



IN PARTNERSHIP WITH:
**Institut national des sciences
appliquées de Rennes**
Université Rennes 1

Activity Report 2013

Project-Team DREAM

Diagnosing, Recommending Actions and Modelling

IN COLLABORATION WITH: Institut de recherche en informatique et systèmes aléatoires (IRISA)

RESEARCH CENTER
Rennes - Bretagne-Atlantique

THEME
**Data and Knowledge Representation
and Processing**

Table of contents

1. Members	1
2. Overall Objectives	1
2.1. Introduction	1
2.2. Application domains	2
2.3. Highlights of the Year	3
3. Research Program	3
3.1. Computer assisted monitoring and diagnosis of physical systems	3
3.2. Machine learning and data mining	4
3.2.1. Relational learning	4
3.2.2. Data mining	5
3.2.3. Mining data streams	5
4. Application Domains	5
4.1. Introduction	5
4.2. Software components monitoring	6
4.3. Environmental decision making	6
5. Software and Platforms	7
5.1. Introduction	7
5.2. QTempIntMiner: quantitative temporal sequence mining	7
5.3. Sacadeau: qualitative modeling and decision-aid to preserve the water quality from pollutants as herbicides	7
5.4. Ecomata	8
5.5. Paturmata	8
5.6. ManageYourself	8
5.7. GeoImageRMP: a RapidMiner extension to georeferenced data	9
5.8. A plugin for visualizing and editing spatial graphs in QGis	9
5.9. Odiseptale: a software for implementing and evaluating sanitary event detectors in cattle	10
6. New Results	10
6.1. Diagnosis of large scale discrete event systems	10
6.1.1. Distributed monitoring with chronicles - Interleaving diagnosis and repair - Making web services more adaptive	10
6.1.2. Scenario patterns for exploring qualitative ecosystems	11
6.1.3. Controller synthesis for dealing with “How to” queries	11
6.2. Machine learning for model acquisition	11
6.2.1. Representing and mining time series	11
6.2.2. Incremental sequential mining	12
6.2.3. Multiscale segmentation of satellite image time series	12
6.2.4. Analysis of landscape based on spatial patterns	12
6.2.5. Subdimensional clustering for fast similarity search over time series data. Application to Information retrieval tasks	13
6.2.6. Knowledge Extraction from Heterogeneous Data	13
6.3. Decision aiding with models and simulation data	14
6.3.1. A datawarehouse for simulation data	14
6.3.2. Efficient computation of skyline queries in an interactive context	14
6.3.3. Hierarchical skylines	14
6.3.4. Modeling influence propagation by Bayesian causal maps	14
6.3.5. Recommending actions from classification rules	15
6.4. Diagnostic, causal reasoning and argumentation	15
7. Bilateral Contracts and Grants with Industry	16
8. Partnerships and Cooperations	16

8.1. Regional Initiatives	16
8.2. National Initiatives	16
8.2.1. ACASSYA: Supporting the agro ecological evolution of breeding systems in coastal watersheds	16
8.2.2. Asterix : spatio-temporal analysis of remote sensing images	17
8.3. International Initiatives	17
8.3.1.1. Declared Inria International Partners	17
8.3.1.2. Informal International Partners	17
9. Dissemination	17
9.1. Scientific Animation	17
9.1.1. Journal editorial board	17
9.1.2. Conference program committees and organizations	17
9.1.3. Scientific and administrative boards	18
9.1.4. Misc	18
9.2. Teaching - Supervision - Juries	18
9.2.1. Teaching	18
9.2.2. Supervision	19
9.2.3. Juries	19
10. Bibliography	19

Project-Team DREAM

Keywords: Artificial Intelligence, Diagnosis, Knowledge Acquisition, Data Mining, Decision Aid, Adaptive Systems, Environment

Creation of the Project-Team: 2004 October 04.

1. Members

Research Scientists

Yves Moinard [Inria, Researcher]

René Quiniou [Inria, Researcher]

Faculty Members

Marie-Odile Cordier [Team leader, Univ. Rennes I, Professor, HdR]

Tassadit Bouadi [Univ. Rennes I, Assistant Professor]

Véronique Masson [Univ. Rennes I, Associate Professor]

Laurence Rozé [INSA Rennes, Associate Professor]

External Collaborators

Philippe Besnard [CNRS Irit, Toulouse, HdR]

Louis Bonneau de Beaufort [Agrocampus-Ouest Rennes]

Thomas Guyet [Agrocampus-Ouest Rennes]

Christine Largouët [Agrocampus-Ouest Rennes]

Simon Malinowski [Agrocampus-Ouest Rennes, until Aug 2013]

PhD Student

Yulong Zhao [Univ. Rennes I, until Jan 2014]

Post-Doctoral Fellow

Sid Ahmed Benabderrahmane [Inria]

Visiting Scientist

Sophie Robin [Collaborator, ex-Associate Professor]

Administrative Assistant

Marie-Noëlle Georgeault [Inria]

2. Overall Objectives

2.1. Introduction

The research goals of the Dream project-team concern monitoring complex systems. The challenge is to design smart systems, both adaptable and dependable, to answer the demand for self-healing embedded systems. The considered systems meet a fixed common goal (or contract), possibly expressed by a set of QoS (Quality of Service) constraints. The Dream team investigates and develops model-based approaches. Dealing with dynamic systems, a central role is given to temporal information and the model specification uses event-based formalisms such as discrete-event systems (mainly described by automata), or sets of chronicles (a chronicle is a temporally constrained set of events). We investigate two main research questions. Firstly, we design and develop distributed architectures and efficient diagnosis/repair algorithms for highly distributed systems. Secondly, we study the automatic acquisition of models from data using symbolic machine learning and data mining methods, with a particular focus on data stream processing. Target applications are of two kinds: large component-based system monitoring applications, like telecommunication networks, and software systems like web services and environmental protection systems with development of decision support systems to help managing agricultural plots and support high water quality threatened by pollution.

In this context, the research questions we are investigating are the following.

Classical model-based diagnosis methodologies appear to be inadequate for complex systems due to the intractable size of the model and the computational complexity of the process. It is especially true when one considers on-line diagnosis or when many interacting components (or agents) make up the system. This is why we focus on decentralized approaches which relies on computing local diagnoses from local models and synchronizing them to get a global view of the current state of the system. The problems we are investigating are the following: (i) which strategy to select for synchronizing the local diagnoses in an optimal way to maintain the efficiency and the completeness of the process, (ii) which kind of communication protocols to use, (iii) how to improve the computation efficiency by using adequate symbolic representations, (iv) how to guarantee the efficiency of incremental on-line diagnosis where observations come from a continuous stream?

When designing a dependable and adaptive system, a main point is to formally characterize the intended properties of the system such as diagnosability (i.e. whether, given the system specifications, it is possible to detect and explain an error in due time), or reparability (i.e. whether it is possible to get the system back to correctness, in due time).

More recently, we enlarge our interest and consider in the same view both monitoring, deficiency detection, diagnosis and the consequent system adaptation or repair. We extended diagnosability to self-healability and investigated how to weave diagnosis and repair, to get adaptive systems maintaining a good QoS, even in unexpected, or even abnormal conditions.

It is well-recognized that model-based approaches suffer from the difficulty of model acquisition. The first issue we have studied is the automatic acquisition of models from data with symbolic learning methods and data mining methods. We list the investigated problems here. How to improve relational learning methods to cope efficiently with data coming from signals (as an electrocardiogram in the medical domain) or alarm logs (in the telecommunication domain)? How to integrate signal processing algorithms to the learning or diagnosis tasks when these latter ones rely on a qualitative description of signals? How to adapt the learning process to deal with multiple sources of information (multi-sensor learning)? How to apply learning techniques to spatiotemporal data? How to combine data mining and visualization to help experts build their models?

Concerning evolving context management and adaptive systems, an emerging issue is to detect when a model is becoming obsolete and to update it by taking advantage of the current data. This difficult and new issue has strong connections with data streams processing. This is a big challenge in the monitoring research area where the model serves as a reference for the diagnosis task.

The last point we consider is the decision part itself, mainly having abilities to propose repair policies to restore the functionalities of the system or the expected quality of service. A first direction is to interleave diagnosis and repair and to design some decision-theoretic procedure to dynamically choose the best action to undertake. Another direction concerns how to automatically build the recommending actions from simulation or recorded data.

2.2. Application domains

Our application domains have links with funds and contracts we have got thanks to long-term relations with academic and industrial partners. These application domains serve us both as providers of real challenging problems and as test-beds for our research development. One should not consider them as distinct research areas but as distinct experimentation fields, for confronting similar methodologies and techniques on various application contexts. We investigate the following application domains:

- large component-based system monitoring applications: the two main investigated domains are telecommunication networks, and software systems as those found in embedded systems or web services.
- environmental protection: more precisely, we are developing decision support systems to help managing agricultural water catchments and ecosystems under fishing pressure.

2.3. Highlights of the Year

Prof. Torsten Schaub (Potsdam University, Germany) has been awarded an Inria international senior grant from 2013 to 2017 for a research project with EPI-DREAM. This research will be concerned with using ASP for data stream post-mining.

3. Research Program

3.1. Computer assisted monitoring and diagnosis of physical systems

keywords: monitoring, diagnosis, deep model, fault model, simulation, chronicle acquisition

Our work on monitoring and diagnosis relies on model-based approaches developed by the Artificial Intelligence community since the seminal studies by R. Reiter and J. de Kleer [78], [89]. Two main approaches have been proposed then: (i) the consistency-based approach, relying on a model of the expected correct behavior ; (ii) the abductive approach which relies on a model of the failures that might affect the system, and which identifies the failures or the faulty behavior explaining the anomalous observations. See the references [29], [31] for a detailed exposition of these investigations.

Since 1990, the researchers in the field have studied dynamic system monitoring and diagnosis, in a similar way as researchers in control theory do. What characterizes the AI approach is the use of qualitative models instead of quantitative ones and the importance given to the search for the actual source/causes of the faulty behavior. Model-based diagnosis approaches rely on qualitative simulation or on causal graphs in order to look for the causes of the observed deviations. The links between the two communities have been enforced, in particular for what concerns the work about discrete events systems and hybrid systems. Used formalisms are often similar (automata, Petri nets ,...) [37], [35].

Our team focuses on monitoring and on-line diagnosis of discrete events systems and in particular on monitoring by alarm management.

