



RESEARCH CENTER  
Lille - Nord Europe

FIELD

Activity Report 2012

# Section Application Domains

Edition: 2013-04-24



1. ADAM Project-Team .....	4
2. ATEAMS Project-Team (section vide) .....	6
3. BONSAI Project-Team .....	7
4. DART Project-Team .....	8
5. DOLPHIN Project-Team .....	9
6. FUN Team (section vide) .....	13
7. MINT Project-Team .....	14
8. MODAL Project-Team .....	15
9. MOSTRARE Project-Team .....	16
10. NON-A Project-Team .....	17
11. RMOD Project-Team (section vide) .....	19
12. SEQUEL Project-Team .....	20
13. SHACRA Project-Team .....	23
14. SIMPAF Project-Team .....	26

## **ADAM Project-Team**

# **4. Application Domains**

## **4.1. Introduction**

The ADAM project-team targets the software engineering of adaptive service-oriented applications and middleware. The application domain covered by ADAM is broad and spans from distribution applications to middleware. In all these cases, adaptability is the property which is sought: applications and middleware must be adaptable to new execution contexts, they must react to changes in the environment and they must be able to discover and integrate new services.

The ADAM project-team produces software and middleware building blocks. This explains why the application domain is broad, yet targeting applications where adaptability is the key requirement. This includes electronic commerce, embedded systems, health care information systems, and terrestrial transport information systems. These domains are in direct relation with our currently funded activities. They act as testbeds for the solutions that we propose in terms of middleware services, middleware platforms, runtime kernels, component libraries, languages design or domain modeling.

## **4.2. Electronic Commerce**

Applications in the domain of electronic commerce are by essence distributed. They involve many different participants with heterogeneous information systems which cannot be changed. The challenge is then to provide an adaptation layer to be able to compose and let these systems interoperate. In the context of the ANR TLog SCOrWare, the ICT SOA4All and the FUI CAPPUCINO projects, our activities in this domain will aim at supporting service-oriented architectures. We want to have adaptive architectures that can be composed and orchestrated seamlessly. In this domain, the business relationship with customers is vital and many different usage scenarios must be supported. Customers are roaming, and the services must be kept operational across different devices. This puts some constraints on the server tier where technical services must be adapted to manage, for instance, long lasting transactions. The application server infrastructure must then provide a support for adapting technical services.

## **4.3. Embedded Systems**

Embedded systems form a domain where adaptation is a key requirement. The design and the implementation of modern embedded software uses advanced software engineering techniques such model-driven development or software component frameworks. In this domain, we are involved in several projects, such as the ANR TLog Flex-eWare, and the FUI MIND projects. Several challenges must be addressed here. For example, when a model-driven developed application is adapted, designers have to ensure that the models and the operational level are kept synchronized. The co-evolution of these two levels is one of the challenges that we are addressing. A second challenge is related to software components that need to be customized in order to fit the requirements imposed by constrained environments. It is, for example, a matter of providing component frameworks that can accommodate various granularities of services.

## **4.4. Health Care Information Systems**

Health care information systems form a third application domain in which the ADAM project-team is involved, for instance through demonstrators which will be implemented in the context of the ANR TLog FAROS project. The challenge here is to provide a distributed infrastructure where information will be available to medical staff wherever they are. This imposes to be able to provide this information on many different devices (from high resolution screens to embedded devices on the scene of an accident), while ensuring the privacy of the medical data of a patient (several level of data access must be granted depending on the categories of medical staff). Given the vital role of such an information system, we want to provide guarantees that the services will be highly available and trustworthy. We envision to provide a service-oriented architecture which will be extended to support software contracts and multi-scale environments.

## **4.5. Information Systems for Terrestrial Transport**

Information systems for terrestrial transport are also a domain that we are relying on, to apply our research activities in accordance with the ANR ARA REVE project and the INRETS collaboration. Applications are here characterized by frequent disconnections, poor quality network links, and high mobility. We want to provide an infrastructure where the technical services, and among others the communication services, can be adapted to support new requirements. One of the paths that we propose to investigate is to include such a scenario in the general context of the adaptiveness of component frameworks.

**ATEAMS Project-Team (section vide)**

## 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 experienced remarkable advances with NGS, that allows 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 allowing to gene expression analysis. 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 noncodingRNA analysis. Noncoding RNA genes play 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 acids 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 onto linkage maps and mathematical models appeared fifteen years ago. But the usage of computational tools was still limited because of 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 year a tight collaboration with Probiogem lab on nonribosomal peptides, and has became 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 [21]. 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).

