol. 56, n⁰ 5, p. 2159–2183, http://dx.doi.org/10.1512/iumj.2007.56.3048.

[91] I. ITENBERG, G. MIKHALKIN, E. SHUSTIN.*Tropical algebraic geometry*, Oberwolfach Seminars, Birkhäuser Verlag, Basel, 2007, vol. 35, viii+103.

[92] Z. IZHAKIAN.*Tropical arithmetic and matrix algebra*, in "Comm. Algebra", 2009, vol. 37, n⁰ 4, p. 1445–1468, http://dx.doi.org/10.1080/00927870802466967.

[93] H. KAISE, W. M. MCENEANEY.*Idempotent expansions for continuous-time stochastic control: compact control space*, in "Proceedings of the 49th IEEE Conference on Decision and Control", Atlanta, Dec. 2010.

[94] V. KOLOKOLTSOV, V. MASLOV.*Idempotent analysis and applications*, Kluwer Acad. Publisher, 1997.

[95] B. LEMMENS, R. NUSSBAUM.*Nonlinear Perron-Frobenius theory*, Cambridge Tracts in Mathematics, Cambridge University Press, Cambridge, 2012, vol. 189, xii+323, http://dx.doi.org/10.1017/CBO9781139026079.

[96] Q. LU, M. MADSEN, M. MILATA, S. RAVN, U. FAHRENBERG, K. G. LARSEN.*Reachability Analysis for Timed Automata using Max-Plus Algebra*, in "J. Logic Alg. Prog.", 2012, vol. 81, n° 3, p. 298-313.

[97] V. MASLOV.*Méthodes Operatorielles*, Edition Mir, Moscou, 1987.

[98] W. MCENEANEY, A. DESHPANDE, S. GAUBERT.*Curse-of-Complexity Attenuation in the Curse-of-Dimensionality-Free Method for HJB PDEs*, in "Proc. of the 2008 American Control Conference", Seattle, Washington, USA, June 2008.

[99] W. M. MCENEANEY, H. KAISE, S. H. HAN.*Idempotent Method for Continuous-time Stochastic Control and Complexity Attenuation*, in "Proceedings of the 18th IFAC World Congress, 2011", Milano, Italie, 2011, p. 3216-3221.

[100] W. M. MCENEANEY.*A curse-of-dimensionality-free numerical method for solution of certain HJB PDEs*, in "SIAM J. Control Optim.", 2007, vol. 46, n° 4, p. 1239–1276, http://dx.doi.org/10.1137/040610830.

[101] J.-F. MERTENS, S. SORIN, S. ZAMIR.*Repeated Games*, Cambridge, 2015.

[102] G. MIKHALKIN.*Enumerative tropical algebraic geometry in $\mathbb{R}^2$*, in "J. Amer. Math. Soc.", 2005, vol. 18, n° 2, p. 313–377, http://dx.doi.org/10.1090/S0894-0347-05-00477-7.

[103] R. H. MÖHRING, M. SKUTELLA, F. STORK.*Scheduling with AND/OR precedence constraints*, in "SIAM J. Comput.", 2004, vol. 33, n° 2, p. 393–415, http://dx.doi.org/10.1137/S009753970037727X.

[104] A. NIV.*On pseudo-inverses of matrices and their characteristic polynomials in supertropical algebra*, in "Linear Algebra and its Applications", April 2015, vol. 471, p. 264–290 [*DOI :* 10.1016/J.LAA.2014.12.038], https://hal.inria.fr/hal-01253421.

[105] A. PAPADOPOULOS.*Metric spaces, convexity and non-positive curvature*, IRMA Lectures in Mathematics and Theoretical Physics, Second, European Mathematical Society (EMS), Zürich, 2014, vol. 6, xii+309, http://dx.doi.org/10.4171/132.

[106] J.-E. PIN.*Tropical Semirings*, in "Idempotency", J. GUNAWARDENA (editor), Publications of the Isaac Newton Institute, Cambridge University Press, 1998, vol. 11, p. 50–69.

[107] D. REEB, M. J. KASTORYANO, M. M. WOLF.*Hilbert's projective metric in quantum information theory*, in "J. Math. Phys.", 2011, vol. 52, n° 8, 082201, 33, http://dx.doi.org/10.1063/1.3615729.

[108] J. RICHTER-GEBERT, B. STURMFELS, T. THEOBALD.*First steps in tropical geometry*, in "Idempotent mathematics and mathematical physics", Providence, RI, Contemp. Math., Amer. Math. Soc., 2005, vol. 377, p. 289–317.

[109] G. SAGNOL, S. GAUBERT, M. BOUHTOU.*Optimal monitoring on large networks by Successive c-Optimal Designs*, in "Proceedings of the 22nd international teletraffic congress (ITC22), Amsterdam, The Netherlands, September", IEEE, 2010, [pdf] [*DOI :* 10.1109/ITC.2010.5608717].

[110] S. SANKARANARAYANAN, H. SIPMA, Z. MANNA.*Scalable Analysis of Linear Systems using Mathematical Programming*, in "VMCAI", LNCS, 2005, vol. 3385.

[111] R. SEPULCHRE, A. SARLETTE, P. ROUCHON.*Consensus in noncommutative spaces*, in "Proceedings of the 49th IEEE Conference on Decision and Control", Atlanta, USA, Dec 2010, p. 6596-6601, http://dx.doi.org/ 10.1109/CDC.2010.5717072.