Two different methods have been studied by our team in the last years:

- In the first method, the automaton used as a model is transformed off-line into an automaton adapted to diagnosis. This automaton is called a *diagnoser*. This method has first been proposed by M. Sampath and colleagues [80]. The main drawback of this approach is its centralized nature that requires to explicitly build the global model of the system, which is most of the time unrealistic. It is why we proposed a decentralized approach in [75].
- In the second method, the idea is to associate each failure that we want to detect with a *chronicle* (or a scenario), i.e. a set of observable events interlinked by time constraints. The chronicle recognition approach consists in monitoring and diagnosing dynamic systems by recognizing those chronicles on-line [53], [77], [51].

One of our research focus is to extend the chronicle recognition methods to a distributed context. Local chronicle bases and local recognizers are used to detect and diagnose each component. However, it is important to take into account the interaction model (messages exchanged by the components). Computing a global diagnosis requires then to check the synchronisation constraints between local diagnoses.

Another issue is the chronicle base acquisition. An expert is often needed to create the chronicle base, and that makes the creation and the maintenance of the base very expensive. That is why we are working on an automatic method to acquire the base.

Developing diagnosis methodologies is not enough, especially when on-line monitoring is required. Two related concerns must be tackled, and are the topics of current research in the team:

- The ultimate goal is usually not merely to diagnose, but to put the system back in some acceptable state after the occurrence of a fault. One of our aim is to develop self-healable systems able to self-diagnose and -repair.
- When designing a system and equipping it with diagnosis capabilities, it may be crucial to be able to check off-line that the system will behave correctly, i.e., that the system is actually 'diagnosable'. A lot of techniques have been developed in the past (see Lafortune and colleagues [79]), essentially in automata models. We extended them to cope with temporal patterns. A recent focus has been to study the self-healability of systems (ability to self-diagnose and self-repair).

3.2. Machine learning and data mining

keywords: machine learning, Inductive Logic Programming (ILP), temporal data mining, temporal abstraction, data-streams

The machine learning and data mining techniques investigated in the group aim at acquiring and improving models automatically. They belong to the field of machine (artificial) learning [48]. In this domain, the goal is the induction or the discovery of hidden objects characterizations from their descriptions by a set of features or attributes. For several years we investigated Inductive Logic Programming (ILP) but now we are also working on data-mining techniques.

We are especially interested in structural learning which aims at making explicit dependencies among data where such links are not known. The relational (temporal or spatial) dimension is of particular importance in applications we are dealing with, such as process monitoring in health-care, environment or telecommunications. Being strongly related to the dynamics of the observed processes, attributes related to temporal or spatial information must be treated in a special way. Additionally, we consider that the legibility of the learned results is of crucial importance as domain experts must be able to evaluate and assess these results.

The discovery of spatial patterns or temporal relations in sequences of events involve two main steps: the choice of a learning space and the choice of a learning technique.

We are mainly interested in symbolic supervised and unsupervised learning methods. Furthermore, we are investigating methods that can cope with temporal or spatial relationships in data. In the sequel, we will give some details about relational learning, relational data-mining and data streams mining.

3.2.1. Relational learning

Relational learning, also called inductive logic programming (ILP), lies at the intersection of machine learning, logic programming and automated deduction. Relational learning aims at inducing classification or prediction rules from examples and from domain knowledge. As relational learning relies on first order logic, it provides a very expressive and powerful language for representing learning hypotheses especially those learnt from temporal data. Furthermore, domain knowledge represented in the same language can also be used. This is a very interesting feature which enables taking into account already available knowledge and avoids starting learning from scratch.

Concerning temporal data, our work is more concerned with applying relational learning rather than developing or improving the techniques. Nevertheless, as noticed by Page and Srinivasan [74], the target application domains (such as signal processing in health-care) can benefit from adapting relational learning scheme to the particular features of the application data. Therefore, relational learning makes use of constraint programming to infer numerical values efficiently [81]. Extensions, such as QSIM [62], have also been used for learning a model of the behavior of a dynamic system [54]. Precisely, we investigate how to associate temporal abstraction methods to learning and to chronicle recognition. We are also interested in constraint clause induction, particularly for managing temporal aspects. In this setting, the representation of temporal phenomena uses specific variables managed by a constraint system [76] in order to deal efficiently with the associated computations (such as the covering tests).

For environmental data, we have investigated tree structures where a set of attributes describe nodes. Our goal is to find patterns expressed as sub-trees [47] with attribute selectors associated to nodes.

3.2.2. Data mining

Data mining is an unsupervised learning method which aims at discovering interesting knowledge from data. Association rule extraction is one of the most popular approach and has deserved a lot of interest in the last 10 years. For instance, many enhancements have been proposed to the well-known Apriori algorithm [33]. It is based on a level-wise generation of candidate patterns and on efficient candidate pruning having a sufficient relevance, usually related to the frequency of the candidate pattern in the data-set (i.e., the support): the most frequent patterns should be the most interesting. Later, Agrawal and Srikant proposed a framework for "mining sequential patterns" [34], which extends Apriori by coping with the order of elements in patterns.

In [69], Mannila and Toivonen extended the work of Agrawal et al. by introducing an algorithm for mining patterns involving temporal episodes with a distinction between parallel and sequential event patterns. Later, in [52], Dousson and Vu Duong introduced an algorithm for mining chronicles. Chronicles are sets of events associated with temporal constraints on their occurrences. They generalize the temporal patterns of Mannila and Toivonen. The candidate generation is an Apriori-like algorithm. The chronicle recognizer CRS [50] is used to compute the support of patterns. Then, the temporal constraints are computed as an interval whose bounds are the minimal and the maximal temporal extent of the delay separating the occurrences of two given events in the data-set. Chronicles are very interesting because they can model a system behavior with sufficient precision to compute fine diagnoses. Their extraction from a data-set is reasonably efficient. They can be efficiently recognized on an input data stream.

Relational data-mining [30] can be seen as generalizing these works to first order patterns. In this field, the work of Dehaspe for extracting first-order association rules have strong links with chronicles. Another interesting research concerns inductive databases which aim at giving a theoretical and logical framework to data-mining [63], [49]. In this view, the mining process means to query a database containing raw data as well as patterns that are implicitly coded in the data. The answer to a query is, either the solution patterns that are already present in the database, or computed by a mining algorithm, e.g., Apriori. The original work concerns sequential patterns only [67]. We have investigated an extension of inductive databases where patterns are very close to chronicles [85].

3.2.3. Mining data streams

During the last years, a new challenge has appeared in the data mining community: mining from data streams [32]. Data coming for example from monitoring systems observing patients or from telecommunication systems arrive in such huge volumes that they cannot be stored in totality for further processing: the key feature is that "you get only one look at the data" [56]. Many investigations have been made to adapt existing mining algorithms to this particular context or to propose new solutions: for example, methods for building synopses of past data in the form of summaries have been proposed, as well as representation models taking advantage of the most recent data. Sequential pattern stream mining is still an issue [70]. At present, research topics such as, sampling, summarizing, clustering and mining data streams are actively investigated.

A major issue in data streams is to take into account the dynamics of process generating data, i.e., the underlying model is evolving and, so, the extracted patterns have to be adapted constantly. This feature, known as *concept drift* [86], [64], occurs within an evolving system when the state of some hidden system variables changes. This is the source of important challenges for data stream mining [55] because it is impossible to store all the data for off-line processing or learning. Thus, changes must be detected on-line and the current mined models must be updated on line as well.

4. Application Domains

4.1. Introduction

The DREAM research applications have been oriented towards surveillance of large networks as telecommunication networks and more recently of web services. During the past few years, we have focussed increasingly on agricultural and environmental applications by means of research collaborations with INRA and Agrocampus Ouest.

4.2. Software components monitoring

software components, web services, distributed diagnosis

Web-services, i.e., services that are provided, controlled and managed through Internet, cover nowadays more and more application areas, from travel booking to goods supplying in supermarkets or the management of an e-learning platform. Such applications need to process requests from users and other services on line, and respond accurately in real time. Anyway, errors may occur, which need to be addressed in order to still be able to provide the correct response with a satisfactory quality of service (QoS): on-line monitoring, especially diagnosis and repair capabilities, become then a crucial concern.

We have been working on this problem within the WS-DIAMOND project [84], a large European funded project involving eight partners in Italy, France, Austria and Netherlands <http://wsdiamond.di.unito.it/>. Our own work consisted in two distinct contributions.

The first issue has been to extend the decentralized component-oriented approach, initially developed for monitoring telecommunication networks [4] to this new domain. To this end we have proposed the concept of distributed chronicles, with synchronization events, and the design of an architecture consisting of distributed CRSs (Chronicle Recognition Systems) communicating their local diagnoses to a broker agent which is in charge of merging them to compute a global diagnosis.

Our current work aims at coupling diagnosing and repair, in order to implement *adaptive web services*. We started this study by proposing an architecture inspired from the one developed during the WS-DIAMOND project and dedicated to the adaptive process of a request event when faults occur and propagate through the orchestration.

4.3. Environmental decision making

environment, decision methods

The need of decision support systems in the environmental domain is now well-recognized. It is especially true in the domain of water quality. For instance the program, named “Bretagne Eau Pure”. was launched a few years ago in order to help regional managers to protect this important resource in Brittany. The challenge is to preserve the water quality from pollutants as nitrates and herbicides, when these pollutants are massively used by farmers to weed their agricultural plots and improve the quality and increase the quantity of their crops. The difficulty is then to find solutions which satisfy contradictory interests and to get a better knowledge on pollutant transfer.

In this context, we are cooperating with INRA (Institut National de Recherche Agronomique) and developing decision support systems to help regional managers in preserving the river water quality. The approach we advocate is to rely on a qualitative modeling, in order to model biophysical processes in an explicative and understandable way. The SACADEAU model associates a qualitative biophysical model, able to simulate the biophysical process, and a management model, able to simulate the farmer decisions. One of our main contribution is the use of qualitative spatial modeling, based on runoff trees, to simulate the pollutant transfer through agricultural catchments.

The second issue is the use of learning/data mining techniques to discover, from model simulation results, the discriminant variables and automatically acquire rules relating these variables. One of the main challenges is that we are faced with spatiotemporal data. The learned rules are then analyzed in order to recommend actions to improve a current situation.

This work has been done in the framework of the APPEAU project, funded by ANR and of the ACASSYA project, funded by ANR, having started at the beginning of 2009 and ended at the end of 2012. We were also involved in the PSDR GO CLIMASTER project, that started in september 2008 and end in 2011. CLIMASTER stands for “Changement climatique, systèmes agricoles, ressources naturelles et développement territorial” and is dedicated to the impact of climate changes on the agronomical behaviors in west of France (“Grand Ouest”). PSDR GO stands for “Programme Pour et Sur le Développement Régional Grand Ouest”.

