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

Activity Report 2012

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
Knowledge and Data Representation and Management

Table of contents

1. Members 1
2. Overall Objectives 1
2.1. Introduction 1
2.2. Application domains 2
3. Scientific Foundations 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 4
3.2.3. Mining data streams 5
4. Application Domains 5
4.1. Introduction 5
4.2. Software components monitoring 5
4.3. Environmental decision making 6
5. Software 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 7
5.5. ManageYourself 8
6. New Results 8
6.1. Diagnosis of large scale discrete event systems 8
6.1.1. Distributed monitoring with chronicles - Interleaving diagnosis and repair - Making web services more adaptive 9
6.1.2. Scenario patterns for exploring qualitative ecosystems 9
6.1.3. Controler synthesis for dealing with “How to” queries 9
6.2. Machine learning for model acquisition 9
6.2.1. Mining temporal patterns with numerical information 10
6.2.2. Incremental sequential mining 10
6.2.3. Incremental learning of preventive rules 10
6.2.4. Multiscale segmentation of satellite image time series 10
6.2.5. Mining a big unique graph for spatial pattern extraction 10
6.3. Decision aiding with models and simulation data 11
6.3.1. Exploring models thanks to scenarios: a generic framework 11
6.3.2. A datawarehouse for simulation data 11
6.3.3. Efficient computation of skyline queries in an interactive context 11
6.3.4. Influence Diagrams for Multi-Criteria Decision 12
6.3.5. Modeling influence propagation by Bayesian causal maps 12
6.3.6. Mining simulation data by rule induction 12
6.4. Diagnostic and causal reasoning 12
7. Bilateral Contracts and Grants with Industry 13
7.1. Bilateral Grants with Industry 13
7.2. National Initiatives 13
7.2.1. SACADEAU-APPEAU: Decision-aid to improve streamwater quality 13
7.2.2. ACASSYA: Supporting the agro ecological evolution of breeding systems in coastal watersheds 14
7.2.3. PayOTe-Network: characterizing agricultural landscapes via data mining 14
8. Dissemination 14

8.1. Scientific Animation 14
8.1.1. Journal editorial board 14
8.1.2. Conference program committees and organizations 14
8.1.3. Scientific and administrative boards 15
8.1.4. Misc 15
8.2. Teaching - Supervision - Juries 15
8.2.1. Teaching 15
8.2.2. Supervision 16
8.2.3. Juries 16
9. Bibliography **16**

Project-Team DREAM

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

Creation of the Project-Team: October 04, 2004 .

1. Members

Research Scientists
Yves Moinard [CR Inria]
René Quiniou [CR Inria]

Faculty Members
Marie-Odile Cordier [Team Leader, Professor, University of Rennes 1, HdR]
Véronique Masson [Associate professor, University of Rennes 1 (80 % time)]
Laurence Rozé [Associate professor, Insa, Rennes]

External Collaborators
Philippe Besnard [Research scientist, DR CNRS, IRIT, Toulouse]
Louis Bonneau de Beaufort [Research Engineer, Agrocampus-Ouest, Rennes]
Thomas Guyet [Associate Professor, Agrocampus-Ouest, Rennes]
Simon Malinowski [Associate Professor, Agrocampus-Ouest, Rennes (since september 1st 2012)]
Karima Sedki [Associate Professor, Agrocampus-Ouest, Rennes (till august 31st 2012)]

Engineer
Florimond Ployette [IR Inria (till august 31st 2012]

PhD Students
Tassadit Bouadi [University of Rennes1]
Yulong Zhao [University of Rennes1]

Post-Doctoral Fellow
Sid Ahmed Benabderrahmane [since 26th november 2012]

Administrative Assistant
Marie-Noëlle Georgeault

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 streams 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 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 efficiency of the computation 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 the diagnosability (i.e. whether, given the system specifications, it is possible to detect and explain an error in due time), or the repairability (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, and 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.

3. Scientific Foundations

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 [63], [74]. 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 [21], [23] 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 ,...) [28], [27].