[112] I. SIMON.*Limited subsets of a free monoid*, in "Proc. 19th Annual Symposium on Foundations of Computer Science", Piscataway, NJ, 1978, p. 143–150.

[113] H. L. SMITH.*Monotone dynamical systems*, Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 1995, vol. 41, x+174, An introduction to the theory of competitive and cooperative systems.

[114] N. M. TRAN, J. YU.*Product-Mix Auctions and Tropical Geometry*, 2015.

[115] O. VIRO.*Dequantization of real algebraic geometry on logarithmic paper*, in "European Congress of Mathematics, Vol. I (Barcelona, 2000)", Basel, Progr. Math., Birkhäuser, 2001, vol. 201, p. 135–146.

[116] O. VIRO.*Dequantization of real algebraic geometry on logarithmic paper*, in "European Congress of Mathematics, Vol. I (Barcelona, 2000)", Basel, Progr. Math., Birkhäuser, 2001, vol. 201, p. 135–146, http:// arxiv.org/abs/math.AG/0005163.

[117] C. WALSH.*The horofunction boundary and isometry group of the Hilbert geometry*, in "Handbook of Hilbert Geometry", A. PAPADOPOULOS, M. TROYANOV (editors), IRMA Lectures in Mathematics and Theoretical Physics, European Mathematical Society, 2014, vol. 22, https://hal.inria.fr/hal-00782827.

[118] C. WALSH.*Hilbert and Thompson geometries isometric to infinite-dimensional Banach spaces*, December 2015, working paper or preprint, https://hal.inria.fr/hal-01249343.

Team XPOP

statistical modelling for life sciences

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

RESEARCH CENTER
Saclay - Île-de-France

THEME
Computational Neuroscience and Medecine

Table of contents

Team XPOP

Creation of the Team: 2016 January 01

Keywords:

Computer Science and Digital Science:

3.1.1. - Modeling, representation
3.2.3. - Inference
3.3. - Data and knowledge analysis
3.3.1. - On-line analytical processing
3.3.2. - Data mining
3.3.3. - Big data analysis
3.4.1. - Supervised learning
3.4.2. - Unsupervised learning
3.4.4. - Optimization and learning
3.4.5. - Bayesian methods
3.4.6. - Neural networks
3.4.7. - Kernel methods
3.4.8. - Deep learning
5.9.2. - Estimation, modeling
6.1.1. - Continuous Modeling (PDE, ODE)
6.2.2. - Numerical probability
6.2.3. - Probabilistic methods
6.2.4. - Statistical methods
6.3.3. - Data processing
6.3.5. - Uncertainty Quantification

Other Research Topics and Application Domains:

1.1.5. - Genetics
1.1.6. - Genomics
1.1.9. - Bioinformatics
1.1.11. - Systems biology
2.2.3. - Cancer
2.2.4. - Infectious diseases, Virology
2.4.1. - Pharmaco kinetics and dynamics
9.1.1. - E-learning, MOOC

1. Members

Research Scientist
Marc Lavielle [Team leader, Inria, Senior Researcher]

Faculty Members
Julie Josse [Ecole Polytechnique, Associate Professor]
Erwan Le Pennec [Ecole Polytechnique, Associate Professor]

Eric Moulines [Ecole Polytechnique, Professor]

Technical Staff

François-Marie Floch [Inria, until Oct 2016]

PhD Students

Nicolas Brosse [Ecole Polytechnique]
Mohammed Karimi [Inria, from Oct 2016]
Genevieve Robin [Ecole Polytechnique, from Sep 2016]
Marine Zulian [Dassault Systemes, from Sep 2016, granted by CIFRE]

2. Overall Objectives

2.1. Developing sound, useful and usable methods

The main objective of XPOP is to develop new sound and rigorous methods for statistical modeling in the field of biology and life sciences. These methods for modeling include statistical methods of estimation, model diagnostics and model selection as well as methods for numerical models (systems of ordinary and partial differential equations). Historically, the key area where these methods have been used is population pharmacokinetics. However, the framework is currently being extended to sophisticated numerical models in the contexts of viral dynamics, glucose-insulin processes, tumor growth, precision medicine, intracellular processes, etc.

Furthermore, an important aim of XPOP is to transfer the methods developed into software packages so that they can be used in everyday practice.

2.2. Combining numerical, statistical and stochastic components of a model

Mathematical models that characterize complex biological phenomena are defined by systems of ordinary differential equations when dealing with dynamical systems that evolve with respect to time, or by partial differential equations when there is a spatial component in the model. Also, it is sometimes useful to integrate a stochastic aspect into the dynamical system in order to model stochastic intra-individual variability.

In order to use such methods, we must deal with complex numerical difficulties, generally related to resolving the systems of differential equations. Furthermore, to be able to check the quality of a model (i.e. its descriptive and predictive performances), we require data. The statistical aspect of the model is thus critical in how it takes into account different sources of variability and uncertainty, especially when data come from several individuals and we are interested in characterizing the inter-subject variability. Here, the tools of reference are mixed-effects models.

Confronted with such complex modeling problems, one of the goals of XPOP is to show the importance of combining numerical, statistical and stochastic approaches.