Our main partners are the SAS INRA research group, located in Rennes and the BIA INRA and AGIR INRA research groups in Toulouse.

5. Software and Platforms

5.1. Introduction

The pieces of software described in this section are prototypes implemented by members of the project. Any interested person should contact relevant members of the project.

5.2. QTempIntMiner: quantitative temporal sequence mining

QTEMPINTMINER (Quantitative Temporal Interval Miner) is a data mining (cf. 3.2.2) software that implements several algorithms presented in [59] and [3].

The software is mainly implemented in Matlab. A standalone application is now available. It uses the Mixmod toolbox [44] to compute multi-dimensional Gaussian distributions. The main features of QTEMPINTMINER are:

- a tool for generating synthetic noisy sequences of temporal events,
- an implementation of the QTEMPINTMINER, QTIAPRIORI and QTIPREFIXSPAN algorithms,
- a graphical interface that enables the user to generate or import data set and to define the parameters of the algorithm and that displays the extracted temporal patterns.
- a sequence transformer to process long sequences of temporal events. Long sequences are transformed into a database of short temporal sequences that are used as input instances for the available algorithms.

The following website gives many details about the algorithms and provides the latest stable implementation of QTEMPINTMINER: <http://www.irisa.fr/dream/QTempIntMiner/>.

5.3. Sacadeau: qualitative modeling and decision-aid to preserve the water quality from pollutants as herbicides

SACADEAU is an environmental decision software (cf. 4.3) that implements the SACADEAU transfer model. The SACADEAU simulation model couples two qualitative models, a transfer model describing the pesticide transfer through the catchment and a management model describing the farmer decisions. Giving as inputs a climate file, a topological description of a catchment, and a cadastral repartition of the plots, the SACADEAU model simulates the application of herbicides by the farmers on the maize plots, and the transfer of these pollutants through the catchment until the river. The two main simulated processes are the runoff and the leaching. The output of the model simulation is the quantity of herbicides arriving daily to the stream and its concentration at the outlets. The originality of the model is the representation of water and pesticide runoffs with tree structures where leaves and roots are respectively up-streams and down-streams of the catchment.

The software allows the user to see the relationships between these tree structures and the rules learnt from simulations (cf. 3.2.1). A more elaborated version allows to launch simulations and to learn rules on-line. This year, we have developed this new version by enabling access to two recommendation action algorithms. The user can choose different parameters (set of classification rules from which actions will be built, parameters concerning action feasibility, etc) before asking for action recommending process, and then easily visualize the characteristics of situations to improve (polluted ones) compared with the different recommended actions. The software is mainly in Java.

The following website is devoted to the presentation of the SACADEAU: <http://www.irisa.fr/dream/SACADEAU/>. See also [10] for a presentation.

5.4. Ecomata

EcoMata is a tool-box for qualitative modeling and exploring ecosystems and for aiding to design environmental guidelines. We have proposed a new qualitative approach for ecosystem modeling (cf. 4.3) based on timed automata (TA) formalism combined to a high-level query language for exploring scenarios.

To date, EcoMata is dedicated to ecosystems that can be modeled as a collection of species (prey-predator systems) under various human pressures and submitted to environmental disturbances. It has two main parts: the Network Editor and the Query Launcher. The Network Editor let a stakeholder describe the trophic food web in a graphical way (the species icons and interactions between them). Only few ecological parameters are required and the user can save species in a library. The number of qualitative biomass levels is set as desired. An efficient algorithm generates automatically the network of timed automata. EcoMata provides also a dedicated window to help the user to define different fishing pressures, a nice way being by using chronograms. In the Query Launcher, the user selects the kind of query and the needed parameters (for example the species biomass levels to define a situation). Results are provided in a control panel or in files that can be exploited later. Several additional features are proposed in EcoMata: building a species library, import/export of ecosystem model, batch processing for long queries, etc. EcoMata is developed in Java (Swing for the GUI) and the model-checker called for the timed properties verification is UPPAAL.

The following website is devoted to the presentation of ECOMATA: <http://oban.agrocampus-ouest.fr:8080/ecomata>.

5.5. Paturmata

Paturmata is a tool-box for qualitative modeling and exploring agrosystems, specifically management of herd based on pasture [6]. The system is modelled using a hierarchical hybrid model described in timed automata formalism.

In PaturMata software, users can create a pasture system description by entering herds and plots information. For each herd, the only parameter is the number of animals. For each plot, users should enter the surface, the density, the herb height, the distance to the milking shed, a herb growth profile and an accessibility degree .

Users then specify pasturing and fertilization strategies. Finally, users can launch a pasture execution. PaturMata displays the results and a detailed trace of pasture. Users can launch a batch of different strategies and compare the results in order to find the best pasture strategy.

PaturMata is developed in Java (Swing for the GUI) and the model-checker that is called for the timed properties verification is UPPAAL.

Another feature which will be soon added to PaturMata is strategy synthesis. Users choose a pasture configuration or a type of pasture configuration and PaturMata proposes the best pasture and fertilization strategy in order to minimize the pasture procedure cost and use of nitrogen fertilizer.

5.6. ManageYourself

ManageYourself is a collaborative project between Dream and the Telelogos company aiming at monitoring smartphones from a stream of observations made on the smartphone state (cf. 3.2.3).

Today's smartphones are able to perform calls, as well as to realize much more complex activities. They are small computers. But as in computers, the set of applications embedded on the smartphone can lead to problems. The aim of the project ManageYourself is to monitor smartphones in order to avoid problems or to detect problems and to repair them.

The ManageYourself application includes three parts :

- A monitoring part which triggers preventive rules at regular time to insure that the system is working correctly, e.g. *if the memory is full then delete the tmp directory*. This part is always running on the smartphone.
- A reporting part which records regularly the state of the smartphone (the memory state - free vs allocated -, the connection state, which applications are running, etc.). This part also is always running on the smartphone. The current state is stored in a report at regular period and is labeled *normal*. When an application or the system bugs, the current buggy state is stored in a report and is labeled *abnormal*. At regular timestamps, all the reports are sent to a server where the learning process is executed.
- A learning part which learns new bug rules from the report dataset. This part is executed offline on the server. Once the bug rules are learnt, human experts translates them into preventive rules which are downloaded and integrated in the monitoring part of the smartphones.

The following website is devoted to the presentation of MANAGEYOURSELF: <http://www.irisa.fr/dream/ManageYourself/Site/ManageYourself.html>.

5.7. GeoImageRMP: a RapidMiner extension to georeferenced data

RapidMiner is one of the most used frontend for data mining, modelling and analysis. RapidMiner enables the user to design data processing tool chains interactively. A tool chain is a flow chart of processing tools represented by boxes in the interface. This software is easily extendable by designing Plugins. The GeoImageRMP plugin is a plugin dedicated to the design of tool chains to process georeferenced images (raster and vector images) [18]. It is a practical and useful respond to the analytic tasks of georeferenced data. This is the first plugin that is interested in including georeferenced data in RapidMiner and although the only user-friendly tool to create and compare georeferenced data tool chains. It benefits from the large amount of data processing tools that are already implemented in RapidMiner (classification, clustering, frequent pattern mining, etc.). One of the main aims of this plugin is to quickly prototype machine learning tools chain for remote sensing classification task. The GeoImageRMP plugin provide several new processing boxes:

- georeferenced data import/export: create and export dataset that can be processed by standard RapidMiner tool box from/to standard geospatial format (GeoTiff, Shapefiles)
- geospatial sampling method: based on multi-heterogeneous layers of georeferenced data, the sampling method can be transects, random, equidistant, from punctual layer.
- georeferenced data transformation tools: a set of tools dedicated to the manipulation of our new data structures (coordinates, SRS, etc.)
- visualization tool

The following website is devoted to the presentation of GEOIMAGERMP: <http://geoimagermp.gforge.inria.fr/>.

5.8. A plugin for visualizing and editing spatial graphs in QGis

Spatial graphs are accurate representations of spatial information through spatial objects linked by relationships (spatial or not). This representation is suited to the modeling and analysis of spatial information by computer processing (data mining, search for shortest paths, etc.). While Geographic information System (GIS), such as QGis, offers the possibility to visualize and manage georeferenced information, the use of spatial graph suffers from the lack of tools to facilitate the construction and integrated visualization.

We developed a QGIS plugin for the visualization and the interactive construction of spatial graphs. QGIS is the most used open source GIS. This plugin introduces a new type of layer: GraphLayer [16]. These new layers can be integrated into any GIS projects. They offer rich functionality for visualization and interactive editing.

5.9. Odisseptale: a software for implementing and evaluating sanitary event detectors in cattle

Odisseptale is a software for implementing disease detectors using monitoring of data provided by sensors placed on calves or cows. Sensors record streams of data such as body temperature, physical activity, feeding behavior, etc. These data are transmitted regularly to a monitoring software that aims to detect if a noticeable change has occurred on the data streams. Several detectors can be simultaneously active and each contribute to the final decision (detection of a disease). Two kinds of detectors have been implemented: a generic detector based on adaptive CUSUM and a symbolic pattern-based detector. Odisseptale provides also facilities for parameter setting and performance evaluation. This year, the software has been re-implemented in Python for enhanced portability and dissemination.

6. New Results

6.1. Diagnosis of large scale discrete event systems

Participants: Marie-Odile Cordier, Christine Largouët, Sophie Robin, Laurence Rozé, Yulong Zhao.

The problem we deal with is monitoring complex and large discrete-event systems (DES) such as an orchestration of web services or a fleet of mobile phones. Two approaches have been studied in our research group. The first one consists in representing the system model as a discrete-event system by an automaton. In this case, the diagnostic task consists in determining the trajectories (a sequence of states and events) compatible with the sequence of observations. From these trajectories, it is then easy to determine (identify and localize) the possible faults. In the second approach, the model consists in a set of predefined characteristic patterns. We use temporal patterns, called chronicles, represented by a set of temporally constrained events. The diagnostic task consists in recognizing these patterns by analyzing the flow of observed events.

6.1.1. *Distributed monitoring with chronicles - Interleaving diagnosis and repair - Making web services more adaptive*

Our work addresses the problem of maintaining the quality of service (QoS) of an orchestration of Web services (WS), which can be affected by exogenous events (i.e., faults). The main challenge in dealing with this problem is that typically the service where a failure is detected is not the one where a fault has occurred: faults have cascade effects on the whole orchestration of services. We have proposed a novel methodology to treat the problem that is not based on Web service (re)composition, but on an adaptive re-execution of the original orchestration. The re-execution process is driven by an orchestrator Manager that takes advantage of an abstract representation of the whole orchestration and may call a diagnostic module to localize the source of the detected failure. It is in charge of deciding the service activities whose results can be reused and may be skipped, and those that must be re-executed.