## **DART Project-Team**

# **4. Application Domains**

## **4.1. Gaspard2 for avionic hybrid test platform design**

The emergence and the maturity of FPGA circuits for distributed and reconfigurable architectures offer the opportunity to explore real time problems in the field of avionic systems. FPGA becomes de facto a major processing element as same as general CPUs. As of now, the FPGA is widely used in the field of I/O component in order to connect the real equipment with the CPU host. Among the main features mapped into the FPGA in the original architecture, we quote the fast serial link and RAM IPs (Intellectual property) which are needed to ensure communication between CPU and FPGA. Additionally, the Base Time IP is needed for the global system synchronization. This minimal configuration based on FPGA can be duplicated several times and connected together to build bigger test system or a complete simulator. Eurocopter expectation for the above-described architecture is to prototype some models which can be eligible and relocated in the FPGA. The objective is to increase the performances of these models and to reduce the communication latencies by the means of embedding the different parts in the same chip. To do so, we studied in this first year a real avionic test loop in order to extract the complex models that will be implemented in the FPGA. Different hardware model configurations have been explored to reach an optimal well-balanced global system using the ML403 Virtex-4 Xilinx board. Different tradeoffs in terms of performance and resource occupation in the FPGA are obtained. Later, these results will be used for dynamically adapt the system functioning according to the available resources and performance requirements.

As a second part, we used the MARTE profile to represent an hybrid system (CPU/FPGA). In the MARTE specification, an application is a set of tasks connected through ports. Tasks are considered as mathematical functions reading data from their input ports and writing data on their output ports. This specification has been used to model the avionic test loop. In addition, MARTE allows describing the hardware architecture in a structural way. Typical components such as HwProcessor, HwFPGA and HwRAM can be specified with their non-functional properties. We used this subset of MARTE in order to represent an hybrid multiprocessor architecture. The main component of this architecture is composed of the Xeon-X3370 processor (multicore CPU) and the Virtex-4 Xilinx FPGA. Furthermore, MARTE provides the Allocate concept as well as the concept specially crafted for repetitive structures Distribute. This latter concept gives a way to express regular distribution of tasks onto a set of processors or FPGA resources. The mapping step relies on two types of distribution (timeScheduling and spatialDistribution) depending on the target hardware platform (CPU/FPGA). The different models of our avionic test loop can be mapped onto the host multicore processor, the embedded processor (Microblaze) or the hardware resources in the FPGA.

## **4.2. Electromagnetic modeling**

We collaborate with the L2EP specialized in electromagnetic modeling, on algorithms definition and the parallelization of their computations, especially on GPUs.

For the first point, we have designed a parallel version of the Finite Integration Technique (F.I.T). This is used to simulate electromagnetic phenomena. This technique is efficient if the mesh is generated by a regular hexahedron. Moreover the matrix system, obtained from a regular mesh can be exploited to use the parallel direct solver. In fact, in reordering the unknowns by the nested dissection method, it is possible to construct directly the lower triangular matrix with many processors without assembling the matrix system. During this year, we have used our parallel direct solver as a preconditionner for a sparse linear system coming from a FEM problem with a good efficiency[25].

For the second point, we have include our Gaspard2 generated code in Code\_CARMEL, a software for electromagnetic fields simulations. This GPGPU code is now robust enough to run most of the testbenches implemented inside this framework[23].



## DOLPHIN Project-Team

# 4. Application Domains

## 4.1. Academic generic problems

In this project, some well known optimization problems are re-visited in terms of multi-objective modelization and resolution:

- **Workshop optimization problems:**

Workshop optimization problems deal with optimizing the production. In this project, two specific problems are under study.

- **Flow-shop scheduling problem:** The flow-shop problem is one of the most well-known scheduling problems. However, most of the works in the literature use a mono-objective model. 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 tri-criteria flow-shop problem, minimizing in addition the maximum tardiness, is also studied. It allows us to develop and test multi-objective (and not only bi-objective) exact methods.
- **Cutting problems:** Cutting problems occur when pieces of wire, steel, wood, or paper have to be cut from larger pieces. The objective is to minimize the quantity of lost material. Most of these problems derive from the classical one-dimensional cutting-stock problem, which have been studied by many researchers. The problem studied by the DOLPHIN project is a two-dimensional bi-objective problem, where rotating a rectangular piece has an impact on the visual quality of the cutting pattern. First we have to study the structure of the cutting-stock problem when rotation is allowed, then we will develop a method dedicated to the bi-objective version of the problem.