2.3. Developing future standards

Linear mixed-effects models have been well-used in statistics for a long time. They are a classical approach, essentially relying on matrix calculations in Gaussian models. Whereas a solid theoretical base has been developed for such models, *nonlinear* mixed-effects models (NLMEM) have received much less attention in the statistics community, even though they have been applied to many domains of interest. It has thus been the users of these models, such as pharmacometricians, who have taken them and developed methods, without really looking to develop a clean theoretical framework or understand the mathematical properties of the methods. This is why a standard estimation method in NLMEM is to linearize the model, and few people have been interested in understanding the properties of estimators obtained in this way.

Statisticians and pharmacometricians frequently realize the need to create bridges between these two communities. We are entirely convinced that this requires the development of new standards for population modeling that can be widely accepted by these various communities. These standards include the language used for encoding a model, the approach for representing a model and the methods for using it:

- **The approach.** Our approach consists in seeing a model as hierarchical, represented by a joint probability distribution. This joint distribution can be decomposed into a product of conditional distributions, each associated with a submodel (model for observations, individual parameters, etc.). Tasks required of the modeler are thus related to these probability distributions.

- **The methods.** Many tests have shown that algorithms implemented in MONOLIX are the most reliable, all the while being extremely fast. In fact, these algorithms are precisely described and published in well known statistical journals. In particular, the SAEM algorithm, used for calculating the maximum likelihood estimation of population parameters, has shown its worth in numerous situations. Its mathematical convergence has also been proven under quite general hypotheses.

- **The language.** Mlxtran is used by MONOLIX and other modeling tools and is today by far the most advanced language for representing models. Initially developed for representing pharmacometric models, its syntax also allows it to easily code dynamical systems defined by a system of ODEs, and statistical models involving continuous, discrete and survival variables. This flexibility is a true advantage both for numerical modelers and statisticians.

3. Research Program

3.1. Scientific positioning

"Interfaces" is the defining characteristic of XPOP:

The interface between statistics, probability and numerical methods. Mathematical modelling of complex biological phenomena require to combine numerical, stochastic and statistical approaches. The CMAP is therefore the right place to be for positioning the team at the interface between several mathematical disciplines.

The interface between mathematics and the life sciences. The goal of XPOP is to bring the right answers to the right questions. These answers are mathematical tools (statistics, numerical methods, etc.), whereas the questions come from the life sciences (pharmacology, medicine, biology, etc.). This is why the point of XPOP is not to take part in mathematical projects only, but also pluridisciplinary ones.

The interface between mathematics and software development. The development of new methods is the main activity of XPOP. However, new methods are only useful if they end up being implemented in a software tool. A strong partnership with Lixoft (the spin-off company who continue developing MONOLIX) is indispensable to maintaining this positioning.

3.2. The mixed-effects models

Mixed-effects models are statistical models with both fixed effects and random effects. They are well-adapted to situations where repeated measurements are made on the same individual/statistical unit.

Consider first a single subject i of the population. Let $y_i = (y_{ij}, 1 \leq j \leq n_i)$ be the vector of observations for this subject. The model that describes the observations y_i is assumed to be a parametric probabilistic model: let $p_Y(y_i; \psi_i)$ be the probability distribution of y_i, where ψ_i is a vector of parameters.

In a population framework, the vector of parameters ψ_i is assumed to be drawn from a population distribution $p_\Psi(\psi_i; \theta)$ where θ is a vector of population parameters.

Then, the probabilistic model is the joint probability distribution

$$p(y_i, \psi_i; \theta) = p_Y(y_i|\psi_i)p_\Psi(\psi_i; \theta) \tag{18}$$

To define a model thus consists in defining precisely these two terms.

In most applications, the observed data y_i are continuous longitudinal data. We then assume the following representation for y_i:

$$y_{ij} = f(t_{ij}, \psi_i) + g(t_{ij}, \psi_i)\varepsilon_{ij} \quad , \ 1 \le i \le N \quad , \quad 1 \le j \le n_i. \tag{19}$$

Here, y_{ij} is the observation obtained from subject i at time t_{ij}. The residual errors (ε_{ij}) are assumed to be standardized random variables (mean zero and variance 1). The residual error model is represented by function g in model (2).

Function f is usually the solution to a system of ordinary differential equations (pharmacokinetic/pharmacodynamic models, etc.) or a system of partial differential equations (tumor growth, respiratory system, etc.). This component is a fundamental component of the model since it defines the prediction of the observed kinetics for a given set of parameters.

The vector of individual parameters ψ_i is usually function of a vector of population parameters ψ_{pop}, a vector of random effects $\eta_i \sim \mathcal{N}(0, \Omega)$, a vector of individual covariates c_i (weight, age, gender, ...) and some fixed effects β.

The joint model of y and ψ depends then on a vector of parameters $\theta = (\psi_{\text{pop}}, \beta, \Omega)$.

3.3. Computational Statistical Methods

Central to modern statistics is the use of probabilistic models. To relate these models to data requires the ability to calculate the probability of the observed data: the likelihood function, which is central to most statistical methods and provides a principled framework to handle uncertainty.

The emergence of computational statistics as a collection of powerful and general methodologies for carrying out likelihood-based inference made complex models with non-standard data accessible to likelihood, including hierarchical models, models with intricate latent structure, and missing data.

In particular, algorithms previously developed by POPIX for mixed effects models, and today implemented in several software tools (especially MONOLIX) are part of these methods:

- the adaptive Metropolis-Hastings algorithm allows one to sample from the conditional distribution of the individual parameters $p(\psi_i|y_i; c_i, \theta)$,

- the SAEM algorithm is used to maximize the observed likelihood $\mathcal{L}(\theta; y) = p(y; \theta)$,

- Importance Sampling Monte Carlo simulations provide an accurate estimation of the observed log-likelihood $\log(\mathcal{L}(\theta; y))$.