This year, we have improved the prototype, adding the visualization of the roadmap and the activities that do not have to be reexecuted. This work has been published in ICWS2013 [15] and we are working on a journal paper that will be submitted in 2014.

6.1.2. Scenario patterns for exploring qualitative ecosystems

This work aims at giving means of exploring complex systems, in our case ecosystems, and more recently agrosystems, specifically herd management systems. We proposed to transform environmental questions about future evolution of ecosystems into formalized queries that can be submitted to a simulation model. The system behavior is represented as a discrete event system described by a set of interacting timed automata, the global model corresponding to their composition on shared events. To query the model, we have defined high-level generic query patterns associated to the most usual types of request scenarios. These patterns are then translated into temporal logic formulas. The answer is computed thanks to model-checking techniques that are efficient for analyzing large-scale systems. Five generic patterns have been defined using TCTL (Timed Computation Tree Logic) “WhichStates”, “WhichDate”, “Stability”, “Always”, “Safety”. Three of them have been implemented using the model-checker UPPAAL.

The approach has first been experimented on a marine ecosystem under fishing pressure. The model describes the trophodynamic interactions between fish trophic groups as well as interactions with the fishery activities and with an environmental context. A paper has been previously published in the Environmental Modelling Software Journal [65]. More recently, a similar approach has been experimented on agrosystems, specifically herd management systems, for which a hybrid model has been built using hierarchical timed automata. This later work has been achieved in the context of Yulong Zhao’s PhD thesis [6] and done in collaboration with our colleagues of INRA.

6.1.3. Controller synthesis for dealing with “How to” queries

We extended the approach to deal with “How to” queries. As before, we rely on a qualitative model in the form of timed automata and on model-checking tools to answer queries. We proposed and compared two approaches to answer questions such as “How to avoid a given situation ?”(safety query). The first one exploits controller synthesis and the second one is a “generate and test” approach. We evaluated these two approaches in the context of an application that motivates this work, i.e. the management of a marine ecosystem and the evaluation of fishery management policies. The results have been previously published in [88].

More recently, we used similar methodological tools to analyze in the context of herd management on a catchment. An hybrid model has been built using hierarchical timed automata and scenarios can be simulated and evaluated using the approach presented in the previous paragraph. In this context, the goal is to identify and analyse the best/optimal farming practices in order to reduce nitrate pollution due to livestock effluents. We proposed to use controller synthesis tools and to couple them with machine learning techniques in order to get the best strategies and to put them on easy-to-use form. This work has been made in the context of Yulong Zhao’s PhD thesis [6] and in collaboration with our colleagues of INRA (UMR PEGASE).

6.2. Machine learning for model acquisition

Participants: Sid Ahmed Benabderrahmane, Marie-Odile Cordier, Thomas Guyet, Simon Malinowski, René Quiniou.

Model acquisition is an important issue for model-based diagnosis, especially while modeling dynamic systems. We investigate machine learning methods for temporal data recorded by sensors or spatial data resulting from simulation processes. Our main objective is to extract knowledge, especially sequential and temporal patterns or prediction rules, from static or dynamic data (data streams). We are particularly interested in mining temporal patterns with numerical information and in incremental mining from sequences recorded by sensors.

6.2.1. Representing and mining time series

Time series are sequences of numerical values, e.g. recorded by sensors. Since these series can be huge and subject to noise, they are often transformed into sequences of symbols. The best known symbolic transformation method is SAX (Symbolic Aggregate approXimation) [68]. SAX is based on a piecewise constant approximation method that does not take into account the slope of the time series values in successive windows. We have extended the SAX method by adding a symbolic slope information to the SAX symbols.

We have experimented our new representation, 1d-SAX, on three mining tasks. In most of these experiments 1d-SAX leads to a better accuracy than SAX [19].

We have also investigated a probabilistic representation of temporal patterns based on the latent Dirichlet allocation model (LDA). Such patterns can approximate the dynamics of a set of similar multivariate time series. We have experimented the method on hydrological flood time series to extract temporal patterns [7]. The extracted patterns were considered relevant and easy to understand by experts of the domain.

6.2.2. Incremental sequential mining

Sequential pattern mining algorithms operating on data streams generally compile a summary of the data seen so far from which they compute the set of actual sequential patterns. We propose another solution where the set of actual sequential patterns are incrementally updated as soon as new data arrive on the input stream. Our work stands in the framework of mining an infinite unique sequence. Our method [60] provides an algorithm that maintains a tree representation (inspired by the PSP algorithm [71]) of frequent sequential patterns and their minimal occurrences [69] in a window that slides along the input data stream. It makes use of two operations: deletion of the itemset at the beginning of the window (obsolete data) and addition of an itemset at the end of the window (new data). The experiments were conducted on simulated data and on real data of instantaneous power consumption. The results show that our incremental algorithm significantly improves the computation time compared to a non-incremental approach [61].

Recently, we have worked on the adaptation of our algorithm to closed sequential patterns. A closed pattern is a local maximal pattern such there exists no extension of this pattern having the same support. Closed patterns are known to provide a condensed representation of the solution patterns and lead to more efficient algorithms without losing information or completeness on extracted patterns. The tree of closed-patterns is less deep than the pattern-tree but the transformations of the tree by addition or deletion of items are more complex. The algorithm is under evaluation. We plan to submit a paper in 2014.

6.2.3. Multiscale segmentation of satellite image time series

Satellite images allow the acquisition of large-scale ground vegetation. Images are available along several years with a high acquisition frequency (1 image every two weeks). Such data are called satellite image time series (SITS). In [58], we presented a method to segment an image through the characterization of the evolution of a vegetation index (NDVI) on two scales: annual and multi-year. The main issue of this approach was the required computation resources (time and memory). We first propose to adapt image segmentation algorithm to SITS. Segmented images reduces the number of time series to analyze and the computation time. We secondly applied 1D-SAX to reduce data dimensionality [20]. We evaluated this approach on the supervised classification of large SITS of Senegal and we showed that 1D-SAX approaches the classification results of time series while significantly reducing the required memory storage of the images.

6.2.4. Analysis of landscape based on spatial patterns

Researchers in agro-environment need a great variety of landscapes to test the agro-ecological models of their scientific hypotheses. Real landscapes are difficult to acquire and do not enable the agronomist to test all their hypothesis. Working with simulated landscapes is then an alternative to get a sufficient variety of experimental data. Our objective is to develop an original scheme to generate realistic landscapes. This approach is based on a spatial representation of landscapes by a graph expressing the spatial relationships between the agricultural parcels (as well as the roads, the rivers, the buildings, etc.), of a specific geographic area. We extract spatial patterns from a real geographic area and we use these patterns to generate new realistic landscapes. Using patterns preserves the interface properties between parcels.

We have begun the exploration of graph mining techniques, such as gSPAN [87], to discover the relevant spatial patterns present in a spatial-graph. But the graph-mining techniques are very time-consuming in comparison to sequence mining.

This year, we would like to test if using a path instead of a graph would be a faithful representation of the spatial organization of the landscape. In [17], we compare the potential expressivity of graphs and Hilbert-Peano curves [66] to characterize an agricultural landscape. The results show that mining frequent patterns in Hilbert-Peano curves would be as discriminant as mining frequent patterns in graphs.

The perception of the environment is an important dimension of the landscape we live in. One of our objectives is to study the relationships between the landscape patterns and their perception. We cope with this dimension by analysing the textual content of ‘atlas du paysage’ (landscape atlas), that are produce by each french administrative regions. This year we worked on the construction of an ontology of landscape perception [21].

6.2.5. *Subdimensional clustering for fast similarity search over time series data. Application to Information retrieval tasks*

Information retrieval and similarity search tasks in time series databases remains a challenge that require to discover relevant pattern-sequences that are recurrent over the overall time series sequences, and to find temporal associations among these frequently occurring patterns. Previous work on information retrieval and similarity search in time series has been performed in different contexts such as diagnosis or failure detection of industrial materials. In whole query matching, a time series given as query is entirely compared to every time series of a database. The series should have same length, and a similarity measure is used to retrieve either a most similar time series or the top-k ranked time series. However, these methods suffer from a lack of flexibility of the used similarity measures, a lack of scalability of the representation model, and a penalizing runtime to retrieve the information. Moreover, in some real world applications, one can be interested in retrieving specific interesting subsequences that are frequently present at different instants.

Motivated by these observations, we have designed a framework tackling the query by content problem on time series data, ensuring (i) fast response time, (ii) multi-level information representation, and (iii) representing temporal associations between extracted patterns. During the preparation step, all the multi-valued time series present in the database are transformed into a multi-resolution symbolic representation thus ensuring a lower dimensionality. Then, to accelerate and enhance the similarity search and the retrieval over the database, our model creates an index over recurrent patterns in the time series collection. These patterns can be generated by different techniques. Finally, the extracted patterns are grouped by clustering and the resulting clusters are indexed in a table within their centroids. A paper presenting the preliminary results is under submission to an international journal.

6.2.6. *Knowledge Extraction from Heterogeneous Data*

Recently, mining microarrays data has become a big challenge due to the growing sources of available data. We are using machine learning methods such as clustering, dimensionality reduction, association rules discovery on transcriptomic data, by combining a domain ontology as source of knowledge, in order to supervise the KDD process. Our objectives concern the identification of genes that could participate in the development of tumors. A two-way classification method was proposed, combining genes expression levels, represented as numerical data, and Gene Ontology (GO) annotations as symbolic data. The hopeful results obtained with genes clustering, through GO annotations, are an encouraging track to predict transcriptional regulatory networks, and for refining the existing sets of genes [11], [12].

We also introduced a new method for extracting enriched biological functions from transcriptomic databases using an integrative bi-classification approach. The initial gene datasets are firstly represented as a formal context (objects attributes), where objects are genes, and attributes are their expression profiles and complementary information of different knowledge bases. Formal Concept Analysis (FCA) is applied for extracting formal concepts regrouping genes having similar transcriptomic profiles and functional behaviors. An enrichment analysis is then performed in order to identify the relevant formal concepts from the generated Galois lattice, and to extract biological functions that could participate in the proliferation of cancers. Preliminary results seem very promising, and could help experts during the identification of degenerated biological functions [13].

6.3. Decision aiding with models and simulation data

Participants: Louis Bonneau de Beaufort, Tassadit Bouadi, Marie-Odile Cordier, Véronique Masson, René Quiniou.