- **Logistics and transportation problems:**

- **Packing problems:** In logistic and transportation fields, packing problems may be a major issue in the delivery process. They arise when one wants to minimize the size of a warehouse or a cargo, the number of boxes, or the number of vehicles used to deliver a batch of items. These problems have been the subjects of many papers, but only few of them study multi-objective cases, and to our knowledge, never from an exact point of view. Such a case occurs for example when some pairs of items cannot be packed in the same bin. The DOLPHIN project is currently studying the problem in its one-dimensional version. We plan to generalize our approach to two and three dimensional problems, and to more other conflict constraints, with the notion of *distance* between items.
- **Routing problems:** The vehicle routing problem (VRP) is a well-known problem and it has been studied since the end of the 50's. 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. As far as we know, this model is one of the pioneer work in the literature.

The second routing problem is a generalization of the covering tour problem (CTP). In the DOLPHIN project, this problem is solved as a bi-objective problem where a set of constraints are modeled as an objective. The two objectives are: i) minimization of the length of the tour; ii) minimization of the largest distance between a node to be covered and a visited node. As far as we know, this study is among the first works that tackle a classic mono-objective routing problem by relaxing constraints and building a more general MOP.

The third studied routing problem is the Ring Star Problem (RSP). This problem consists in locating a simple cycle through a subset of nodes of a graph while optimizing two kinds of costs. The first objective is to minimize a ring cost that is related to the length of the cycle. The second one is to minimize an assignment cost from non-visited nodes to visited ones. In spite of its natural bi-criteria formulation, this problem has always been studied in a single-objective form where either both objectives are combined or one objective is treated as a constraint.

Now our goal is to propose VRP libraries that are able to cope with a large number of constraints. The libraries are conceived in such a way that totally new sets of constraints can be added dynamically, similarly to what is done in constraint programming. This code will be used in future industrial collaborations, and already includes algorithms to use GIS.

For all studied problems, standard benchmarks have been extended to the multi-objective case. The benchmarks and the obtained results (optimal Pareto front, best known Pareto front) are available on the Web pages associated to the project and from the MCDM (International Society on Multiple Criteria Decision Making) Web site. This is an important issue to encourage comparison experiments in the research community.

## **4.2. Application to mobile telecommunication networks**

With the extraordinary success of mobile telecommunication systems, service providers have been affording huge investments for network design and infrastructure. Mobile network design is of outmost importance, and is thus a major issue in mobile telecommunication systems. In fact, with the continuous and rapid growth of communication traffic, large scale planning becomes more and more difficult. Hence, automatic, interactive and self-adaptive optimization algorithms and tools would be very useful and helpful. Advances in this area will certainly lead to important improvements in terms of quality of service, network management and cost deployment.

In the past, the DOLPHIN team has initiated solid industrial collaborations within the domain of mobile networks. In fact, the problem of network design and frequency assignment was studied in collaboration with France Telecom. In particular, a new formulation/resolution of the problem as a multi-objective constrained combinatorial optimization problem was considered. In collaboration with Mobinets, the DOLPHIN team has also addressed the problem of access network design. The problem consists in minimizing the cost of the access network and maximizing its availability.

More recently, the DOLPHIN team has been interested in new optimization models and algorithms to address new difficult problems raised by new emerging technologies in wireless networks. In fact, wireless communications are evolving from inflexible and monolithic systems to a composite radio environment made of cognitive radio devices and networks of different technologies. Within this context, the challenge is to design new optimization techniques which are not only resource, power, scale, and applications aware, but which are self-adaptive and fully distributed in order to allow the dynamic optimization of radio-devices behaviors depending on the environment constraints e.g., spectrum availability, network traffic, user demand, etc. To achieve this goal, distributed and nature-inspired algorithms, such as ant-colony and bees, will be investigated in order to dynamically and distributively optimize predefined criterion such as throughput, fairness, quality of service to cite a few. It is expected that the techniques developed in this work will lead to the design of new models and algorithms for opportunistic/dynamic spectrum access and cross layer network optimization which are at the core of future generation wireless networks.

### 4.3. Application to Bioinformatics

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 to analyze 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 datamining 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 number of data increases rapidly. Within the new collaboration, started in 2010, with Genes Diffusion, specializes in genetics and animal reproduction for bovine, swine, equine and rabbit species, we will study combinations of Single Nucleotide Polymorphisms (SNP) that can explain some phenotypic characteristics.

#### 4.3.2. Docking and conformational sampling

In molecular modelling, conformational sampling and docking procedures provide help for understanding the interaction mechanisms between (macro)molecules involved in physiological processes. The processes to be simulated are of a combinatorial complexity (molecule size, number of degrees of freedom) which represents an important challenge for the currently available computing power. Such a challenge can be expressed by three major objectives: (1) the proposition of mathematical models of maximum simplicity that nevertheless provide a relevant description of molecular behavior, (2) the development of powerful distributed optimization algorithms (evolutionary algorithms, local search methods, hybrid algorithms) for sampling the molecular energy surface for stable, populated conformations, and (3) the deployment of those intrinsic distributed algorithms on computational Grids.

