



RESEARCH CENTER
Lille - Nord Europe

FIELD

Activity Report 2014

Section Application Domains

Edition: 2015-03-24

ALGORITHMICS, PROGRAMMING, SOFTWARE AND ARCHITECTURE	
1. ATEAMS Project-Team (section vide)	4
2. DREAMPAL Team (section vide)	5
APPLIED MATHEMATICS, COMPUTATION AND SIMULATION	
3. DOLPHIN Project-Team	6
4. MEPHYSTO Team	8
5. MODAL Project-Team	10
6. NON-A Project-Team	11
7. SEQUEL Project-Team	14
DIGITAL HEALTH, BIOLOGY AND EARTH	
8. BONSAI Project-Team	17
NETWORKS, SYSTEMS AND SERVICES, DISTRIBUTED COMPUTING	
9. FUN Project-Team (section vide)	18
10. RMOD Project-Team	19
11. SPIRALS Team	20
PERCEPTION, COGNITION AND INTERACTION	
12. LINKS Team	21
13. MAGNET Team	22
14. MINT Project-Team	23

ATEAMS Project-Team (section vide)

DREAMPAL Team (section vide)

DOLPHIN Project-Team

4. Application Domains

4.1. Academic Benchmark Problems

- ρMNK -landscapes constitute a problem-independent model used for constructing multiobjective multimodal landscapes with objective correlation. They extend single-objective NK-landscapes [59] and multiobjective NK-landscapes with independent objective functions [54]. The four parameters defining a ρMNK -landscape are: (i) the size of (binary string) solutions N , (ii) the variable correlation $K < N$, (iii) the number of objective functions M , and (iv) the correlation coefficient ρ . A number of problem instances and an instance generator are available at the following URL: <http://mocobench.sf.net/>.
- The Unconstrained Binary Quadratic Programming (UBQP) problem is known to be a unified modeling and solution framework for many combinatorial optimization problems [60]. Given a collection of n items such that each pair of items is associated with a profit value that can be positive, negative or zero, UBQP seeks a subset of items that maximizes the sum of their paired values. We proposed an extension of the single-objective UBQP to the multiobjective case (mUBQP), where multiple objectives are to be optimized simultaneously. We showed that the mUBQP problem is both NP-hard and intractable. Some problem instances with different characteristics and an instance generator are also available at the following URL: <http://mocobench.sf.net/>.

4.2. Transportation and logistics

- **Scheduling problems under uncertainty:** The flow-shop scheduling problem is one of the most well-known problems from scheduling. However, most of the works in the literature use a deterministic single-objective formulation. In general, the minimized objective is the total completion time (makespan). Many other criteria may be used to schedule tasks on different machines: maximum tardiness, total tardiness, mean job flowtime, number of delayed jobs, maximum job flowtime, etc. In the DOLPHIN project, a bi-criteria model, which consists in minimizing the makespan and the total tardiness, is studied. A bi-objective flow-shop problem with uncertainty on the duration, minimizing in addition the maximum tardiness, is also studied. It allows us to develop and test multi-objective (and not only bi-objective) optimization methods under uncertainty.
- **Routing problems under uncertainty:** The vehicle routing problem (VRP) is a well-known problem and it has been studied since the end of the fifties. It has a lot of practical applications in many industrial areas (ex. transportation, logistics, etc). Existing studies of the VRP are almost all concerned with the minimization of the total distance only. The model studied in the DOLPHIN project introduces a second objective, whose purpose is to balance the length of the tours. This new criterion is expressed as the minimization of the difference between the length of the longest tour and the length of the shortest tour. Uncertainty on the demands has also been introduced in the model.

4.3. Bioinformatics and Health care

Bioinformatic research is a great challenge for our society and numerous research entities of different specialities (biology, medical or information technology) are collaborating on specific themes.

4.3.1. Genomic and post-genomic studies

Previous studies of the DOLPHIN project mainly deal with genomic and postgenomic applications. These have been realized in collaboration with academic and industrial partners (IBL: Biology Institute of Lille; IPL: Pasteur Institute of Lille; IT-Omics firm).

First, genomic studies aim at analyzing genetic factors which may explain multi-factorial diseases such as diabetes, obesity or cardiovascular diseases. The scientific goal was to formulate hypotheses describing associations that may have any influence on diseases under study.

Secondly, in the context of post-genomic, a very large amount of data are obtained thanks to advanced technologies and have to be analyzed. Hence, one of the goals of the project was to develop analysis methods in order to discover knowledge in data coming from biological experiments.