Models can be very useful for decision aiding as they can be used to play different plausible scenarios for generating the data representing future states of the modeled process. However, the volume of simulation data may be huge. Thus, efficient tools must be investigated in order to store the simulation data, to focus on relevant parts of the data and to extract interesting knowledge from these data.

6.3.1. A datawarehouse for simulation data

The ACASSYA project 8.2.1 aims at providing experts or stakeholders or farmers with a tool to evaluate the impact of agricultural practices on water quality. As the simulations of the deep model TNT2 are time-consuming and generate huge data, we have proposed to store these simulation results in a datawarehouse and to extract relevant information, such as prediction rules, from the stored data. We have devised a general architecture for agro-environmental data on top of the framework Pentaho.

This year we have been working on the efficient computation of OLAP queries related to realistic scenarios proposed by experts in the domain. Precisely, we have devised indexing schemes to access the data in the OLAP cube. We have also worked on the visualization by a GIS (Geographical Information System) of the query results on maps of the geographical area under interest. A paper have been submitted to the COMPAG Journal. This work is detailed in Tassadit Bouadi's thesis [5].

6.3.2. Efficient computation of skyline queries in an interactive context

Skyline queries retrieve from a database the objects that maximize some criteria, related to user preferences for example, or objects that are the best compromises satisfying these criteria. When data are in huge volumes, such objects may shed light on interesting parts of the dataset. However, computing the skylines (i.e. retrieving the skyline points) may be time consuming because of many dominance tests. This is, especially the case in an interactive setting such as querying a data cube in the context of a datawarehouse.

We have worked at improving the formal setting of the partial materialization of skyline queries when dynamic preferences are refined online by the user. We have explicitated which parts of the skyline evolve (which point are added or removed) when a new dimension is introduced in the computation. This led to an efficient incremental method for the online computation of the skyline corresponding to new user preferences [46]. An extended version of this paper is published in Journal "Transactions on Large Scale Data and Knowledge Centered Systems" (TLDKS) [8] and in Tassadit Bouadi's thesis [5].

6.3.3. Hierarchical skylines

Conventional skyline queries retrieve the skyline points in a context of dimensions with a single hierarchical level. However, in some applications with multidimensional and hierarchical data structure (e.g. data warehouses), skyline points may be associated with dimensions having multiple hierarchical levels. Thus, we have proposed an efficient approach reproducing the effect of the OLAP operators "drill-down" and "roll-up" on the computation of skyline queries. It allows the user to navigate along the dimensions hierarchies (i.e. specialize / generalize) while ensuring an online calculation of the associated skyline. The method is described in Tassadit Bouadi's thesis [5]. A paper describing this contribution is currently under submission to the "Very Large Data Bases (VLDB 2014)" conference.

6.3.4. Modeling influence propagation by Bayesian causal maps

The goal of this project is modeling shellfish fishing to assess the impact of management pollution scenarios on the *Rade de Brest*. Cognitive maps were built from interviews with fishermen. To represent and reason about these cognitive maps, we propose to use Bayesian Causal Maps making use of fishermen knowledge, particularly to perform influence propagation [82].

However, this model does not take into account the variety of influences asserted by the fishermen, but only the "mean" causal map. A report describing the project is available [28]. An approach that could combine individual knowledge with belief functions in the way of Philippe Smets's Transferable Belief Model [83] has been proposed. A report describing the project is available [28].

This work is done in the framework of the RADE2BREST project, involving Agrocampus Ouest and CNRS (GEOMER/LETG), funded by "Ministère de l'Ecologie" (This project is not mentioned in section 8.2 because DREAM is not an official partner of this project.).

6.3.5. Recommending actions from classification rules

In the framework of the SACADEAU project, a paper dedicated to building actions from classification rules has been published in the KAIS Journal [9]. Our goal is to burden of analysing a large set of classification rules when the user is confronted to an "unsatisfactory situation" and needs help to decide about the appropriate actions to remedy to this situation. The method consists in comparing the situation to a set of classification rules. For this purpose, we propose DAKAR, a new framework for learning action recommendations dealing with complex notion of feasibility and quality of actions.

Sacadeau-Software, which is the decision support tool implemented with F. Ployette (former Inria engineer in the EPI Dream, now retired) in the SACADEAU project, has been published in the RIA Journal [10]. Sacadeau-Software allows to run simulations throughout a watershed and obtain the transfer rate of pollution through the catchment. Classification rules, characterizing the sub-parts of the watershed with pollution and the sub-parts without pollution, are automatically learned from the simulations. A visualization tool enables to relate the learned rules to the examples characterized by these rules. Finally, a user can select a situation of pollution and the action recommendation tool analyses the learned rules and proposes actions that improve this situation of pollution.

6.4. Diagnostic, causal reasoning and argumentation

Participants: Philippe Besnard, Marie-Odile Cordier, Yves Moinard.

Stemming on [38], [39], [40], [41], [42], we have designed an inference system based on causal statements. This is related to diagnosis (observed symptoms explained by faults). The aim is to produce possible explanations for some observed facts. Previously existing proposals were ad-hoc or, as in [45], [57], they were too close to standard logic to make a satisfactory diagnosis. A key issue for this kind of work is to distinguish logical implication from causal links and from ontological links. This is done by introducing a simple causal operator, and an *is-a* hierarchy. These two operators are added to a restricted first order logic of the Datalog kind (no function symbols). Then, our system produces elementary *explanations* for some set of observed facts. Each explanation links some facts to the considered observation, together with a set of atoms called the *justifications*: The observation is explained from these facts, provided the justifications are possible (not contradicted by the available data). This formalism has been translated into answer set programming [72], [73]). It is able to deal with complex problems such as finding explanations for the hurricane Xynthia (2010, February 28). In such situations, there are many data and many possible elementary explanations can be examined. This involves an extension of our formalism, in order to deal with more complex chains of causations and *is-A* links. Our formalism makes precise what all these possible explanations are. Then, in order to deal with so many possible complex explanations, we integrate this causal formalism into an argumentation framework. Logic-based formalizations of argumentation [43] take pros and cons for some conclusion into account. These formalizations assume a set of formulae and then exhaustively lay out arguments and counterarguments. This involves providing an initiating argument for the inference and then providing undercuts to this argument, and then undercuts to undercuts. So here our causal formalism provides a (rather large) set of explanations, and the argumentation part allows to select the best ones, under various criteria [22], [14].

Then, since answer set programming can easily deal with logical formalisms, the argumentation part will be incorporated into our already existing answers set programming translation of the causal formalism. Regarding this field of knowledge representation and reasoning, and more generally, artificial intelligence, we have participated to several chapters in the to be published "Panorama de l'intelligence artificielle. Ses bases méthodologiques, ses développements" [27], [26], [23], [24].

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

7.1.1. *ManageYourSelf: diagnosis and monitoring of embedded platforms*

Participants: Marie-Odile Cordier, Sophie Robin, Laurence Rozé.

ManageYourSelf is a project that deals with the diagnosis and monitoring of embedded platforms, in the framework of a collaboration with Telelogos, a French company expert in mobile management and data synchronization. ManageYourSelf aims to perform diagnostic and repair on a fleet of mobile smartphones and PDAs. The idea is to embed on the mobile devices a rule-based expert system and its set of politics, for example "if memory full then delete (directory). recognition is performed, using the parameters of the phones as the fact base. Of course, it is impossible to foresee all the rules in advance. Upon detection of a non anticipated problem, a report containing all the system's information prior to the problem is sent to a server. The learning step was first implemented using decision trees, the aim being to characterize the faults and consequently update the global knowledge base and its distributed instances. An incremental version of this learning step has been studied in order to get an on-line process [36]. This means being able to learn new faults characterizations and add new preventive rules, and also forget no longer needed ones.

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. *Projet RTR: Coupling observation/simulation for decision-aid in environment complex systems*

Participants: Sid Ahmed Benabderrahmane, Marie-Odile Cordier, Thomas Guyet, Simon Malinowski, René Quiniou.

This RTR (Réseaux Thématiques de Recherche - Thematic Research Networks) project is a collaboration between COSTEL (UMR LETG, Rennes), the team Obelix (IRISA, Vannes), UMR SAS (INRA, Rennes) and the EPI Dream. The project began in 2013 and has been funded for one year. It aims at studying the relationships between observations and simulations. The objective is to better understand what one side can provide to the other side in order to improve decision-making. This project gathers partners having expertise and skills in teledetection and image analysis, in modeling and simulation, and in knowledge acquisition for aiding decision in environmental research. The targeted applications belong to the domain of hydrology and agriculture. A final workshop (<http://tinyurl.com/k3smbox>) has been organized in november.

8.2. National Initiatives

8.2.1. *ACASSYA: Supporting the agro ecological evolution of breeding systems in coastal watersheds*

Participants: Marie-Odile Cordier, Véronique Masson, René Quiniou.

The ACASSYA project (ACcompagner l'évolution Agro-écologique deS SYstèmes d'élevage dans les bassins versants côtiers) is funded by ANR/ADD. It started at the beginning of 2009 and will end in June 2013. The main partners are our colleagues from INRA (SAS from Rennes. One of the objectives is to develop modeling tools supporting the management of ecosystems, and more precisely the agro ecological evolution of breeding systems in coastal watersheds. In this context, the challenge is to transform existing simulation tools (as SACADEAU or TNT2 into decision-aid tools, able to answer queries or scenarios about the future evolution of ecosystems. (<http://tinyurl.com/ptzdqo5>)

8.2.2. Asterix : spatio-temporal analysis of remote sensing images

Participant: Thomas Guyet.

The ASTERIX project (Analyse Spatio-temporelle pour la Télédétection de l'Environnement par Reconnaissance dans les Images complexes) is funded by ANR/JCJC. The project leader is S. Lefèvre from the IRISA/Vannes Team Obelix. The other partners are OSUR/University of Rennes-2, the Laboratory Image, Ville, Environnement (LIVE), University of Strasbourg, DYNAFOR (INRA/ENSAT), Toulouse and Institut de Physique du Globe de Strasbourg (IPGS), University of Strasbourg. The project started at the end of 2013 (<http://anr-asterix.irisa.fr/>) and will end in 2017.

The goal of the ASTERIX project is to provide methods, algorithms and software in the field of image analysis and machine learning/data mining to support the analysis of remote sensing images. The project addresses the specific issues of such data: dimensionality, heterogeneity, volume, spatio-temporal nature and the temporal evolution. It is dedicated to the field of environmental remote sensing and deals with concrete applications such as the evolution of the coastline or the colonization of grasslands by ash.

Our contribution to this project will be the proposition of data mining algorithms to deal with the spatio-temporal dimensions of satellite image time series.