Within the framework of ANR DOCK and Decrypton projects, the focus was to propose multi-objective formulations of the conformational and docking problems. The goal was to take into account different criteria characteristics of the complex docking process. Furthermore, in order to deal with the multimodal nature of the problems it is important to define new hybrid mechanisms allowing us to provide algorithms with both diversification and intensification properties. Finally, to deal with the exponential combinatoric of these problems when large proteins are concerned parallel and grid computing is highly required. Using grid computing is not straightforward, so a "gridification" process is necessary. Such process allows us to adapt the proposed algorithms to the characteristics of the grid. The gridification process must be exploited by the user in a transparent way. Therefore, coupling ParadisEO-PEO with a generic grid middleware such as Globus is important to provide robust and efficient algorithms to be exploited transparently.

New contacts with the Servier company show that these questions are really challenging ones for the design of new drug molecules.

### **4.3.3. Optimization for health care**

The new collaboration (PhD thesis started in october 2010) with Alicante company, major actor in the hospital decision making, will deal 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 combinatorial optimization will be implemented, requiring extensive work to model the problem, to define criteria optimization and to design specific optimization methods.

**FUN Team (section vide)**

## 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 [33]. 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.

## 4.4. Interactive Art

The heart of Mint project is about interaction gesture, and aims at making relation between application and user more intimate through the production of tools and methods for application to use more information from user gesture. There seems to be, at first sight, very strong difference of fields, tools, vocabulary, between Science and Art. Up to basic intellectual schemes are classically thought to be different. Yet, a closer look needs to be taken on things. Through time, Art is more and more involved in relation between people and content. For example, *relational art*<sup>1</sup> is centered on inter-human relations and social context. Because of this similar analysis about relation between person and content, research on interactive systems probably has a lot to develop relations with Art, this is also true for research on gestural interaction.

---

<sup>1</sup>[http://en.wikipedia.org/wiki/Relational\\_Art](http://en.wikipedia.org/wiki/Relational_Art)

## MODAL Project-Team

# 4. Application Domains

## 4.1. Application Domains

Potential application areas of statistical modelling 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 [8.1](#) .

## **MOSTRARE Project-Team**

# **4. Application Domains**

## **4.1. Context**

XML transformations are basic to data integration: HTML to XML transformations are useful for information extraction from the Web; XML to XML transformations are useful for data exchange between Web services or between peers or between databases. Doan and Halevy [40] survey novel integration tasks that appear with the Semantic Web and the usage of ontologies. Therefore, the semi-automatic generation of XML transformations is a challenge in the database community and in the semantic Web community.

Also, XML transformations are useful for document processing. For instance, there is need of designing transformations from documents organized w.r.t visual format (HTML, DOC, PDF) into documents organized w.r.t. semantic format (XML according to a DTD or a schema). The semi-automatic design of such transformations is obviously a very challenging objective.

Furthermore, quite some activities of Mostrare concern efficient evaluation of XPath queries on XML documents and XML streams. XPath is fundamental to all XML standards, in particular to XQuery, XSLT, and XProc.



## NON-A Project-Team

### 4. Application Domains

#### 4.1. Application domains

Unlike the traditional methods, the estimators we defined are "non-asymptotic": solutions are provided by an explicit formulae. They result in relatively simple and fast algorithms. In this sense, rather than being a project linked to a specific domain of application, we can say that the project Non-A is method-driven. However, one must not forget that applicability remains a guideline in all our researches. As it has been noted, estimation is a huge area, which touches a variety of possible application fields, which our new methods address. Figure 3 illustrates the connections between our techniques and the possible applications.