These problems can be modeled as classical data mining tasks (Association rules, feature selection). As the combinatoric of such problems is very high and the quality criteria not unique, we proposed to model these problems as multi-objective combinatorial optimization problems. Evolutionary approaches have been adopted in order to cope with large scale problems.

Nowadays the technology is still going fast and the amount of data increases rapidly. Within the collaboration with Genes Diffusion, specialized in genetics and animal reproduction for bovine, swine, equine and rabbit species, we study combinations of Single Nucleotide Polymorphisms (SNP) that can explain some phenotypic characteristics. Therefore feature selection for regression is addressed using metaheuristics.

4.3.2. Optimization for health care

The collaboration with the Alicante company, a major actor in the hospital decision making, deals with knowledge extraction by optimization methods for improving the process of inclusion in clinical trials. Indeed, conducting a clinical trial, allowing for example to measure the effectiveness of a treatment, involves selecting a set of patients likely to participate to this test. Currently existing selection processes are far from optimal, and many potential patients are not considered. The objective of this collaboration consists in helping the practitioner to quickly determine if a patient is interesting for a clinical trial or not. Exploring different data sources (from a hospital information system, patient data...), a set of decision rules have to be generated. For this, approaches from multi-objective combinatorial optimization are implemented, requiring extensive work to model the problem, to define criteria optimization and to design specific optimization methods.

MEPHYSTO Team

4. Application Domains

4.1. Mechanics of heterogeneous media

The mechanics of heterogeneous materials aims at characterizing the macroscopic properties of heterogeneous materials using the properties of their constituents.

The homogenization theory is a natural tool for this task. In particular, for linear problems (linear conductivity or linear elasticity), the macroscopic properties are encoded into a single (conductivity or elasticity) homogenized tensor. The numerical approximation of this homogenized tensor is a typical objective of quantitative homogenization.

For nonlinear problems, such as rubber elasticity, the macroscopic properties are no longer characterized by a single tensor, but rather by a nonlinear energy density. Our aim is to relate qualitatively and quantitatively the (precise but unpractical) statistical physics picture to explicit macroscopic constitutive laws that can be used for practical purposes. This endeavor is relevant both in science and technology. The rigorous derivation of rubber elasticity from polymer-physics was indeed emphasized by John Ball as an important open problem of nonlinear elasticity in his survey [40] on the field. Its solution could shed light on some aspects of polymer-physics. The associated *ab initio* derivation of constitutive laws (as an alternative to phenomenological laws) would also be of interest to computational mechanics and rubber industry.

For this application domain, we work in close collaboration with physicists (François Lequeux, ESPCI) and researchers from mechanics and computational mechanics (Patrick Le Tallec, Ecole polytechnique).

4.2. Numerical simulation in heterogeneous media

Solving numerically PDEs in highly heterogeneous media is a problem encountered in many situations, such as the transport of pollutants or the design of oil extraction strategies in geological undergrounds. When such problems are discretized by standard numerical methods the number of degrees of freedom may become prohibitive in practice, whence the need for other strategies.

Numerical solution methods inspired by asymptotic analysis are among the very few feasible alternatives, and started fifteen years ago with the contributions of Hou and Wu [49], Arbogast [37] etc. We refer to [45], [57],[3] for a recent state of the art. Numerical homogenization methods usually amount to looking for the solution of the problem (1) in the form $u_\varepsilon(x) \simeq u_0(x) + \varepsilon \nabla u_0(x) \cdot \Phi(x, \frac{x}{\varepsilon})$, where $\Phi(x, \cdot)$ is a proxy for the corrector field computed locally at point $x \in D$ (in particular, one does not use explicitly that the problem is periodic so that the method can be used for more general coefficients) and u_0 is a function which does not oscillate at scale ε .

Relying on our quantitative insight in stochastic homogenization, a first task consists in addressing the three following prototypical academic examples: periodic, quasi-periodic, and stationary ergodic coefficients with short range dependence. The more ambitious challenge is to address more complex coefficients (of interest to practitioners), and design adaptive and efficient algorithms for diffusion in heterogeneous media.

4.3. Laser physics

Our contribution to the analysis of models in laser physics is motivated by the LabEx CEMPI (Centre Européen pour les Mathématiques, la Physique et leurs Interactions, a large eight-year research and training project approved by the French government in February 2012 as a "Laboratoire d'Excellence" and an initiative of mathematicians and physicists of the Université Lille 1). For this application domain, we work in close collaboration with physicists, which ensures our direct impact on these scientific issues. We focus on two applications: optical fibers and cold atoms.