8.3. International Initiatives

8.3.1. Inria International Partners

8.3.1.1. Declared Inria International Partners

- University of Potsdam, Germany. Prof. Torsten Schaub has been awarded an Inria international senior grant from 2103 to 2017.

8.3.1.2. Informal International Partners

- University of Calgary, Canada. Dr Edouard Timsit, Dept. of production Animal Health, Faculty of Veterinary, Medicine.
- University di Torino, Italy, Dr Roberto Micalizio, Dept. of Computer Science.

9. Dissemination

9.1. Scientific Animation

9.1.1. Journal editorial board

- *AAI: Applied Artificial Intelligence* (M.-O. Cordier).
- *Interstices webzine* (M.-O. Cordier).
- *Revue d'Intelligence Artificielle*: Invited editors of the special issue on "Intelligence Artificielle and Agronomy": T. Guyet, F. Le Ber, M. Teisseire [25]

9.1.2. Conference program committees and organizations

- Program committee members of FOSTA workshop at EGC 2013 (T. Guyet, R. Quiniou).

- Organization committee member of CIDN workshop at EGC 2013 (T. Guyet).
- Local chair of EGC 2014 in Rennes (R. Quiniou).
- Organization committee member of EGC 2014 in Rennes (T. Guyet).
- Organization chairs and Program committee members of FST-CERGEIO workshop at EGC 2014 (T. Guyet, R. Quiniou).
- Steering Committee of RFIA'2014 (T. Guyet).
- Session chair of ICBBIO 2014 (S. Benabderrahmane)

9.1.3. Scientific and administrative boards

- ECCAI fellow + Honoric member of AFIA (Association Française d'Intelligence Artificielle): M. - O. Cordier
- Member of "Agrocampus-Ouest" scientific board: M.-O. Cordier.
- Member of "Conseil d'administration de l'ISTIC": M.-O. Cordier.
- Head of IRISA department "Data and Knowledge Management" and member of the IRISA scientific management committee: M.-O. Cordier
- Member of the "Prix de thèse AFIA 2013" award committee (selects the best French PhD thesis in the Artificial Intelligence domain): M.-O. Cordier.
- Chair of the INRA CSS-MBIA (Commission scientifique spécialisée "Mathématiques, Biologie et Intelligence Artificielle"): M.-O. Cordier.
- Member of the CoNRS (Comité national recherche scientifique (since october 2012) : M.-O. Cordier
- Member of the AFIA board (since october 2011): T. Guyet.
- Member of the COREGE (Research Committee-COMITÉ de la REcherche du Grand Etablissement) of Agrocampus-Ouest: T. Guyet.
- Evaluator for the "Fonds de recherche du Québec – Nature et technologies": T. Guyet.
- Member of the Payote-Network board: T. Guyet.

9.1.4. Misc

- Invited talk at the "Pattern Mining Days" workshop, Grenoble, July 2013 (T. Guyet).

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Many members of the EPI DREAM are also faculty members and are actively involved in computer science teaching programs in ISTIC, INSA and Agrocampus-Ouest. Besides these usual teachings DREAM is involved in the following programs:

Master: *Module DSS: Apprentissage sur des données séquentielles symboliques*, 10 h, M2, ISTIC University of Rennes (R. Quiniou).

Master: *C++ Programming*, M1, ENSAI, Rennes (T. Guyet and S. Benabderrahmane),

Master: *Géoinformation*, M2, Agrocampus Ouest Rennes (L. Bonneau, T. Guyet, S. Malinowski)

9.2.2. Supervision

PhD: Tassadit Bouadi, “Analyse interactive de résultats de simulation et découverte de connaissances. Application à l’aide à la décision dans le domaine agroécologique pour l’amélioration de la qualité des eaux des bassins versants”, november 28th 2013, co-supervisors Marie-Odile Cordier, René Quiniou and Chantal Gascuel, ANR project Acassya grant

PhDs: Yulong Zhao, “Modélisation d’agroécosystèmes dans un formalisme de type systèmes à événements discrets et simulation de scénarios utilisant des outils de model-checking. Application à l’étude des impacts des changements climatiques et des pratiques agricoles sur les flux de nutriments vers les eaux de surface.”, january 13th 2014, supervisor Marie-Odile Cordier and Chantal Gascuel

PhD in progress: Philippe Rannou, “Modèle rationnel pour humanoïdes virtuels”, october 1st 2010, co-supervisors Marie-Odile Cordier and Fabrice Lamarche

9.2.3. Juries

- Committee member and reviewer of Nicolas Méger’s HDR defence (Univ. Savoie): M.-O. Cordier
- Committee member of Alexandre Termier’s HDR defence (Univ. Grenoble): M.-O. Cordier
- Committee member of Tassadit Bouadi’s PhD defence (Université de Rennes 1): M.-O. Cordier , R. Quiniou.
- Thesis committee of Pierre Clément, Oniris Nantes, “Valeur informative du budget-temps de la vache laitière pour la détection précoce des troubles de santé”. R. Quiniou.
- Bachelor: A. Fratila (Cluj-Napoca University, Romania), "Evaluating similarity measures between time series symbolic sequences", September 2013, supervisors: Sidahmed Benabderrahmane and René Quiniou.

10. Bibliography

Major publications by the team in recent years

- [1] P. BESNARD, M.-O. CORDIER, Y. MOINARD. *Ontology-based inference for causal explanation*, in "Integrated Computer-Aided Engineering", 2008, vol. 15, n^o 4, pp. 351-367, <http://hal.inria.fr/inria-00476906/en/>
- [2] C. GASCUEL-ODOUX, P. AUROUSSEAU, M.-O. CORDIER, P. DURAND, F. GARCIA, V. MASSON, J. SALMON-MONVIOLA, F. TORTRAT, R. TRÉPOS. *A decision-oriented model to evaluate the effect of land use and agricultural management on herbicide contamination in stream water*, in "Environmental modelling & software", 2009, vol. 24, pp. 1433-1446, <http://hal.inria.fr/hal-00544122/en>
- [3] T. GUYET, R. QUINIOU. *Extracting temporal patterns from interval-based sequences*, in "International Joint Conference on Artificial Intelligence (IJCAI)", Barcelone, Spain, July 2011, <http://hal.inria.fr/inria-00618444>
- [4] Y. PENCOLÉ, M.-O. CORDIER. *A formal framework for the decentralised diagnosis of large scale discrete event systems and its application to telecommunication networks*, in "Artificial Intelligence Journal", 2005, vol. 164, n^o 1-2, pp. 121-170, <http://hal.inria.fr/inria-00511104/en/>

Publications of the year

Doctoral Dissertations and Habilitation Theses

- [5] T. BOUADI. , *Analyse multidimensionnelle interactive de résultats de simulation. Aide à la décision dans le domaine de l’agroécologie*, Université Rennes 1, November 2013, <http://hal.inria.fr/tel-00933375>

- [6] Y. ZHAO. , *Modélisation qualitative des agro-écosystèmes et aide à leur gestion par utilisation d'outils de model-checking*, Université Rennes 1, January 2014, <http://hal.inria.fr/tel-00933443>

Articles in International Peer-Reviewed Journals

- [7] A. AUBERT, R. TAVENARD, R. EMONET, A. DE LAVENNE, S. MALINOWSKI, T. GUYET, R. QUINIOU, J.-M. ODOBEZ, P. MÉROT, C. GASCUEL-ODOUX. *Clustering flood events from water quality time-series using Latent Dirichlet Allocation model*, in "Water Resources Research", 2013, 1 p. [DOI : 10.1002/2013WR014086], <http://hal.inria.fr/halshs-00906292>
- [8] T. BOUADI, M.-O. CORDIER, R. QUINIOU. *Computing Skyline Incrementally in Response to Online Preference Modification*, in "Transactions on Large-Scale Data- and Knowledge- Centered Systems", 2013, vol. 10, pp. 34-59 [DOI : 10.1007/978-3-642-41221-9_2], <http://hal.inria.fr/hal-00920548>
- [9] R. TRÉPOS, A. SALLEB-AOUISSI, M.-O. CORDIER, V. MASSON, C. GASCUEL-ODOUX. *Building Actions From Classification Rules*, in "Knowledge and Information Systems (KAIS) journal", 2013, vol. 34, n^o 2, pp. 267-298 [DOI : 10.1007/s10115-011-0466-5], <http://hal.inria.fr/hal-00649388>

Articles in National Peer-Reviewed Journals

- [10] V. MASSON, F. PLOYETTE, M.-O. CORDIER, C. GASCUEL-ODOUX, R. TRÉPOS. *Sacadeau-Software, un logiciel d'aide à la décision pour améliorer la qualité de l'eau*, in "Revue d'Intelligence Artificielle", September 2013, vol. 27, n^o 4-5, pp. 443-469 [DOI : 10.3166/RIA.27.443-469], <http://hal.inria.fr/hal-00881658>

International Conferences with Proceedings

- [11] S. BENABDERRAHMANE. *Biomedical Knowledge Extraction Using Fuzzy Differential Profiles and Semantic Ranking*, in "14th Conference on Artificial Intelligence in Medicine, AIME 2013", Murcia, Spain, N. PEEK (editor), June 2013, vol. 7885, pp. 84-93 [DOI : 10.1007/978-3-642-38326-7_13], <http://hal.inria.fr/hal-00934265>
- [12] S. BENABDERRAHMANE. *Enhancing Transcriptomic Data Mining with Semantic Ranking: Towards a new Functional Spectral Representation*, in "International Work-Conference on Bioinformatics and Biomedical Engineering, IWBBIO 2013", Granada, Spain, ISBN 978-84-15814-13-9, March 2013, vol. 15814, pp. 978-84, <http://hal.inria.fr/hal-00934279>
- [13] S. BENABDERRAHMANE. *Formal Concept Analysis and Knowledge Integration for Highlighting Statistically Enriched Functions from Microarrays Data*, in "International Work-Conference on Bioinformatics and Biomedical Engineering, IWBBIO 2014", Granada, Spain, April 2014, 1 p. , <http://hal.inria.fr/hal-00935378>
- [14] P. BESNARD, M.-O. CORDIER, Y. MOINARD. *Arguments using ontological and causal knowledge*, in "FoKS", Bordeaux, France, C. BEIERLE, C. MEGHINI (editors), LNCS, Springer, March 2014, vol. 8367, pp. 79-96, <http://hal.inria.fr/hal-00931679>
- [15] M.-O. CORDIER, R. MICALIZIO, S. ROBIN, L. ROZÉ. *Adapting Web Services to Maintain QoS Even When Faults Occur*, in "ICWS - IEEE 20th International Conference on Web Services", Santa Clara - CA, United States, 2013, pp. 403-410 [DOI : 10.1109/ICWS.2013.61], <http://hal.inria.fr/hal-00920474>