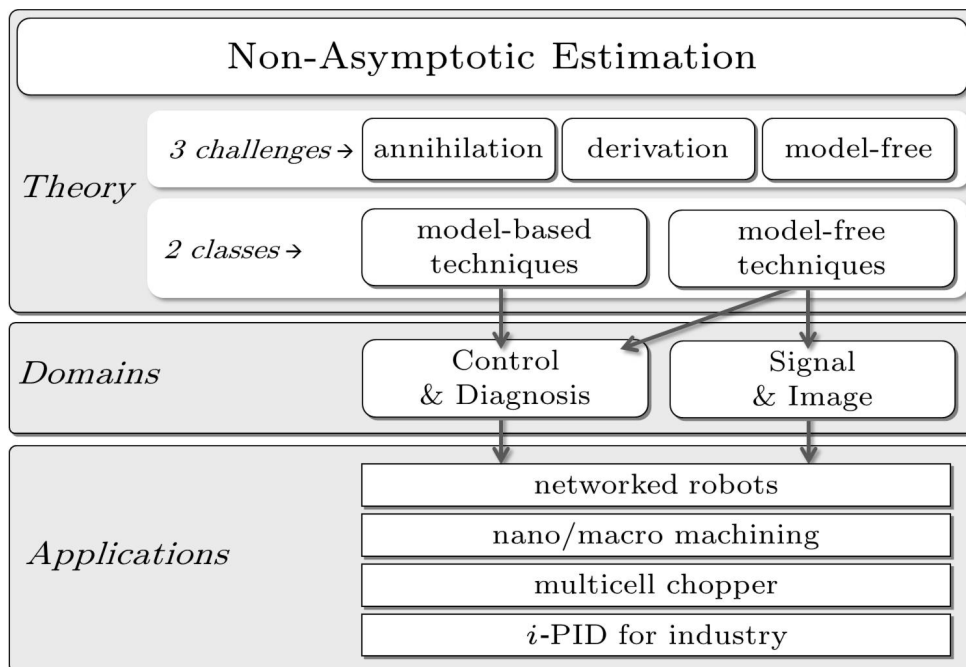


Figure 3. Non-A is a method-driven project, centered around non-asymptotic estimation techniques (i.e., providing estimates in finite-time), and connected to applications.

During these few years, our techniques have already generated 3 patents [98], [100], [99]. It shows their efficiency in various industrial domains, including (see the previous reports):

- Vehicle control (engine throttle [112], lateral and longitudinal velocities [96], stop-and-go [131], tire/road contact condition [134]) with PSA, APEDGE, Mines-ParisTech, Inria IMARA, Universidad Carlos III (Madrid), Université Paris Sud, spark ignition engine control with GM [28];
- Hydroelectric power plants [111], [110] with EDF-CIH (patent pending FR0858532);
- Shape memory actuators [107] with Université de Bretagne Occidentale and ANR MAFESMA;
- Magnetic actuators with Univ. des Saarlandes;

- Power Electronics [121] with Univ. du Québec à Trois-Rivières;
- Aircraft identification [130] with ONERA DCSD;
- Secured communications (chaos-based cryptography [128], [136], [135], CPM demodulation [122]) with CINVESTAV Mexico, Math.Dept. Tlemcen Univ. Algeria and PRISME ENSI-Bourges.
- Image and video processing (denoising [102], edge detection [113]) with Inria QGAR, compression [103], compressive sensing [132], [133] with CINVESTAV Mexico and Whuan Univ., China;
- Financial engineering [97] with MEREOR Investment Management and Advisory SAS;
- Neural signal processing [35], [83];
- Stabilization of a ball & beam system through saturated control [91], [81]
- Control of a new class of power converters [48], [49];
- Power management of a hydrogen fuel cell system combined with super capacitors [69];
- Sensorless parameter identification for permanent magnet stepper motors [60], [58], [59];
- A flexible robot joint control [78];
- Localization of mobile robots [82];
- Teleoperation [46], [44], [45];
- Oscillatory failure case (OFC) detection for aircraft [38].

**RMOD Project-Team (section vide)**

## SEQUEL Project-Team

# 4. Application Domains

## 4.1. Introduction

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,
- medical applications,
- 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 [69]. 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 softwares: in the 1990's, RL has been the basis of a very successful Backgammon program, TD-Gammon [75] 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 [69]
- 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 [73], [72] 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 Demspter 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. Medical Applications

One of the initial motivations of the multi-arm bandit theory stems from clinical trials when one researches the effects of different treatments while maximizing the improvement of the patients' health states.

Medical health-care and in particular patient-management is up today one of the most important applications of the sequential decision making. This is because the treatment of the more complex health problems is typically sequential: A physician repeatedly observes the current state of the patient and makes the decision in order to improve the health condition as measured for example by *qualys* (quality adjusted life years).

Moreover, machine learning methods may be used for at least two means in neuroscience:

1. as in any other (experimental) scientific domain, the machine learning methods relying heavily on statistics, they may be used to analyse experimental data,
2. dealing with induction learning, that is the ability to generalize from facts which is an ability that is considered to be one of the basic components of “intelligence”, machine learning may be considered as a model of learning in living beings. In particular, the temporal difference methods for reinforcement learning has strong ties with various concepts of psychology (Thorndike's law of effect, and the Rescorla-Wagner law to name the two most well-known).

## 4.5. 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 sometime 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.