Computational statistics is an area which remains extremely active today. Recently, one can notice that the incentive for further improvements and innovation comes mainly from three broad directions: the high dimensional challenge, the quest for adaptive procedures that can eliminate the cumbersome process of tuning "by hand" the settings of the algorithms and the need for flexible theoretical support, arguably required by all recent developments as well as many of the traditional MCMC algorithms that are widely used in practice.

Working in these three directions is a clear objective for XPOP.

3.4. Markov Chain Monte Carlo algorithms

While these Monte Carlo algorithms have turned into standard tools over the past decade, they still face difficulties in handling less regular problems such as those involved in deriving inference for high-dimensional models. One of the main problems encountered when using MCMC in this challenging settings is that it is difficult to design a Markov chain that efficiently samples the state space of interest.

The Metropolis-adjusted Langevin algorithm (MALA) is a Markov chain Monte Carlo (MCMC) method for obtaining random samples from a probability distribution for which direct sampling is difficult. As the name suggests, MALA uses a combination of two mechanisms to generate the states of a random walk that has the target probability distribution as an invariant measure:

1. new states are proposed using Langevin dynamics, which use evaluations of the gradient of the target probability density function;

2. these proposals are accepted or rejected using the Metropolis-Hastings algorithm, which uses evaluations of the target probability density (but not its gradient).

Informally, the Langevin dynamics drives the random walk towards regions of high probability in the manner of a gradient flow, while the Metropolis-Hastings accept/reject mechanism improves the mixing and convergence properties of this random walk.

Several extensions of MALA have been proposed recently by several authors, including fMALA (fast MALA), AMALA (anisotropic MALA), MMALA (manifold MALA), position-dependent MALA (PMALA), ...

MALA and these extensions have demonstrated to represent very efficient alternative for sampling from high dimensional distributions. We therefore need to adapt these methods to general mixed effects models.

3.5. Parameter estimation

The Stochastic Approximation Expectation Maximization (SAEM) algorithm has shown to be extremely efficient for maximum likelihood estimation in incomplete data models, and particularly in mixed effects models for estimating the population parameters. However, there are several practical situations for which extensions of SAEM are still needed:

High dimensional model: a complex physiological model may have a large number of parameters (in the order of 100). Then several problems arise:

- when most of these parameters are associated with random effects, the MCMC algorithm should be able to sample, for each of the N individuals, parameters from a high dimensional distribution. Efficient MCMC methods for high dimensions are then required.

- Practical identifiability of the model is not ensured with a limited amount of data. In other words, we cannot expect to be able to properly estimate all the parameters of the model, including the fixed effects and the variance-covariance matrix of the random effects. Then, some random effects should be removed, assuming that some parameters do not vary in the population. It may also be necessary to fix the value of some parameters (using values from the literature for instance). The strategy to decide which parameters should be fixed and which random effects should be removed remains totally empirical. XPOP aims to develop a procedure that will help the modeller to take such decisions.

Large number of covariates: the covariate model aims to explain part of the inter-patient variability of some parameters. Classical methods for covariate model building are based on comparisons with respect to some criteria, usually derived from the likelihood (AIC, BIC), or some statistical test (Wald test, LRT, etc.). In other words, the modelling procedure requires two steps: first, all possible models are fitted using some estimation procedure (e.g. the SAEM algorithm) and the likelihood of each model is computed using a numerical integration procedure (e.g. Monte Carlo Importance Sampling); then, a model selection procedure chooses the "best" covariate model. Such a strategy is only possible with a reduced number of covariates, i.e., with a "small" number of models to fit and compare.

As an alternative, we are thinking about a Bayesian approach which consists of estimating simultaneously the covariate model and the parameters of the model in a single run. An (informative or uninformative) prior is defined for each model by defining a prior probability for each covariate to be included in the model. In other words, we extend the probabilistic model by introducing binary variables that indicate the presence or absence of each covariate in the model. Then, the model selection procedure consists of estimating and maximizing the conditional distribution of this sequence of binary variables. Furthermore, a probability can be associated to any of the possible covariate models.

This conditional distribution can be estimated using an MCMC procedure combined with the SAEM algorithm for estimating the population parameters of the model. In practice, such an approach can only deal with a limited number of covariates since the dimension of the probability space to explore increases exponentially with the number of covariates. Consequently, we would like to have methods able to find a small number of variables (from a large starting set) that influence certain parameters in populations of individuals. That means that, instead of estimating the conditional distribution of all the covariate models as described above, the algorithm should focus on the most likely ones.

Fixed parameters: it is quite frequent that some individual parameters of the model have no random component and are purely fixed effects. Then, the model may not belong to the exponential family anymore and the original version of SAEM cannot be used as it is. Several extensions exist:

- introduce random effects with decreasing variances for these parameters,

- introduce a prior distribution for these fixed effects,

- apply the stochastic approximation directly on the sequence of estimated parameters, instead of the sufficient statistics of the model.

None of these methods always work correctly. Furthermore, what are the pros and cons of these methods is not clear at all. Then, developing a robust methodology for such model is necessary.