In collaboration with physicists from the PhLAM laboratory in Lille, we aim at developing new techniques for the numerical integration of a family of 1D Schrödinger-like equations modelling the propagation of laser pulses in optical fibers. The questions arising are challenging since physicists would like to have fairly fast and cheap methods for their problems, with correct qualitative and quantitative behaviors. Another point is that they are interested in methods and codes that are able to handle different physical situations, hence different terms in the NLS equation. To meet these requirements, we will have to use numerical time-integration techniques such as splitting methods or exponential Runge-Kutta methods, space discretization techniques such as finite differences and fast Fourier transforms, and absorbent boundary conditions. Our goal, together with the physicists is to be able to reproduce numerically the results of the experiments they make in actual optical fibers, and then to be able to tune parameters numerically to get more insight into the appearance of rogue waves beyond the dispersive blowup phenomenon.

Recall that the Schrödinger equation also describes Bose-Einstein condensates. A second experimental team at PhLAM projects to study questions related to Anderson localization in such condensates. In fact, they will realize the “kicked rotor” (see [43]), which provides a paradigm for Anderson localization, in a Bose-Einstein condensate. We plan to collaborate with them on the theoretical underpinnings of their findings, which pose many challenging questions.

MODAL Project-Team

4. Application Domains

4.1. Application domains

Potential application areas of statistical modeling for heterogeneous data are extensive but some particular areas are identified. For historical reasons and considering the background of the team members, MODAL is mainly focused on biological applications where new challenges in high throughput technologies are opened. In addition, other secondary applications areas are considered in industry, retail, credit scoring and astronomy. Several contacts and collaborations are already established with some partners in these application areas and are described in Sections 7 and 8.

NON-A Project-Team

4. Application Domains

4.1. Networked Robots

Both economically and scientifically, cooperation in robot swarms represents an important issue since it concerns many service applications (health, handicap, urban transports...) and can increase the potential of sensor networks⁰. It involves several challenges such as:

- Because autonomy is a key for being able to increase the network size, maximize the autonomy of the robots in their different tasks of localization, motion, communication;
- Aiming at making 1+1 be more than 2, extend the global potential of the swarm by introducing collaboration (exchanging information with other robots) and cooperation (acting with other robots);
- Include time and energy saving considerations at the design stage. The self deployment of autonomous groups of mobile robots in an unknown environment (including different kinds of static or moving obstacles) involves localization, path planning and robust control problems. Both the control and signal aspects of our researches are oriented to solve some problems coming from - or taking advantage of - such collaboration frameworks. To mention a few:
- Localization using as few as possible landmarks and exteroceptive information by means of derivative estimates;
- Image-based sensing algorithms inspired by our multidimensional estimation techniques;
- Detection and adaptation to sudden loss of communication, time-varying topology, or communication delays;
- Robust, autonomous, energy-aware controllers based on either model-free or model-based techniques.

Several algorithms have already been applied to the control of formations of mobile robots: an illustrative platform is currently developed at EuraTechnologie center within the framework of Non-A⁰. They are now being extended to medical devices (such as wheelchairs) within the European project SYSIASS (see <http://www.sysiass.eu>), in collaboration with partners from hospital settings. Another future application concerns Wireless Sensor and Robot Networks (WSRN, Fig. 2), dedicated to the surveillance of zones, to the exploration of hostile areas, or to the supervision of large scale sensor networks. The main idea here is to integrate mobile nodes (the mobile robots) within the sensor network, allowing to overcome a sensor defection, to maintain the connectivity of the network, or to extend the coverage area during a random deployment. This involves consideration about mobile actuators within a mobile network of sensors and control networks (wireless) with strong constraints on the possibilities of communication in a noisy and non-homogeneous environment. This work is made in close collaboration with the Inria project-team POPS (Lille), which brings its expertise in terms of sensor networks. It takes place in the framework of the Inria ADT SENSAS and represents our contribution to the LABEX proposal ICON.

4.2. Nano/Macro machining

Nano machining

⁰Integrating wireless sensor networks and multi-robot systems increases the potential of the sensors: robots, in comparison, are resource-rich and can be involved in taking decisions and performing appropriate actions on themselves on sensors and/or the environment.