## 4.6. 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.

## SHACRA Project-Team

# 4. Application Domains

## 4.1. Clinical Applications

Some of the scientific challenges described previously can be seen in a general context (such as solving constraints between different types of objects, parallel computing for interactive simulations, etc.) but often it is necessary to define a clinical context for the problem. This is required in particular for defining the appropriate assumptions in various stages of the biophysical modeling. It is also necessary to validate the results. This clinical context is a combination of two elements: the procedure we attempt to simulate and the objective of the simulation: training, planning or per-operative guidance. Below are a series of applications we plan to develop. The choice of these applications is not random: the clinical procedures we target are all technically challenging, they highlight various parts of our research, and often they represent an ideal testbed for transitioning from training to planning to guidance. It is important also to note that developing these applications raises many challenges and as such this step should be seen as an integral part of our research. It is also through the development of these applications that we can communicate with physicians, and validate our results. SOFA will be used as a backbone for the integration of our research into clinical applications.

### 4.1.1. Interventional radiology

Over the past twenty years, interventional methods such as angioplasty, stenting, and catheter-based drug delivery have substantially improved the outcomes for patients with vascular disease. Pathologies that used to require a surgical procedure can now be treated in a much less invasive way. As a consequence, interventional radiology procedures represent an increasing part of the interventions currently performed, with more than 6 million patients treated every year in Europe and about 5 millions the United States. However, these techniques require an intricate combination of tactile and visual feedback, and extensive training periods to attain competency. To reinforce the need to reach and maintain proficiency, the FDA recently required that US physicians go through simulation-based training before using newly developed carotid stents. Besides simulation for training, interventional radiology is a perfect target to illustrate the potential of planning and rehearsal of procedures. As an initial step in this direction, Alcove and Magrit were partners in an ARC project (Simple) to develop a planning tool for the treatment of aneurysms using coils. This collaboration still goes on after the end of the ARC, and led to a series of papers in key conferences [5] [28], [34], [21].

#### 4.1.1.1. Interventional neuro-radiology

We will continue the development of our simulation and planning system for interventional radiology, with two principal clinical partners: Massachusetts General Hospital in Boston and University Hospital in Nancy. We have completed the integration in SOFA of improved versions of algorithms for describing the behavior of catheters, guide-wires, coils, as well as the interactive simulation of fluoroscopic images, the modeling of complex contacts. Our future efforts will focus on the development of an advanced planning system for interventional radiology, in particular for coil embolization. This will require the integration of new methods of reconstruction of vascular anatomy from medical images (in collaboration with the MAGRIT team). We will also add our recent results on blood flow simulation in aneurysms.

#### 4.1.1.2. Interventional cardiology using radio-frequency ablation

Cardiac arrhythmias (or dysrhythmias) are problems that affect the electrical system of the heart muscle, producing abnormal heart rhythms, and causing the heart to pump less effectively. About 5% of people over 40 years old are affected by this pathology, with a rather high morbidity rate. Radio-frequency ablation is a non-surgical procedure that has been used for about 15 years to treat tachyarrhythmias, i.e. rapid, uncoordinated heartbeats. The procedure is performed by guiding a catheter with an electrode at its tip to the area of heart muscle where there is an accessory pathway. The catheter is guided under fluoroscopic imaging. When the catheter is positioned at the site where cells give off the electrical signals that stimulate the abnormal heart

rhythm, a low radio-frequency energy is transmitted to the pathway. This destroys heart muscle cells within a very small area near the tip of the catheter and stops the area from conducting the extra impulses that caused the arrhythmia. In this context, a simulation system would be able to provide added value in two main areas: 1) to train physicians in the early stages of their apprenticeship and 2) to provide quantitative information during the planning phase of a complex procedure, using patient-specific data. Most aspects of this simulation will rely on components developed during our research program but we will also extend our collaboration with the ASCLEPIOS team and the CardioSense3D project on the modeling of the heart the Cadiosense3D project. This involves an important integration task, and it will also validate the reusability aspects of the code developed within SOFA.

#### **4.1.2. Minimally-invasive surgery**

##### *4.1.2.1. Laparoscopic hepatic resection*

The liver is one of the major organs in the human body. It is in charge of more than 100 vital functions. Because of its many functions, its pathologies are also varied, numerous and unfortunately often lethal. This is for instance the case of hepatitides which today affect about 300,000 people in France for hepatitis B and 600,000 people for hepatitis C. The most advanced state of evolution of these pathologies is generally cirrhosis followed by cancer, which represents the third cause of cancer related death. In 2005, 14,267 liver cancer cases and 20,497 cirrhosis cases have been diagnosed in France. The surgical solution remains the option offering the best success rate for these pathologies. More than 7,000 surgical interventions have been carried out on the liver in 2005 and partial resection of the liver remains the most common approach. In this context, the ability to train surgeons, and to be able to plan complex procedures using computer-based simulations, would be a formidable help to the current apprenticeship model: "See One, Do One, Teach One". Right now, only a few commercial systems are available to the medical community, and they are limited to basic skills training. Developing a realistic simulation system that could be used to plan and rehearse procedures would be a very important step in the introduction of new training paradigms in medicine. This is the main objective of the PASSPORT european project in which we are actively contributing at two levels. First, our research results on biomechanical modeling of solid organs and on coupling will be used to propose a realistic model of the deformation of the liver and its vascular network. Second, SOFA has been chosen in this project as the software for integrating all results from the different partners. Both aspects will help validate our models, test SOFA and obtain feedback from the clinicians.