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 [65]. 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 [60].
- 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 [43], [62], [41].

  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 [64]), 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 -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 or artificial learning [38]. 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 data representation 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 [59], 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 [66]. Extensions, such as QSIM [49], have also been used for learning a model of the behavior of a dynamic system [44]. 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 [61] 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 [37] 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 [25]. 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" [26], which extends Apriori by coping with the order of elements in patterns.

In [54], Mannila and Toivonen extended the work of Aggrawal et al. by introducing an algorithm for mining patterns involving temporal episodes with a distinction between parallel and sequential event patterns. Later, in [42], 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 [40] 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 [22] 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 [50], [39]. 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 [53]. We have investigated an extension of inductive database where patterns are very close to chronicles [69].

3.2.3. Mining data streams

) During the last years, a new challenge has appeared in the data mining community: mining from data streams [24]. 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” [46]. 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 or summaries have been proposed, as well as representation models taking advantage of the most recent data. Sequential pattern stream mining is still an issue [55]. 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* [71], [51], 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 [45] 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 more and more 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 [68], 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 Développpement 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

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 [48] and [3].

The software is mainly implemented in Matlab. A standalone application is now avalaible. It uses the Mixmod toolbox [35] 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 presented in section 7.2.1. 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 (see section 6.3.1). 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/.

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. 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 ManagerYourself 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.

6. New Results

6.1. Diagnosis of large scale discrete event systems

Participants: Marie-Odile Cordier, 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 and worked on a journal paper that will be submitted in 2013.

6.1.2. Scenario patterns for exploring qualitative ecosystems

This work aims at giving means of exploring complex systems, in our case ecosystems. 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 discret 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 formula. The answer is computed thanks to model-checking techniques that are efficient for analysing 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 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 accepted for publication in the Environmental Modelling Software Journal [52].

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

We extended the approach to deal with “How to” queries. As before, we rely on a qualitative model in the form of timed automata and on model-checking tools to answer queries. We proposed and compared two approaches to answer questions such as “How to avoid a given situation ?”(safety query). The first one exploits controller synthesis and the second one is a “generate and test” approach. We evaluated these two approaches in the context of an application that motivates this work, i.e the management of a marine ecosystem and the evaluation of fishery management policies. The results have been accepted for publication in [17].

More recently, we use similar methodological tools to model herd management on a catchment and analyse the best/optimal farming practices in order to reduce nitrate pollution due to livestock effluents. An hybrid model has been built using hierarchical timed automata. Scenarios can already be simulated and evaluated. We currently work on adapting controler synthesis tools in order to get the best strategies. This work is made in collaboration with our colleagues of INRA.

6.2. Machine learning for model acquisition

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

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. Mining temporal patterns with numerical information

We are interested in mining interval-based temporal patterns from event sequences where each event is associated with a type and time interval. Temporal patterns are sets of constrained interval-based events. This year we have begun to work on multiscale temporal abstraction to represent time series by codewords at different temporal and amplitude scales. We have improved the method of Wang et al. [70] by introducing Dynamic Time Warping to compute better codewords for time series abstraction. The codeword-based time series representation is then used by QTIPrefixSpan [3] to extract temporal patterns. A paper is in preparation. We are also working on a multivariate version of the method for mining multivariate temporal patterns at different resolution levels.

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. Though being of great importance, this problem has not received a similar attention as mining from a transaction database. Our method [13] provides an algorithm that maintains a tree representation (inspired by the PSP algorithm [56]) of frequent sequential patterns and their minimal occurrences [54] 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 [14].

6.2.3. Incremental learning of preventive rules

The problem is to learn preventive rules in order to avoid malfunctioning on smartphones. A monitoring module is embedded on the phones and sends reports to a server. Reports are labeled with a normal or abnormal label. From this set of reports new rules are learned. As a lot of smartphones are supervised, it is impossible to store all the reports. Therefore incremental learning has to be used.

Last year, we achieved two main tasks: a report database has been built in order to test the future algorithms, and a new algorithm [20] has been developed for implementing an incremental version of the algorithm AQ21 [72].