⁰"RobotCity" was exhibited for the first time during the opening ceremony held on April 6th, 2011

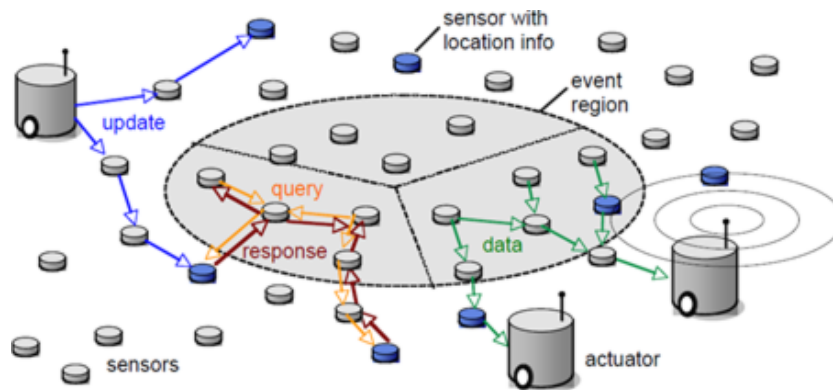


Figure 2. An illustration of collaboration in a Wireless Sensor and Robot Network.

Recent research investigations have reported the development of a number of process chains that are complementary to those used for batch manufacturing of Micro Electro Mechanical Systems (MEMS) and, at the same time, broaden the application domain of products incorporating micro and nano scale features. Such alternative process chains combine micro and nano structuring technologies for master making with replication techniques for high volume production such as injection moulding and roll-to-roll imprinting. In association with the Manufacturing Engineering Center of Cardiff, Arts et Metiers ParisTech center of Lille develops a new process chain for the fabrication of components with nano scale features. In particular, AFM probe-based nano mechanical machining is employed as an alternative master making technology to commonly used lithography-based processes (Fig.). Previous experimental studies demonstrated the potential of this approach for thermoplastic materials. Such a manufacturing route also represents an attractive prototyping solution to test the functionalities of components with nano scale features prior to their mass fabrication and, thus, to reduce the development time and cost of nano technology-enabled products. Application of our control and estimation techniques improves the trajectory tracking accuracy and the speed of the machining tools.

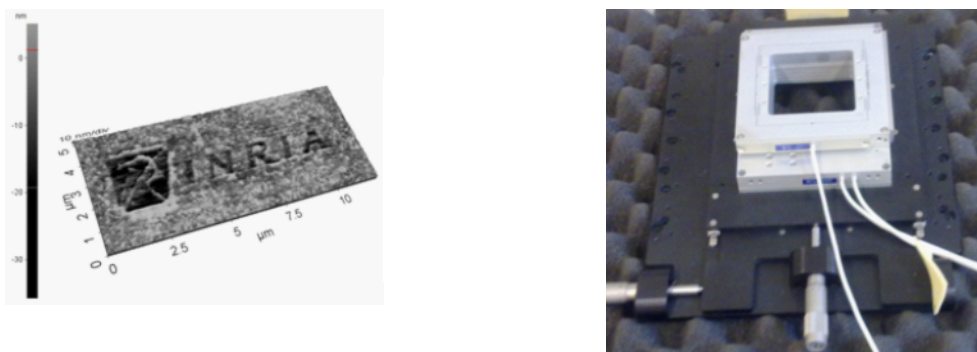


Figure 3. Left: A machined nano structure: $16 \mu\text{m} \times 8 \mu\text{m} \times \text{some nm}$. Right: Nano-positioning system available at Arts et Métiers ParisTech Lille ($75 \mu\text{m}$ range of motion).

Machining with industrial robots

Industrials are enthusiastic to replace machine-tools with industrial robots: compared to machine-tools, industrial articulated robots are very cheaper, more flexible, and exhibit more important workspaces. They can carry out machining applications like prototyping, cleaning and pre-machining of cast parts, as well as end-machining of middle tolerance parts. Such applications require high accuracy in the positioning and path tracking. Unfortunately, industrial robots have a low stiffness and are not that accurate⁰ and they deserve an increased quality of control. We deal with the modelling and the on-line identification of flexible-joint robot models. This can be used both for dynamic simulation and model-based control of industrial robots. We address the problem of real-time identification of the parameters involved in the dynamic linear model of an industrial robot axis. This is possible thanks to a special sensor developed by Arts et Métiers, subject to an EADS project within the FUI (Fonds Unique Interministériel). Control algorithms for other machining actuators such as active magnet bearings are also under study. Within the framework of LAGIS, we also consider the remote control of industrial robots (via internet or Wi-Fi links, for instance), which sets numerous problems in relation with the communication delays.