##### *4.1.2.2. Ophthalmology and cataract surgery*

A cataract is an opacity in the natural lens of the eye. It represents an important cause of visual impairment and, if not treated, can lead to blindness. It is actually the leading cause of blindness worldwide, and its development is related to aging, sunlight exposure, smoking, poor nutrition, eye trauma, and certain medications. The best treatment for this pathology remains surgery. Cataract surgery has made important advances over the past twenty years, and in 2005, more than 5 million people in the United States and in Europe underwent cataract surgery. Most cataract surgeries are performed using microscopic size incisions, advanced ultrasonic equipment to fragment cataracts into tiny fragments, and foldable intraocular lenses to minimize the size of the incision. All these advances benefit the patient, but increase training requirements for eye surgeons. At the end of 2007, we started the development of a new training system for cataract surgery. The main objectives of this simulation are to reproduce with great accuracy the three main steps of cataract surgery: 1) capsulorhexis 2) phacoemulsification and 3) implantation of an intraocular lens. We have already started the development of this simulation. The main research effort went in the choice of appropriate deformable models for the lens and lens capsule. An important effort also went into the development of topological changes corresponding to the capsulorhexis and phacoemulsification [20]. The modeling of the intraocular implant and its deployment in the capsule has been published to the major conference in medical simulation [26].

##### *4.1.2.3. Neurosurgery and deep brain stimulation*

Deep brain stimulation (DBS) is a neurosurgical treatment which stimulates the brain with low electrical signals. The signals reorganize the brain's electrical impulses (similarly as what was presented above for radio-frequency ablation for cardiac problems). This results in major improvements in several pathologies



such as Parkinson disease. The principle of the procedure is the following: a thin, insulated wire lead with several electrodes at the tip is surgically implanted into the affected area of the brain. A wire runs under the skin to a battery-operated pulse generator implanted near the collarbone. The generator is programmed to send continuous electrical pulses to the brain. To implant the electrodes, a neurosurgeon uses a stereotactic head frame and magnetic resonance or computed tomography imaging to map the brain and pinpoint the problem area. The main difficulty in this procedure comes from the deformation of the brain (small brain shift when the skull is opened, and local deformation of the brain due to the insertion of the electrode) and the deflection of the electrode itself during and after the procedure. This results in a difference between the planned target and the location of the end effector of the electrode. Our main objective is to use our work on soft tissue deformation, vascularized structures, as well as our recent results on constraint solving between soft tissues and flexible devices [29]. This work will be done in collaboration with the VISAGES team and we will dedicate an important effort in validating our results, analyzing post-operative medical images, and interacting with surgeons. This project has a strong potential as DBS is being increasingly used yet most research groups only consider non deformable planning systems (geometrical planning). Our proposal could make a important difference in the accuracy of the planning as it takes into account the biophysics of the brain.

## SIMPAF Project-Team

## 4. Application Domains

### 4.1. Fluid mechanics

#### 4.1.1. Numerical methods for viscous flows

We are concerned with systems of PDEs describing the evolution of mixture flows. The fluid is described by its density, its velocity and its pressure. These quantities obey mass and momentum conservation. On the one hand, when we deal with the 2D variable density incompressible Navier-Stokes equations, we aim at studying some instabilities phenomena such as the Raileigh-Taylor instability.

Furthermore, diffuse interface models have gained renewed interest in the last few years in fluid mechanics applications. From a physical viewpoint, they allow to describe phase transition phenomena. In this case, a specific stress tensor, introduced originally by Korteweg, must be added to the momentum equation.

If we use in addition the Fick law to relate the divergence of the velocity to derivatives of the density, one obtains the so-called KS model. In this case, the density of the mixture is naturally highly heterogeneous, and may model powder-snow avalanches or specific pollutants. Similar models also appear in combustion theory.

#### 4.1.2. Control

Flow control strategies using passive or active devices are crucial tools in order to save energy in transports (especially for cars, trucks or planes), or to avoid the fatigue of some materials arising in a vast amount of applications. Nowadays, shape optimization needs to be completed by other original means, such as porous media located on the profiles, as well as vortex generator jets in order to drive active control.