Convergence toward the global maximum of the likelihood: convergence of SAEM can strongly depend on thie initial guess when the observed likelihood has several local maxima. A kind of simulated annealing version of SAEM was previously developed and implemented in MONOLIX. The method works quite well in most situations but there is no theoretical justification and choosing the settings of this algorithm (i.e. how the temperature decreases during the iterations) remains empirical. A precise analysis of the algorithm could be very useful to better understand why it "works" in practice and how to optimize it.

Convergence diagnostic: Convergence of SAEM was theoretically demonstrated under very general hypothesis. Such result is important but of little interest in practice at the time to use SAEM in a finite amount of time, i.e. in a finite number of iterations. Some qualitative and quantitative criteria should be defined in order to both optimize the settings of the algorithm, detect a poor convergence of SAEM and evaluate the quality of the results in order to avoid using them unwisely.

3.6. Model building

Defining an optimal strategy for model building is far from easy because a model is the assembled product of numerous components that need to been evaluated and perhaps improved: the structural model, residual error model, covariate model, covariance model, etc.

How to proceed so as to obtain the best possible combination of these components? There is no magic recipe but an effort will be made to provide some qualitative and quantitative criteria in order to help the modeller for building his model.

The strategy to take will mainly depend on the time we can dedicate to building the model and the time required for running it. For relatively simple models for which parameter estimation is fast, it is possible to fit many models and compare them. This can also be done if we have powerful computing facilities available (e.g., a cluster) allowing large numbers of simultaneous runs.

However, if we are working on a standard laptop or desktop computer, model building is a sequential process in which a new model is tested at each step. If the model is complex and requires significant computation time (e.g., when involving systems of ODEs), we are constrained to limit the number of models we can test in a reasonable time period. In this context, it also becomes important to carefully choose the tasks to run at each step.

3.7. Model evaluation

Diagnostic tools are recognized as an essential method for model assessment in the process of model building. Indeed, the modeler needs to confront "his" model with the experimental data before concluding that this model is able to reproduce the data and before using it for any purpose, such as prediction or simulation for instance.

The objective of a diagnostic tool is twofold: first we want to check if the assumptions made on the model are valid or not ; then, if some assumptions are rejected, we want to get some guidance on how to improve the model.

As is the usual case in statistics, it is not because this "final" model has not been rejected that it is necessarily the "true" one. All that we can say is that the experimental data does not allow us to reject it. It is merely one of perhaps many models that cannot be rejected.

Model diagnostic tools are for the most part graphical, i.e., visual; we "see" when something is not right between a chosen model and the data it is hypothesized to describe. These diagnostic plots are usually based on the empirical Bayes estimates (EBEs) of the individual parameters and EBEs of the random effects: scatterplots of individual parameters versus covariates to detect some possible relationship, scatterplots of pairs of random effects to detect some possible correlation between random effects, plot of the empirical distribution of the random effects (boxplot, histogram,...) to check if they are normally distributed, ...

The use of EBEs for diagnostic plots and statistical tests is efficient with rich data, i.e. when a significant amount of information is available in the data for recovering accurately all the individual parameters. On the contrary, tests and plots can be misleading when the estimates of the individual parameters are greatly shrunk.

We propose to develop new approaches for diagnosing mixed effects models in a general context and derive formal and unbiased statistical tests for testing separately each feature of the model.

3.8. Missing data

The ability to easily collect and gather a large amount of data from different sources can be seen as an opportunity to better understand many processes. It has already led to breakthroughs in several application areas. However, due to the wide heterogeneity of measurements and objectives, these large databases often exhibit an extraordinary high number of missing values. Hence, in addition to scientific questions, such data also present some important methodological and technical challenges for data analyst.

Missing values occur for a variety of reasons: machines that fail, survey participants who do not answer certain questions, destroyed or lost data, dead animals, damaged plants, etc. Missing values are problematic since most statistical methods can not be applied directly on a incomplete data. Many progress have been made to properly handle missing values. However, there are still many challenges that need to be addressed in the future, that are crucial for the users.

- State of arts methods often consider the case of continuous or categorical data whereas real data are very often mixed. The idea is to develop a multiple imputation method based on a specific principal component analysis (PCA) for mixed data. Indeed, PCA has been used with success to predict (impute) the missing values. A very appealing property is the ability of the method to handle very large matrices with large amount of missing entries.

- The asymptotic regime underlying modern data is not any more to consider that the sample size increases but that both number of observations and number of variables are very large. In practice first experiments showed that the coverage properties of confidence areas based on the classical

methods to estimate variance with missing values varied widely. The asymptotic method and the bootstrap do well in low-noise setting, but can fail when the noise level gets high or when the number of variables is much greater than the number of rows. On the other hand, the jackknife has good coverage properties for large noisy examples but requires a minimum number of variables to be stable enough.

- Inference with missing values is usually performed under the assumption of "Missing at Random" (MAR) values which means that the probability that a value is missing may depend on the observed data but does not depend on the missing value itself. In real data and in particular in data coming from clinical studies, both "Missing Non at Random" (MNAR) and MAR values occur. Taking into account in a proper way both types of missing values is extremely challenging but is worth investigating since the applications are extremely broad.

It is important to stress that missing data models are part of the general incomplete data models addressed by XPOP. Indeed, models with latent variables (i.e. non observed variables such as random effects in a mixed effects model), models with censored data (e.g. data below some limit of quantification) or models with dropout mechanism (e.g. when a subject in a clinical trial fails to continue in the study) can be seen as missing data models.