4.3. Multicell Chopper

On the basis of benchmarks developed at ECS-lab (ENSEA Cergy), we intend to work on the control and observation of serial and parallel multicell choppers, as well as more usual power converters. These power electronic systems associated with their respective loads are typical hybrid dynamical systems and many industrial and/or theoretical challenging problems occur. For example, in the industrial problem of power supply for a supercomputer, the parallel multicell chopper appears as a new solution particularly with respect to the power efficiency. Nevertheless, the observation and control of such hybrid dynamical systems is a difficult task, where non asymptotic estimation and control can be useful.

⁰Industrial robots were designed to realize repeatable tasks. The robot repeatability ranges typically from 0.03 to 0.1mm, but the accuracy is often measured to be within several millimetres. Due to their serial structure, articulated robot has lower stiffness (less than 1 N / mm) than classical machine-tools (greater than 50 N / mm). These poor accuracy and stiffness are caused by many factors, such as geometric parameter errors (manufacturing tolerances), wear of parts and components replacement, as well as flexibility of links and gear trains, gear backlashes, encoder resolution errors and thermal effects.

SEQUEL Project-Team

4. Application Domains

4.1. In Short

SEQUEL aims at solving problems of prediction, as well as problems of optimal and adaptive control. As such, the application domains are very numerous.

The application domains have been organized as follows:

- adaptive control,
- signal processing and functional prediction,
- web mining,
- computer games.

4.2. Adaptive Control

Adaptive control is an important application of the research being done in SEQUEL. Reinforcement learning (RL) precisely aims at controlling the behavior of systems and may be used in situations with more or less information available. Of course, the more information, the better, in which case methods of (approximate) dynamic programming may be used [47]. But, reinforcement learning may also handle situations where the dynamics of the system is unknown, situations where the system is partially observable, and non stationary situations. Indeed, in these cases, the behavior is learned by interacting with the environment and thus naturally adapts to the changes of the environment. Furthermore, the adaptive system may also take advantage of expert knowledge when available.

Clearly, the spectrum of potential applications is very wide: as far as an agent (a human, a robot, a virtual agent) has to take a decision, in particular in cases where he lacks some information to take the decision, this enters the scope of our activities. To exemplify the potential applications, let us cite:

- game software: in the 1990's, RL has been the basis of a very successful Backgammon program, TD-Gammon [53] that learned to play at an expert level by basically playing a very large amount of games against itself. Today, various games are studied with RL techniques.
- many optimization problems that are closely related to operation research, but taking into account the uncertainty, and the stochasticity of the environment: see the job-shop scheduling, or the cellular phone frequency allocation problems, resource allocation in general [47]
- we can also foresee that some progress may be made by using RL to design adaptive conversational agents, or system-level as well as application-level operating systems that adapt to their users habits.

More generally, these ideas fall into what adaptive control may bring to human beings, in making their life simpler, by being embedded in an environment that is made to help them, an idea phrased as "ambient intelligence".

- The sensor management problem consists in determining the best way to task several sensors when each sensor has many modes and search patterns. In the detection/tracking applications, the tasks assigned to a sensor management system are for instance:
 - detect targets,
 - track the targets in the case of a moving target and/or a smart target (a smart target can change its behavior when it detects that it is under analysis),
 - combine all the detections in order to track each moving target,
 - dynamically allocate the sensors in order to achieve the previous three tasks in an optimal way. The allocation of sensors, and their modes, thus defines the action space of the underlying Markov decision problem.

In the more general situation, some sensors may be localized at the same place while others are dispatched over a given volume. Tasking a sensor may include, at each moment, such choices as where to point and/or what mode to use. Tasking a group of sensors includes the tasking of each individual sensor but also the choice of collaborating sensors subgroups. Of course, the sensor management problem is related to an objective. In general, sensors must balance complex trade-offs between achieving mission goals such as detecting new targets, tracking existing targets, and identifying existing targets. The word “target” is used here in its most general meaning, and the potential applications are not restricted to military applications. Whatever the underlying application, the sensor management problem consists in choosing at each time an action within the set of available actions.

- sequential decision processes are also very well-known in economy. They may be used as a decision aid tool, to help in the design of social helps, or the implementation of plants (see [51], [50] for such applications).