### 4.2. Rubber elasticity

At the continuum level, rubber is modelled by an energy  $E$  defined as the integral over a domain  $D$  of  $\mathbb{R}^d$  of some energy density  $W$  depending only locally on the gradient of the deformation  $u$ :  $E(u) = \int_D W(\nabla u(x)) dx$ . At the microscopic level (say 100nm), rubber is a network of cross-linked and entangled polymer chains (each chain is made of a sequence of monomers). At this scale the physics of polymer chains is well-understood in terms of statistical mechanics: monomers thermally fluctuate according to the Boltzmann distribution [61]. The associated Hamiltonian of a network is typically given by a contribution of the polymer chains (using self-avoiding random bridges) and a contribution due to steric effects (rubber is packed and monomers are surrounded by an excluded volume). The main challenge is to understand how this statistical physics picture yields rubber elasticity, and how this can be taken advantage of in order to devise physically-based constitutive laws for rubber. A rigorous derivation can be performed using tools of stochastic homogenization. The challenge is now to devise a numerical method to pass from polymer physics to an explicit continuum constitutive law, and to apply this approach to physically relevant models.

### 4.3. Maxwell equations

The a posteriori error estimators developed for the Maxwell equations are very useful tools. They are implemented in the software "Carmel-3D" (see the softwares section). This numerical code is used in order to study some original applications, like electrical machines or specific actuators. It is also devoted to nondestructive control by the use of Foucault currents, to the simulation of devices using magnetic fluids or of induced currents in human bodies.

### 4.4. Corrosion models

The application domains related to corrosion models are the following.

#### **4.4.1. Corrosion modelling of iron based alloy in nuclear waste repository**

The concept for long term storage of high-level radioactive waste in France is based on an underground repository. The waste shall be confined in a glass matrix and then placed into cylindrical steel canisters. These containers shall be placed into micro-tunnels in the highly impermeable Callovo-Oxfordian claystone layer at a depth of several hundred meters. At the request of the French nuclear waste management agency ANDRA, investigations are conducted to optimize and finalize this repository concept with the aim to ensure its long-term safety and its reversibility. The long-term safety assessment of the geological repository has to take into account the degradation of the carbon steel used for the waste overpacks and the cell disposal liners, which are in contact with the claystone formation. This degradation is mainly caused by generalized corrosion processes which form a passive layer on the metal surface consisting of a dense oxide inner layer and a porous hydroxide outer layer in contact with the groundwater in the pore space of the claystones. The processes take place under anaerobic conditions, since the groundwater is anoxic.

As a tool to investigate the corrosion processes at the surface of the carbon steel canisters, the Diffusion Poisson Coupled Model (DPCM) for corrosion has been developed by Bataillon *et al.* [47]. The numerical approximation of this corrosion model by accurate and efficient methods is challenging.

#### **4.4.2. Corrosion modelling of Ni-base alloys in Pressurized Water Reactor primary water**

The understanding of the oxidation behavior of Ni-base alloys in PWR primary water is of major importance due to the cations released due to corrosion of the steam generators which is a source of the radioactivity of the primary circuit. Moreover, the oxidation process is the reason of the initiation of intergranular stress corrosion cracking in some alloys. A numerical model, called EKINOX (Estimation KINetics OXYdation), has been developed at CEA [48] in order to simulate the oxide scale growth. This model should be able to calculate the evolutions of concentration profiles of the species and of their point defects in the oxide and in the substrate. Numerical experiments have shown the limits of this existing numerical model, especially the need of very small time steps for the computations ; a macroscopic model should be developed and numerical methods proposed for its simulation.

### **4.5. Plasmas**

In the last decade, satellite-positioning devices have become one of the most powerful navigation assistance systems. Recently, Europe launched the first satellites of Galileo, a competitor to the constellation of satellites Navstar GPS American system. In order to guarantee the integrity service level, it is fundamental to take into account the various physical phenomena that can affect the mission and to identify all the potential sources of system unavailability. One of the main sources of data unavailability that has been identified is the phenomenon of ionospheric scintillations. Indeed scintillation causes radio frequency signal amplitude to fade and phase variations as satellite signals pass through the ionosphere. Such effects may induce loss of lock or cycle slips on ranging signals broadcast by Galileo satellites making them totally useless for accurate integrity information determination.

### **4.6. Optical fibers**

When designing optical fibers, the question of the envelope dynamics of laser impulses is crucial because its answer decides how much information the fiber will actually be able to carry. Recent progresses in the optical fiber design now allow to choose specific physical parameters very precisely, and the main limitation for being able to design new and more efficient optical fibers is the understanding of the effect of specific conjunctions of parameter values. In other terms, a main issue is the understanding of the influence of several parameters on specific solutions of nonlinear dispersive equations of Schrödinger type.