4. Application Domains

4.1. Population pharmacometrics

Pharmacometrics involves the analysis and interpretation of data produced in pre-clinical and clinical trials. Population pharmacokinetics studies the variability in drug exposure for clinically safe and effective doses by focusing on identification of patient characteristics which significantly affect or are highly correlated with this variability. Disease progress modeling uses mathematical models to describe, explain, investigate and predict the changes in disease status as a function of time. A disease progress model incorporates functions describing natural disease progression and drug action.

The model based drug development (MBDD) approach establishes quantitative targets for each development step and optimizes the design of each study to meet the target. Optimizing study design requires simulations, which in turn require models. In order to arrive at a meaningful design, mechanisms need to be understood and correctly represented in the mathematical model. Furthermore, the model has to be predictive for future studies. This requirement precludes all purely empirical modeling; instead, models have to be mechanistic.

In particular, physiologically based pharmacokinetic models attempt to mathematically transcribe anatomical, physiological, physical, and chemical descriptions of phenomena involved in the ADME (Absorption - Distribution - Metabolism - Elimination) processes. A system of ordinary differential equations for the quantity of substance in each compartment involves parameters representing blood flow, pulmonary ventilation rate, organ volume, etc.

The ability to describe variability in pharmacometrics model is essential. The nonlinear mixed-effects modeling approach does this by combining the structural model component (the ODE system) with a statistical model, describing the distribution of the parameters between subjects and within subjects, as well as quantifying the unexplained or residual variability within subjects.

4.2. Precision medicine and pharmacogenomics

Pharmacogenomics involves using an individual's genome to determine whether or not a particular therapy, or dose of therapy, will be effective. Indeed, people's reaction to a given drug depends on their physiological state and environmental factors, but also to their individual genetic make-up.

Precision medicine is an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person. While some advances in precision medicine have been made, the practice is not currently in use for most diseases.

Currently, in the traditional population approach, inter-individual variability in the reaction to drugs is modeled using covariates such as weight, age, sex, ethnic origin, etc. Genetic polymorphisms susceptible to modify pharmacokinetic or pharmacodynamic parameters are much harder to include, especially as there are millions of possible polymorphisms (and thus covariates) per patient.

The challenge is to determine which genetic covariates are associated to some PKPD parameters and/or implicated in patient responses to a given drug.

Another problem encountered is the dependence of genes, as indeed, gene expression is a highly regulated process. In cases where the explanatory variables (genomic variants) are correlated, Lasso-type methods for model selection are thwarted.

4.3. Biology - Intracellular processes

Significant cell-to-cell heterogeneity is ubiquitously-observed in isogenic cell populations. Cells respond differently to a same stimulation. For example, accounting for such heterogeneity is essential to quantitatively understand why some bacteria survive antibiotic treatments, some cancer cells escape drug-induced suicide, stem cell do not differentiate, or some cells are not infected by pathogens.

The origins of the variability of biological processes and phenotypes are multifarious. Indeed, the observed heterogeneity of cell responses to a common stimulus can originate from differences in cell phenotypes (age, cell size, ribosome and transcription factor concentrations, etc), from spatio-temporal variations of the cell environments and from the intrinsic randomness of biochemical reactions. From systems and synthetic biology perspectives, understanding the exact contributions of these different sources of heterogeneity on the variability of cell responses is a central question.

5. Highlights of the Year

5.1. Highlights of the Year

R Foundation

Julie Josse has been elected member of the R Foundation for Statistical Computing.

mlxR 3.1

mlxR 3.1 available on CRAN

6. New Software and Platforms

6.1. mlxR

A R package for the simulation and visualization of longitudinal data. The models are encoded using the model coding language 'Mlxtran', automatically converted into C++ codes, compiled on the fly and linked to R using the 'Rcpp' package. That allows one to implement very easily complex ODE-based models and complex statistical models, including mixed effects models, for continuous, count, categorical, and time-to-event data.

6.2. FactoMineR

A R package dedicated to principal component methods (PCA, Correspondence Analysis for contingency tables, Multiple Correspondence Analysis for categorical data, MFA for multi-blocks data). Google users group and you-tube videos available.

6.3. missMDA

A R package to perform principal component methods (PCA, MCA, MFA) with missing values and to impute continuous, categorical and mixed data. Multiple imputation is available.

6.4. denoiseR

A R package that approximates a low-rank matrix from noisy data (Gaussian and Poisson Noise). Singular values shrinkage methods are implemented.

7. New Results

7.1. Identifiability in mixed effects models

We considered the question of model identifiability within the context of nonlinear mixed effects models. Although there has been extensive research in the area of fixed effects models, much less attention has been paid to random effects models.

In this context we distinguish between theoretical identifiability, in which different parameter values lead to non-identical probability distributions, structural identifiability which concerns the algebraic properties of the structural model, and practical identifiability, whereby the model may be theoretically identifiable but the design of the experiment may make parameter estimation difficult and imprecise.

We have explored a number of pharmacokinetic models which are known to be non-identifiable at an individual level but can become identifiable at the population level if a number of specific assumptions on the probabilistic model hold. Essentially if the probabilistic models are different, even though the structural models are non-identifiable, then they will lead to different likelihoods. The findings are supported through simulations.

7.2. Enhanced Method for Diagnosing Pharmacometric Models

For nonlinear mixed-effects pharmacometric models, diagnostic approaches often rely on individual parameters, also called empirical Bayes estimates (EBEs), estimated through maximizing conditional distributions. When individual data are sparse, the distribution of EBEs can "shrink" towards the same population value, and as a direct consequence, resulting diagnostics can be misleading.