4.3. Signal Processing

Applications of sequential learning in the field of signal processing are also very numerous. A signal is naturally sequential as it flows. It usually comes from the recording of the output of sensors but the recording of any sequence of numbers may be considered as a signal like the stock-exchange rates evolution with respect to time and/or place, the number of consumers at a mall entrance or the number of connections to a web site. Signal processing has several objectives: predict, estimate, remove noise, characterize or classify. The signal is often considered as sequential: we want to predict, estimate or classify a value (or a feature) at time t knowing the past values of the parameter of interest or past values of data related to this parameter. This is typically the case in estimation processes arising in dynamical systems.

Signals may be processed in several ways. One of the best-known way is the time-frequency analysis in which the frequencies of each signal are analyzed with respect to time. This concept has been generalized to the time-scale analysis obtained by a wavelet transform. Both analysis are based on the projection of the original signal onto a well-chosen function basis. Signal processing is also closely related to the probability field as the uncertainty inherent to many signals leads to consider them as stochastic processes: the Bayesian framework is actually one of the main frameworks within which signals are processed for many purposes. It is worth noting that Bayesian analysis can be used jointly with a time-frequency or a wavelet analysis. However, alternatives like belief functions came up these last years. Belief functions were introduced by Detspiter few decades ago and have been successfully used in the few past years in fields where probability had, during many years, no alternatives like in classification. Belief functions can be viewed as a generalization of probabilities which can capture both imprecision and uncertainty. Belief functions are also closely related to data fusion.

4.4. Web Mining

We work on the news/ad recommendation. These online learning algorithms reached a critical importance over the last few years due to these major applications. After designing a new algorithm, it is critical to be able to evaluate it without having to plug it into the real application in order to protect user experiences or/and the company’s revenue. To do this, people used to build simulators of user behaviors and try to achieve good performances against it. However designing such a simulator is probably much more difficult than designing the algorithm itself! An other common way to evaluate is to not consider the exploration/exploitation dilemma (also known as “Cold Start” for recommender systems). Lately data-driven methods have been developed. We are working on building automatic replay methodology with some theoretical guarantees. This work also exhibits strong link with the choice of the number of contexts to use with recommender systems wrt your audience.

An other point is that web sites must forecast Web page views in order to plan computer resource allocation and estimate upcoming revenue and advertising growth. In this work, we focus on extracting trends and seasonal patterns from page view series. We investigate Holt-Winters/ARIMA like procedures and some regularized models for making short-term prediction (3-6 weeks) wrt to logged data of several big media websites. We

work on some news event related webpages and we feel that kind of time series deserves a particular attention. Self-similarity is found to exist at multiple time scales of network traffic, and can be exploited for prediction. In particular, it is found that Web page views exhibit strong impulsive changes occasionally. The impulses cause large prediction errors long after their occurrences and can sometimes be predicted (*e.g.*, elections, sport events, editorial changes, holidays) in order to improve accuracies. It also seems that some promising model could arise from using global trends shift in the population.

4.5. Games

The problem of artificial intelligence in games consists in choosing actions of players in order to produce artificial opponents. Most games can be formalized as Markov decision problems, so they can be approached with reinforcement learning.

In particular, SEQUEL was a pioneer of Monte Carlo Tree Search, a technique that obtained spectacular successes in the game of Go. Other application domains include the game of poker and the Japanese card game of hanafuda.

BONSAI Project-Team

4. Application Domains

4.1. Sequence processing for Next Generation Sequencing

As said in the introduction of this document, biological sequence analysis is a foundation subject for the team. In the last years, sequencing techniques have experienced remarkable advances with Next Generation Sequencing (NGS), that allow for fast and low-cost acquisition of huge amounts of sequence data, and outperforms conventional sequencing methods. These technologies can apply to genomics, with DNA sequencing, as well as to transcriptomics, with RNA sequencing. They promise to address a broad range of applications including: Comparative genomics, individual genomics, high-throughput SNP detection, identifying small RNAs, identifying mutant genes in disease pathways, profiling transcriptomes for organisms where little information is available, researching lowly expressed genes, studying the biodiversity in metagenomics. From a computational point of view, NGS gives rise to new problems and gives new insight on old problems by revisiting them: Accurate and efficient remapping, pre-assembling, fast and accurate search of non exact but quality labelled reads, functional annotation of reads, ...