- [16] T. GUYET. *Visualisation de données relationnelles*, in "Conférence Internationale de Géomatique et d'Analyse Spatiale (SAGEO)", Brest, France, September 2013, <http://hal.inria.fr/hal-00916923>
- [17] T. GUYET, F. LE BER, S. DA SILVA, C. LAVIGNE. *Comparaison des chemins de Hilbert adaptatif et des graphes de voisinage pour la caractérisation d'un parcellaire agricole*, in "Conférence Extraction et Gestion de Connaissances", Rennes, France, January 2014, <http://hal.inria.fr/hal-00916964>
- [18] T. GUYET, H. NICOLAS, B. GHEDAMSI, E. ATHANE. *Fouille d'images géoréférencées avec RapidMiner*, in "Conférence Internationale de Géomatique et d'Analyse Spatiale (SAGEO)", Brest, France, September 2013, <http://hal.inria.fr/hal-00916921>
- [19] S. MALINOWSKI, T. GUYET, R. QUINIOU, R. TAVENARD. *1d-SAX: A Novel Symbolic Representation for Time Series*, in "International Symposium on Intelligent Data Analysis", United Kingdom, 2013, pp. 273-284 [DOI : 10.1007/978-3-642-41398-8_24], <http://hal.inria.fr/halshs-00912512>
- [20] S. MALINOWSKI, T. GUYET, R. QUINIOU, R. TAVENARD. *1d-SAX : une nouvelle représentation symbolique pour les séries temporelles*, in "Conférence Extraction et Gestion de Connaissances", Rennes, France, January 2014, <http://hal.inria.fr/hal-00916970>

National Conferences with Proceedings

- [21] J. BOURBEILLON, L. PIEL, R. EL AYEB, D. ROUSSELIÈRE, T. GUYET. *Construction semi-automatique d'une ontologie de la perception des paysages*, in "Conférence Ingénierie des Connaissances 2013", Lille, France, May 2013, pp. 1-10, <http://hal.inria.fr/hal-00920722>

Conferences without Proceedings

- [22] P. BESNARD, M.-O. CORDIER, Y. MOINARD. *Arguments using ontological and causal knowledge*, in "JIAF 2013 (Septièmes Journées de l'Intelligence Artificielle Fondamentale)", Aix-en-Provence, France, S. KONIECZNY, N. MAUDET (editors), April 2013, pp. 41-48, <http://hal.inria.fr/hal-00932294>

Scientific Books (or Scientific Book chapters)

- [23] L. AMGOUD, P. BESNARD, C. CAYROL, P. CHATALIC, M.-C. LAGASQUIE-SCHIEX. *Argumentation et raisonnement en présence de contradictions*, in "Panorama de l'intelligence artificielle Ses bases méthodologiques, ses développements", P. MARQUIS, O. PAPINI, H. PRADE (editors), Cépaduès, January 2013, vol. 2, <http://hal.inria.fr/hal-00770580>
- [24] M.-O. CORDIER, P. DAGUE, Y. PENCOLÉ, L. TRAVÉ-MASSUYÈS. *Diagnostic et supervision : approches à base de modèles*, in "Panorama de l'intelligence artificielle : Ses bases méthodologiques, ses développements", P. MARQUIS, O. PAPINI, H. PRADE (editors), Cépaduès, January 2013, vol. 2, <http://hal.inria.fr/hal-00769636>
- [25] T. GUYET, F. LE BER, M. TEISSEIRE. , *Intelligence artificielle et agronomie*, Revue d'Intelligence Artificielle, July 2013, vol. 27, 273 p. , <http://hal.inria.fr/hal-00916980>
- [26] A. HERZIG, P. BESNARD. *Représentation des connaissances : modalités, conditionnels et raisonnement non monotone*, in "Panorama de l'intelligence artificielle Ses bases méthodologiques, ses développements", P. MARQUIS, O. PAPINI, H. PRADE (editors), Cépaduès, January 2013, vol. 2, <http://hal.inria.fr/hal-00770563>

- [27] A. LALLOUET, Y. MOINARD, P. NICOLAS, I. STÉPHAN. *Programmation logique*, in "Panorama de l'intelligence artificielle Ses bases méthodologiques, ses développements", P. MARQUIS, O. PAPINI, H. PRADE (editors), Cépaduès, January 2013, vol. 2, <http://hal.inria.fr/hal-00758896>

Research Reports

- [28] F. GOURMELON, D. LE GUYADER, G. FONTENELLE, H. LEVREL, C. TISSOT, L. BONNEAU DE BEAUFORT, M. ROUAN. , *Modélisation et scénarisation des activités humaines en rade de Brest*, February 2013, 93 p. , <http://hal.inria.fr/hal-00797154>

References in notes

- [29] B. DUBUISSON (editor). , *Diagnostic, intelligence artificielle et reconnaissance des formes*, Traité IC2 : Information - Commande - Communication, Hermes, 2001
- [30] S. DZEROSKI, N. LAVRAČ (editors). , *Relational Data Mining*, SpringerBerlin, 2001
- [31] W. HAMSCHER, L. CONSOLE, J. DE KLEER (editors). , *Readings in Model-Based Diagnosis*, Morgan KaufmannSan Mateo, CA, Etats-Unis, 1992
- [32] C. AGGARWAL. , *Data Streams: Models and Algorithms*, Advances in Database Systems, Springer, 2007
- [33] R. AGRAWAL, T. IMIELINSKI, A. N. SWAMI. *Mining Association Rules between Sets of Items in Large Databases*, in "Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data", Washington, D.C., P. BUNEMAN, S. JAJODIA (editors), 26–28 1993, pp. 207–216
- [34] R. AGRAWAL, R. SRIKANT. *Mining sequential patterns*, in "Eleventh International Conference on Data Engineering", Taipei, Taiwan, P. S. YU, A. S. P. CHEN (editors), IEEE Computer Society Press, 1995, pp. 3–14
- [35] G. ALARME. *Monitoring and alarm interpretation in industrial environments*, in "AI Communications", 1998, vol. 11, 3-4, pp. 139-173, S. Cauvin, Marie-Odile Cordier, Christophe Dousson, P. Laborie, F. Lévy, J. Montmain, M. Porcheron, I. Servet, L. Travé-Massuyès
- [36] M. L. ANGHELOIU. , *Incremental and adaptive learning for online monitoring of embedded software*, Master 2 recherche informatique, Université de Rennes 1, June 2012, <http://dumas.ccsd.cnrs.fr/dumas-00725171>
- [37] M. BASSEVILLE, M.-O. CORDIER. , *Surveillance et diagnostic de systèmes dynamiques : approches complémentaires du traitement de signal et de l'intelligence artificielle*, Irisa, 1996, n^o 1004, <ftp://ftp.irisa.fr/techreports/1996/PI-1004.ps.gz>
- [38] P. BESNARD, M.-O. CORDIER. *Inferring causal explanations*, in "Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU'99)", A. HUNTER, S. PARSONS (editors), Lecture Notes in Artificial Intelligence, Springer-Verlag, 1999, vol. 1638, pp. 55-67
- [39] P. BESNARD, M.-O. CORDIER, Y. MOINARD. *Configurations for Inference between Causal Statements*, in "KSEM 2006 (First Int. Conf. on Knowledge Science, Engineering and Management)", J. LANG, F. LIN, J. WANG (editors), LNAI, Springer, aug 2006, n^o 4092, pp. 292–304, <http://www.irisa.fr/dream/dataFiles/moinard/causeksempub.pdf>

- [40] P. BESNARD, M.-O. CORDIER, Y. MOINARD. *Ontology-based inference for causal explanation*, in "KSEM07 (Second International Conference on Knowledge Science, Engineering and Management)", Z. ZHANG, J. SIEKMANN (editors), LNAI, Springer, nov 2007, n^o 4798, pp. 153-164, <http://www.irisa.fr/dream/dataFiles/moinard/ksem07bcausesonto.pdf>
- [41] P. BESNARD, M.-O. CORDIER, Y. MOINARD. *Deriving explanations from causal information*, in "ECAI 2008 (18th European Conference on Artificial Intelligence)", Patras, Greece, M. GHALLAB, C. D. SPYTOPOULOS, N. FAKOTAKIS, N. AVOURIS (editors), IOS Press, jul 2008, pp. 723–724
- [42] P. BESNARD, M.-O. CORDIER, Y. MOINARD. *Ontology-based inference for causal explanation*, in "Integrated Computer-Aided Engineering", 2008, vol. 15, n^o 4, pp. 351-367
- [43] P. BESNARD, A. HUNTER. , *Elements of Argumentation*, The MIT Press, june 2008, <http://www.mitpress.mit.edu/>
- [44] C. BIERNACKI, G. CELEUX, G. GOVAERT, F. LANGROGNET. *Model-Based Cluster and Discriminant Analysis with the MIXMOD Software*, in "Computational Statistics and Data Analysis", 2006, vol. 51, n^o 2, pp. 587-600
- [45] A. BOCHMAN. *A Causal Theory of Abduction*, in "7th Int. Symposium on Logical Formalizations of Common Sense Reasoning", S. MCILRAITH, P. PEPPAS, M. THIELSCHER (editors), 2005, pp. 33–38
- [46] T. BOUADI, M.-O. CORDIER, R. QUINIOU. *Incremental Computation of Skyline Queries with Dynamic Preferences*, in "Database and Expert Systems Applications (DEXA)", Vienne, Austria, S. W. LIDDLE, K.-D. SCHEWE, A. M. TJOA, X. ZHOU (editors), Springer, 2012, vol. 1, pp. 219-233 [DOI : 10.1007/978-3-642-32600-4_17], <http://hal.inria.fr/hal-00757838>
- [47] Y. CHI, S. NIJSSEN, R. R. MUNTZ, J. N. KOK. *Frequent Subtree Mining—An Overview*, in "Fundamenta Informaticae", IOS Press, 2005, vol. 66, pp. 161–198
- [48] A. CORNUÉJOLS, L. MICLET. , *Apprentissage artificiel : concepts et algorithmes*, Eyrolles, 2002
- [49] L. DE RAEDT. *A perspective on inductive databases*, in "SIGKDD Explor. Newsl.", 2002, vol. 4, n^o 2, pp. 69–77, <http://doi.acm.org/10.1145/772862.772871>
- [50] C. DOUSSON. , *Chronicle Recognition System*, 1994
- [51] C. DOUSSON. , *Suivi d'évolutions et reconnaissance de chroniques*, Université Paul Sabatier de ToulouseLAAS-CNRS, Toulouse, 1994
- [52] C. DOUSSON, T. V. DUONG. *Discovering Chronicles with Numerical Time Constraints from Alarm Logs for Monitoring Dynamic Systems*, in "IJCAI", T. DEAN (editor), Morgan Kaufmann, 1999, pp. 620-626
- [53] C. DOUSSON, P. GABORIT, M. GHALLAB. *Situation recognition: representation and algorithms*, in "Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)", Chambéry, France, 1993, pp. 166-172