Instead of maximizing each individual conditional distribution of individual parameters, we propose to randomly sample them in order to obtain values better spread out over the marginal distribution of individual parameters.

We have evaluated, through diagnostic plots and statistical tests, hypothesis related to the distribution of the individual parameters and shown that the proposed method leads to more reliable results than using the EBEs. In particular, diagnostic plots are more meaningful, the rate of type I error is correctly controlled and its power increases when the degree of misspecification increases. An application to the warfarin pharmacokinetic data confirms the interest of the approach for practical applications.

7.3. A Shrinkage-Thresholding Metropolis Adjusted Langevin Algorithm for Bayesian Variable Selection

We have introduced a new Markov Chain Monte Carlo method for Bayesian variable selection in high dimensional settings. The algorithm is a Hastings-Metropolis sampler with a proposal mechanism which combines a Metropolis Adjusted Langevin (MALA) step to propose local moves associated with a shrinkage-thresholding step allowing to propose new models.

The geometric ergodicity of this new trans-dimensional Markov Chain Monte Carlo sampler was established. An extensive numerical experiment, on simulated and real data, illustrates the performance of the proposed algorithm in comparison with some more classical trans-dimensional algorithms

7.4. Maximum likelihood estimation of a low-order building model

Our objective was to investigate the accuracy of the estimates learned with an open loop model of a building whereas the data is actually collected in closed loop, which corresponds to the true exploitation of buildings. We have proposed a simple model based on an equivalent RC network whose parameters are physically interpretable. We also described the maximum likelihood estimation of these parameters by the EM algorithm, and derived their statistical properties.

The numerical experiments clearly show the potential of the method, in terms of accuracy and robustness. We have emphasized the fact that the estimations are linked to the generating process for the observations, which includes the command system. For instance, the features of the building are correctly estimated if there is a significant gap between the heating and cooling setpoint.

7.5. LP-convergence of a Girsanov theorem based particle filter

We have analyzed the LP-convergence of a previously proposed Girsanov theorem based particle filter for discretely observed stochastic differential equation (SDE) models. We proved the convergence of the algorithm with the number of particles tending to infinity by requiring a moment condition and a step-wise initial condition boundedness for the stochastic exponential process giving the likelihood ratio of the SDEs. The practical implications of the condition are illustrated with an Ornstein–Uhlenbeck model and with a non-linear Bene's model.

7.6. Adaptive estimation in the nonparametric random coefficients binary choice model by needlet thresholding

In the random coefficients binary choice model, a binary variable equals 1 iff an index $X^T\beta$ is positive. The vectors X and β are independent and belong to the sphere S^{d-1} in R^d. We have proven lower bounds on the minimax risk for estimation of the density f_β over Besov bodies where the loss is a power of the $L^p(S^{d-1})$ norm for $1 \leq p \leq \infty$. We have shown that a hard thresholding estimator based on a needlet expansion with data-driven thresholds achieves these lower bounds up to logarithmic factors.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contract with Industry

Contract with Lixoft

9. Partnerships and Cooperations

9.1. European Initiatives

9.1.1. FP7 & H2020 Projects

The Drug Disease Model Resources (DDMoRe) consortium will build and maintain a universally applicable, open source, model-based framework, intended as the gold standard for future collaborative drug and disease modeling and simulation.

The DDMoRe project is supported by the Innovative Medicines Initiative (IMI), a large-scale public-private partnership between the European Union and the pharmaceutical industry association EFPIA.

Marc Lavielle was leader of WP6: "New tools for Model Based Drug Development".

DDMoRe website: http://www.ddmore.eu

Duration: 2010 - 2016

Project members: Uppsala Universitet, Sweden; University of Navarra, Spain; Universiteit Leiden, Netherlands; Université Paris Diderot, France; Universita degli Studi di Pavia, Italy; UCB Pharma , Belgium; Simcyp, UK; Pfizer, UK; Optimata , Israel; Novo Nordisk , Denmark; Novartis, Switzerland; Merck Serono, Switzerland; Takeda, Switzerland; Mango Business Solutions , UK; Lixoft, France; Interface Europe, Belgium; Institut de Recherches Internationales Servier, France; Inria, France; GlaxoSmithKline Research and Development, UK; Freie Universitat Berlin, Germany; F. Hoffmann - La Roche , Switzerland; EMBL - European Bioinformatics Institute, UK; Eli Lilly , UK; Cyprotex Discovery, UK; Consiglio Nazionale delle Ricerche, Italy; AstraZeneca, Sweden.

9.2. International Initiatives

9.2.1. *Informal International Partners*

Marc Lavielle is Adjunct Professor at the Faculty of Pharmacy of Florida University.

Marc Lavielle is Adjunct Professor at the Faculty of Pharmacy of Buffalo University.

Julie Josse collaborates with Susan Holmes, Stanford University.

Eric Moulines regularly collaborates with Sean P. Meyn, University of Florida.

9.3. International Research Visitors

9.3.1. *Visits of International Scientists*

Ricardo Rios, Universidad Central de Venezuela, Caracas: September 2016.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. *Scientific Events Organisation*

10.1.1.1. *General Chair, Scientific Chair*

Julie josse was chair of useR!2016, Stanford, CA, USA, July 2016. http://user2016.org.