4.2. Noncoding RNA

Our expertise in sequence analysis also applies to noncoding RNA. Noncoding RNA plays a key role in many cellular processes. First examples were given by microRNAs (miRNAs) that were initially found to regulate development in *C. elegans*, or small nucleolar RNAs (snoRNAs) that guide chemical modifications of other RNAs in mammals. Hundreds of miRNAs are estimated to be present in the human genome, and computational analysis suggests that more than 20% of human genes are regulated by miRNAs. To go further in this direction, the 2007 ENCODE Pilot Project provides convincing evidence that the Human genome is pervasively transcribed, and that a large part of this transcriptional output does not appear to encode proteins. All those observations open a universe of “RNA dark matter” that must be explored. From a combinatorial point of view, noncoding RNAs are complex objects. They are single stranded nucleic acid sequences that can fold forming long-range base pairings. This implies that RNA structures are usually modelled by complex combinatorial objects, such as ordered labeled trees, graphs or arc-annotated sequences.

4.3. Genome structures

Our third application domain is concerned with the structural organization of genomes. Genome rearrangements are able to change genome architecture by modifying the order of genes or genomic fragments. The first studies were based on linkage maps and fifteen year old mathematical models. But the usage of computational tools was still limited due to the lack of data. The increasing availability of complete and partial genomes now offers an unprecedented opportunity to analyse genome rearrangements in a systematic way and gives rise to a wide spectrum of problems: Taking into account several kinds of evolutionary events, looking for evolutionary paths conserving common structure of genomes, dealing with duplicated content, being able to analyse large sets of genomes even at the intraspecific level, computing ancestral genomes and paths transforming these genomes into several descendant genomes.

4.4. Nonribosomal peptides

Lastly, the team has been developing for several years a tight collaboration with Probiogem lab on nonribosomal peptides, and has become a leader on that topic. Nonribosomal peptide synthesis produces small peptides not going through the central dogma. As the name suggests, this synthesis uses neither messenger RNA nor ribosome but huge enzymatic complexes called nonribosomal peptide synthetases (NRPSs). This alternative pathway is found typically in bacteria and fungi. It has been described for the first time in the 70's [14]. For the last decade, the interest in nonribosomal peptides and their synthetases has considerably increased, as witnessed by the growing number of publications in this field. These peptides are or can be used in many biotechnological and pharmaceutical applications (e.g. anti-tumors, antibiotics, immuno-modulators).

FUN Project-Team (section vide)

RMOD Project-Team

4. Application Domains

4.1. Programming Languages and Tools

Many of the results of RMoD are improving programming languages or development tools for such languages. As such the application domain of these results is as varied as the use of programming languages in general. Pharo, the language that RMoD develops, is used for a very broad range of applications. From pure research experiments to real world industrial use (the Pharo Consortium has over 10 company members) <http://consortium.pharo.org> Examples are web applications, server backends for mobile applications or even graphical tools and embedded applications.

4.2. Software Reengineering

Moose is a language-independent environment for reverse- and re-engineering complex software systems. Moose provides a set of services including a common meta-model, metrics evaluation and visualization. As such Moose is used for analysing software systems to support understanding and continuous development as well as software quality analysis.

SPIRALS Team

4. Application Domains

4.1. Introduction

Although our research is general enough to be applied to many application domains, we currently focus on applications and distributed services for the retail industry and for the digital home. These two application domains are supported by a strong expertise in mobile computing and in cloud computing that are the two main target environments on which our research prototypes are build, for which we are recognized, and for which we have already established strong collaborations with the industrial ecosystem.

4.2. Distributed software services for the retail industry

This application domain is developed in relation with the **PICOM** (*Pôle de compétitivité Industries du Commerce*) cluster. We have established strong collaborations with local companies in the context of former funded projects, such as **Macchiato**, which focused on the development of a new generation of mobile computing platforms for e-commerce. We are also involved in the Datalyse and OCCIware funded projects that define cloud computing environments with applications for the retail industry. Finally, our activities in terms of crowdsensing and data gathering on mobile devices with the APISENSE[®] platform share also applications for the retail industry.

4.3. Distributed software services for the digital home

We are developing new middleware solutions for the digital home, in particular through our long standing collaboration with Orange Labs. We are especially interested in developing energy management and saving solutions with the POWERAPI software library for distributed environments such the ones that equip digital homes. We are also working to bridge the gap between distributed services hosted on home gateways and distributed services hosted on the cloud to be able to smoothly transition between both environments. This work is especially conducted with the SALOON platform.