- [54] S. DZEROSKI, L. TODOROVSKI. *Discovering dynamics: from inductive logic programming to machine discovery*, in "Journal of Intelligent Information Systems", 1995, vol. 4, pp. 89-108
- [55] J. GAO, B. DING, W. FAN, J. HAN, P. S. YU. *Classifying Data Streams with Skewed Class Distributions and Concept Drifts*, in "IEEE Internet Computing", 2008, vol. 12, n^o 6, pp. 37-49
- [56] M. GAROFALAKIS, J. GEHRKE, R. RASTOGI. *Querying and Mining Data Streams: You Only Get One Look. Tutorial notes*, in "ACM Int. Conf. on Management of Data", 2002
- [57] E. GIUNCHIGLIA, J. LEE, V. LIFSCHITZ, N. MCCAIN, H. TURNER. *Nonmonotonic causal theories*, in "Artificial Intelligence Journal", March 2004, vol. 153, n^o 1-2, pp. 49-104
- [58] T. GUYET, H. NICOLAS, A. DIOUCK. *Segmentation multi-échelle de séries temporelles d'images satellite : Application à l'étude d'une période de sécheresse au Sénégal*, in "Reconnaissance de Forme et Intelligence Artificielle (RFIA)", Lyon, France, January 2012, <http://hal.inria.fr/hal-00646158>
- [59] T. GUYET, R. QUINIOU. *Mining temporal patterns with quantitative intervals*, in "4th International Workshop on Mining Complex Data at ICDM 2008", December 2008
- [60] T. GUYET, R. QUINIOU. *Extraction incrémentale de séquences fréquentes dans un flux d'itemsets*, in "Extraction et Gestion de Connaissances (EGC'2012)", Bordeaux, France, B. PINAUD, G. MELANÇON, Y. LECHEVALLIER (editors), RNTI, Hermann, February 2012, <http://hal.inria.fr/hal-00648893>
- [61] T. GUYET, R. QUINIOU. *Incremental mining of frequent sequences from a window sliding over a stream of itemsets*, in "Journées Intelligence Artificielle Fondamentale", France, May 2012, pp. 153-162, <http://hal.inria.fr/hal-00757120>
- [62] D. T. HAU, E. W. COIERA. *Learning qualitative models of dynamic systems*, in "Machine Learning", 1997, vol. 26, pp. 177-211
- [63] T. IMIELINSKI, H. MANNILA. *A database perspective on knowledge discovery*, in "Comm. of The ACM", 1996, vol. 39, pp. 58-64
- [64] M. KUBAT, J. GAMA, P. E. UTGOFF. *Incremental learning and concept drift*, in "Intell. Data Anal.", 2004, vol. 8, n^o 3
- [65] C. LARGOUËT, M.-O. CORDIER, Y.-M. BOZEC, Y. ZHAO, G. FONTENELLE. *Use Of Timed Automata And Model-Checking To Explore Scenarios On Ecosystem Models*, in "Environmental Modelling and Software", November 2011, n^o 30, pp. 123-138, On-line publication : 26 November 2011 [DOI : 10.1016/J.ENVSFT.2011.08.005], <http://hal.inria.fr/hal-00649275>
- [66] E.-G. LAZRAC, M. BENOÎT, J.-F. MARI. *Time-Space Dependencies in Land-Use Successions at Agricultural Landscape Scales*, in "International Conference on Integrative Landscape Modelling", 2010, <http://hal.inria.fr/inria-00482890>
- [67] S. D. LEE, L. DE RAEDT. , *Constraint Based Mining of First Order Sequences in SeqLog*, LNCS, Springer-Verlag, 2004, vol. 2682, pp. 154-173

- [68] J. LIN, E. KEOGH, L. WEI, S. LONARDI. *Experiencing SAX: a Novel Symbolic Representation of Time Series*, in "Data Mining and Knowledge Discovery", 2007, vol. 15, n^o 2, pp. 107–144
- [69] H. MANNILA, H. TOIVONEN, A. I. VERKAMO. *Discovery of Frequent Episodes in Event Sequences*, in "Data Mining and Knowledge Discovery", 1997, vol. 1, n^o 3, pp. 259–289
- [70] A. MARASCU, F. MASSEGLIA. *Mining sequential patterns from data streams: a centroid approach*, in "J. Intell. Inf. Syst.", 2006, vol. 27, n^o 3, pp. 291–307, <http://dx.doi.org/10.1007/s10844-006-9954-6>
- [71] F. MASSEGLIA, F. CATHALA, P. PONCELET. *The PSP Approach for Mining Sequential Patterns*, in "Proceedings of the Second European Symposium on Principles of Data Mining and Knowledge Discovery", 1998, pp. 176–184
- [72] Y. MOINARD. *A formalism for causal explanations with an Answer Set Programming translation*, in "4th International Conference on Knowledge Science, Engineering & Management (KSEM 2010)", United Kingdom Belfast, B. SCOTNEY, Z. JIN (editors), LNCS, Springer-Verlag, Aug 2010, <http://hal.inria.fr/inria-00511093/en>
- [73] Y. MOINARD. *Using ASP with recent extensions for causal explanations*, in "ASPOCP10, Answer Set Programming and Other Computing Paradigms Workshop, associated with ICLP", United Kingdom Edinburgh, M. BALDUCCINI, S. WOLTRAN (editors), Jul 2010, <http://hal.inria.fr/inria-00542880/en>
- [74] D. PAGE. *ILP: Just Do It*, in "Proceedings of ILP'2000", J. CUSSENS, A. FRISCH (editors), LNAI, Springer, 2000, vol. 1866, pp. 3-18
- [75] Y. PENCOLÉ, M.-O. CORDIER, L. ROZÉ. *Incremental decentralized diagnosis approach for the supervision of a telecommunication network*, in "DX'01 (International Workshop on Principles of Diagnosis)", Sansicario, Italy, 2001, <http://www.irisa.fr/dream/dataFiles/ypencole/DX01.pdf>
- [76] R. QUINIOU, M.-O. CORDIER, G. CARRAULT, F. WANG. *Application of ILP to cardiac arrhythmia characterization for chronicle recognition*, in "ILP'2001", C. ROUVEIROL, M. SEBAG (editors), LNAI, Springer-Verlag, 2001, vol. 2157, pp. 220-227, <http://www.irisa.fr/dream/dataFiles/quiniou/ilp01.pdf>
- [77] N. RAMAUX, M. DOJAT, D. FONTAINE. *Temporal scenario recognition for intelligent patient monitoring*, in "Proc. of the 6th Conference on Artificial Intelligence in Medecine Europe (AIME'97)", 1997
- [78] R. REITER. *A theory of diagnosis from first principles*, in "Artificial Intelligence", 1987, vol. 32, n^o 1, pp. 57-96
- [79] M. SAMPATH, R. SENGUPTA, S. LAFORTUNE, K. SINNAMOHIDEEN, D. TENEKETZIS. *Diagnosability of discrete event systems*, in "Proceedings of the International Conference on Analysis and Optimization of Systems", 1995, vol. 40, pp. 1555-1575
- [80] M. SAMPATH, R. SENGUPTA, S. LAFORTUNE, K. SINNAMOHIDEEN, D. TENEKETZIS. *Active Diagnosis of Discrete-Event Systems*, in "IEEE Transactions on Automatic Control", 1998, vol. 43, n^o 7, pp. 908-929
- [81] M. SEBAG, C. ROUVEIROL. *Constraint Inductive Logic Programming*, in "Advances in Inductive Logic Programming", L. DE RAEDT (editor), IOS Press, 1996, pp. 277-294

-
- [82] K. SEDKI, L. BONNEAU DE BEAUFORT. *Cognitive Maps and Bayesian Networks for Knowledge Representation and Reasoning*, in "24th International Conference on Tools with Artificial Intelligence", Greece, 2012, pp. 1035-1040 [DOI : 10.1109/ICTAI.2012.175], <http://hal.inria.fr/hal-00757189>
- [83] P. SMETS, R. KENNES. *The Transferable Belief Model*, in "Classic Works of the Dempster-Shafer Theory of Belief Functions", R. R. YAGER, L. LIU (editors), Studies in Fuzziness and Soft Computing, Springer Berlin Heidelberg, 2008, vol. 219, pp. 693-736
- [84] THE WS-DIAMOND TEAM. 2, in "WS-DIAMOND: Web Services DIAGnosability, MONitoring and Diag-nosis", J. MYLOPOULOS, M. PAPAZOGLU (editors), MIT Press Series on Information Systems, 2009
- [85] A. VAUTIER, M.-O. CORDIER, R. QUINIOU. *An Inductive Database for Mining Temporal Patterns in Event Sequences (short version)*, in "Proceedings of IJCAI-05 (International Joint Conference on Artificial Intelligence)", Edinburgh, L. P. KAEHLING, A. SAFFIOTTI (editors), 2005, pp. 1640-1641, Poster
- [86] G. WIDMER, M. KUBAT. *Learning in the Presence of Concept Drift and Hidden Contexts*, in "Machine Learning", 1996, vol. 23, n^o 1, pp. 69-101
- [87] X. YAN, J. HAN. *gSpan: Graph-Based Substructure Pattern Mining*, in "Proceedings of the 2002 IEEE International Conference on Data Mining", Washington, DC, USA, ICDM '02, IEEE Computer Society, 2002, pp. 721–
- [88] Y. ZHAO, M.-O. CORDIER, C. LARGOUËT. *Répondre aux questions "Que faire pour" par synthèse de contrôleur sur des automates temporisés - Application à la gestion de la pêche*, in "RFIA 2012 (Reconnaissance des Formes et Intelligence Artificielle)", Lyon, France, January 2012, pp. 978-2-9539515-2-3, Session "Posters", <http://hal.inria.fr/hal-00656543>
- [89] J. DE KLEER, A. MACKWORTH, R. REITER. *Characterizing diagnoses and systems*, in "Artificial Intelligence", 1992, vol. 56, n^o 2-3, pp. 197-222