10.1.2. *Scientific Expertise*

Marc Lavielle is member of the Scientific Committee of the High Council for Biotechnologies

10.1.3. *Research administration*

Marc Lavielle is member of

- the Scientific Programming Committee (CPS) of the Institute Henri Poincaré (IHP),
- the Executive Board (CA) of SMAI.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master : Julie Josse, Statistics with R, 48, M2, X-HEC

Master : Eric Moulines, Regresson models, 36, M2, X-HEC

Engineering School : Eric Moulines, Statistics, 36, 2A, X

Engineering School : Eric Moulines, Markov Chains, 36, 3A, X

Engineering School : Erwan Le Pennec, Statistics, 36, 2A, X

Engineering School : Erwan Le Pennec, Statistical Learning, 36, 3A, X

Engineering School : Marc Lavielle, Time Series, 24, 3A, X

10.2.2. Supervision

PhD in progress : Nicolas Brosse, September 2016, Eric Moulines

PhD in progress : Geneviève Robin, September 2016, Julie Josse

PhD in progress : Belhal Karimi, October 2016, Marc Lavielle and Eric Moulines

PhD in progress : Marine Zulian, October 2016, Marc Lavielle

11. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] H. BRAHAM, S. B. JEMAA, G. FORT, E. MOULINES, B. SAYRAC.*Fixed Rank Kriging for Cellular Coverage Analysis*, in "IEEE Transactions on Vehicular Technology", 2016, 11 [*DOI :* 10.1109/TVT.2016.2599842], https://hal.inria.fr/hal-01418961.

[2] H. BRAHAM, S. B. JEMAA, G. FORT, E. MOULINES, B. SAYRAC.*Spatial Prediction Under Location Uncertainty in Cellular Networks*, in "IEEE Transactions on Wireless Communications", 2016, vol. 15, p. 7633 - 7643 [*DOI :* 10.1109/TWC.2016.2605676], https://hal.inria.fr/hal-01419051.

[3] M. JALA, C. C. LEVY LEDUC, E. MOULINES, E. CONIL, J. WIART.*Sequential design of computer experiments for the assessment of fetus exposure to electromagnetic fields*, in "Technometrics", January 2016, vol. 58, n⁰ 1, p. 30–42 [*DOI :* 10.1080/00401706.2014.979951], https://hal.archives-ouvertes.fr/hal-01418859.

[4] M. LAVIELLE, L. AARONS.*What do we mean by identiability in mixed effects models?*, in "Journal of Pharmacokinetics and Pharmacodynamics", September 2016, https://hal.archives-ouvertes.fr/hal-01365535.

[5] M. LAVIELLE, B. RIBBA.*Enhanced Method for Diagnosing Pharmacometric Models: Random Sampling from Conditional Distributions*, in "Pharmaceutical Research / Pharmaceutical Research (Dordrecht)", 2016 [*DOI :* 10.1007/s11095-016-2020-3], https://hal.archives-ouvertes.fr/hal-01365532.

[6] A. SCHRECK, G. FORT, S. LE CORFF, E. MOULINES.*A Shrinkage-Thresholding Metropolis Adjusted Langevin Algorithm for Bayesian Variable Selection*, in "IEEE Journal of Selected Topics in Signal Processing", 2016, vol. 10, p. 366 - 375 [*DOI :* 10.1109/JSTSP.2015.2496546], https://hal.inria.fr/hal-01418960.

International Conferences with Proceedings

[7] P. ILHE, F. ROUEFF, E. MOULINES, A. SOULOUMIAC.*Nonparametric estimation of a shot-noise process*, in "SSP 16 - Statistical Signal Processing Workshop", Palma de Mallorca, Spain, IEEE Signal Processing Society, June 2016 [*DOI :* 10.1109/SSP.2016.7551709], https://hal.inria.fr/hal-01418963.

[8] J. LAFOND, H.-T. WAI, E. MOULINES.*D-FW: Communication efficient distributed algorithms for high-dimensional sparse optimization*, in "International Conference on Acoustics, Speech and Signal Processing", Shangai, China, March 2016, p. 4144 - 4148 [*DOI :* 10.1109/ICASSP.2016.7472457], https://hal.inria.fr/hal-01419048.

[9] T. NABIL, E. MOULINES, F. ROUEFF, J.-M. JICQUEL, A. GIRARD.*Maximum likelihood estimation of a low-order building model*, in "EUSIPCO 2016: 2016 24th European Signal Processing Conference", Budapest, Hungary, August 2016 [*DOI :* 10.1109/EUSIPCO.2016.7760339], https://hal.inria.fr/hal-01419050.

[10] S. SÄRKKÄ, E. MOULINES.*On the L<inf>P</inf>-convergence of a Girsanov theorem based particle filter*, in "ICASSP 2016 - International Conference on Acoustics, Speech and Signal Processing", Shangai, China, March 2016, p. 3989 - 3993 [*DOI :* 10.1109/ICASSP.2016.7472426], https://hal.inria.fr/hal-01419046.

Other Publications

[11] E. GAUTIER, E. LE PENNEC.*Adaptive estimation in the nonparametric random coefficients binary choice model by needlet thresholding*, September 2016, working paper or preprint, https://hal.inria.fr/inria-00601274.

[12] M. LAVIELLE, L. AARONS.*What do we mean by identifiability in mixed effects models?*, January 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01251986.

[13] L. MONTUELLE, E. LE PENNEC.*PAC-Bayesian aggregation of affine estimators*, October 2016, working paper or preprint, https://hal.inria.fr/hal-01070805.