LINKS Team

3. Application Domains

3.1. Collective Intelligence

⁰ Links represented in the data are important for web users, who try to locate relevant information. They typically want to pose their queries locally and obtain the answers from both local and remote repositories. With the concept of linked data collections, the users are provided with a virtual collection of data and links. The answers to a query need to follow both explicit and implicit links to external repositories. Nevertheless, we argue that the benefits of links are not limited to casual users. In this paragraph, we briefly discuss two applications in which linked data collections need to be created. In the past decade, most of the enterprise data was proprietary, thus residing within the enterprise repository, along with the knowledge derived from that data. Today's enterprises need to face the problem of information explosion, due to the Internet's instability to rapidly convey large amounts of information throughout the world via end-user applications and tools. A linked data collection thus represents a virtual knowledge repository, in which relevant data is collected and meaningful mappings between this data and external world are inferred. A linked data collection would ease the task of expert users to (i) process data, metadata and knowledge that derives from machine to machine processing and eventually make sense of it; (ii) assemble and disassemble pieces of data, metadata and knowledge to create aggregate opinions or to disaggregate opinions in a way that turns to be useful in decision making; (iii) continuously learn from user feedback to produce better knowledge and to enhance the semantics of data processing.

3.2. Linked Bibliographic Collections

The second example concerns scientists who want to quickly inspect relevant literature and datasets. In such a case, local knowledge that comes from a local repository of publications belonging to a research institute (e.g. HAL) need to be integrated with other Web-based repositories, such as DBLP, Google Scholar, ResearchGate and even Wikipedia. Indeed, the local repository may be incomplete or contain semantic ambiguities, such as mistaken or missing conference venues, mistaken long names for the publication venues and journals, missing explanation of research keywords, and opaque keywords. A linked data collection would lead to build both implicit and explicit links by (i) cleaning the errors with links to correct data e.g. via mappings from HAL to DBLP for the publications errors, and via mappings from HAL to Wikipedia for opaque keywords, (ii) thoroughly complete the list of publications of the research institute, and (iii) support complex queries on the corrected data combined with external links. Links are thus useful in all scenarios, in which massive data need to be understood, analyzed and processed. Finally, they can be processed in highly distributed scenarios, such as in the cloud, where their evaluation can be run on many different sites.

⁰Collective intelligence is a shared or group intelligence that emerges from the collaboration and competition of many individuals and appears in consensus decision making" (from Wikipedia - Collective Intelligence)

MAGNET Team

4. Application Domains

4.1. Overview

Our main targeted applications are browsing, monitoring and mining in information networks. Such discovered structures would also be beneficial to predicting links between users and texts which is at the core of recommender systems. All the learning tasks considered in the project such as node clustering, node and link classification and link prediction are likely to yield important improvements in these applications. Application domains cover social networks for cultural data and e-commerce, and biomedical informatics.

MINT Project-Team

4. Application Domains

4.1. Next-generation desktop systems

The term *desktop system* refers here to the combination of a window system handling low-level graphics and input with a window manager and a set of applications that share a distinctive look and feel. It applies not only to desktop PCs but also to any other device or combination of devices supporting graphical interaction with multiple applications. Interaction with these systems currently rely on a small number of interaction primitives such as text input, pointing and activation as well as a few other basic gestures. This limited set of primitives is one reason the systems are simple to use. There is, however, a cost. Most simple combinations being already used, few remain to trigger and control innovative techniques that could facilitate task switching or data management, for example. Desktop systems are in dire need of additional interaction primitives, including gestural ones.

4.2. Ambient Intelligence

Ambient intelligence (AmI) refers to the concept of being surrounded by intelligent systems embedded in everyday objects [49]. Envisioned AmI environments are aware of human presence, adapt to users' needs and are capable of responding to indications of desire and possibly engaging in intelligent dialogue. Ambient Intelligence should be unobtrusive: interaction should be relaxing and enjoyable and should not involve a steep learning curve. Gestural interaction is definitely relevant in this context.

4.3. Serious Games

Serious game refers to techniques extensively used in computer games, that are being used for other purposes than gaming. Fields such as learning, use of Virtual Reality for rehabilitation, 3D interactive worlds for retail, art-therapy, are specific context with which the MINT group has scientific connection, and industrial contacts. This field of application is a good opportunity for us to test and transfer our scientific knowledge and results.