6.2.4. 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 [12] , we present a method to segment an image through the characterization of the evolution of a vegetation index (NDVI) on two scales: annual and multi-year. We test this method to segment Senegal SITS and compare our method to a direct classification of time series. The results show that our method using two time scales better differentiates regions in the median zone of Senegal and locates fine interesting areas (cities, forests, agricultural areas).

6.2.5. Mining a big unique graph for spatial pattern extraction

Researchers in agro-environment needs a great variety of landscapes to test the agro-ecological models of their scientific hypotheses. As the representation of real landcapes necessitates lots of on-land measures, good big representations are difficult to acquire. Working with landscape simulations is then an alternative to get a sufficient variety of experimental landscapes. We propose to extract spatial patterns from a well described geographic area and to use these patterns to generate realistic landscapes. We have begun the exploration of graph mining techniques to discover the relevant spatial patterns present in a graph expressing the spatial relationships between the agricultural plots as well as the roads, the rivers, the buildings, etc., of a specific geographic area.

This year, we have been working on extending algorithm gSPAN [73] with an adaptive support threshold and with a taxonomy to be able to extract interesting patterns involving agricultural plots with rare features. We plan to submit a paper in 2013.

6.3. Decision aiding with models and simulation data

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

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 very 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. Exploring models thanks to scenarios: a generic framework

In the framework of the Appeau project (see 7.2.1) a paper describing a generic framework for scenario exercises using models applied to water-resource management, has been written in cooperation with all the partners and published in Environmental Modelling and Software [5].

6.3.2. A datawarehouse for simulation data

The ACASSYA project 7.2.2 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 will be submitted to the COMPAG Journal in beginning 2013.

6.3.3. Efficient computation of skyline queries in an interactive context

Skyline queries retrieve from a database the objects that maximizes 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.

This year we have worked at improving the formal setting of the partial materialization of skyline queries when dynamic preferences are refined online by the user. We have explicited which parts of the skyline evolve (which point are added or removed) when a new dimension is introduced in the computation. This led to an efficient incremental method for the online computation of the skyline corresponding to new user preferences [9]. An extended version of this paper is under submission to the Journal "Transactions on Large Scale Data and Knowledge Centered Systems" (TLDKS).

We are working now on a hierachical extension of our method that could be introduced in a datawarehouse context.

6.3.4. Influence Diagrams for Multi-Criteria Decision

For multi-criteria decision-making problems, we propose in [7] a model based on influence diagrams able to handle uncertainty, represent interdependencies among the different decision variables and facilitate communication between the decision-maker and the analyst. The model makes it possible to take into account the alternatives described by an attribute set, the decision-maker's characteristics and preferences, and other information (e.g., internal or external factors) that influence the decision. Modeling the decision problem in terms of influence diagrams requires a lot of work to gather expert knowledge. However, once the model is built, it can be easily and efficiently used for different instances of the decision problem. In fact, using our model simply requires entering some basic information, such as the values of internal or external factors and the decision-maker's characteristics.

6.3.5. 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 [11].

However, this model does not take into account the variety of influences asserted by the fishermen, but only the "mean" causal map. This year we have been working on an approach that could combine individual knowledge with belief functions in the way of Philippe Smets's Transferable Belief Model [67].

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 7.2 because DREAM is not an official partner of this project.).

6.3.6. Mining simulation data by rule induction

In the framework of the SACADEAU project (see 7.2.1), a paper dedicated to mining simulation data by rule induction has been published in the COMPAG Journal [8]. Both qualitative and quantitative predictions from a model of an agro-environmental system are analysed. Two approaches in rule learning from spatial data (ILP and attribute-value approaches) are compared and show that results help identify factors with strong influence on herbicide stream-pollution.

We have also participated in a collaboration for modeling the effects crop rotations the results of which were published in the Science of the Total Environment Journal [6].

6.4. Diagnostic and causal reasoning

Participants: Philippe Besnard, Louis Bonneau de Beaufort, Marie-Odile Cordier, Yves Moinard, Karima Sedki.

Stemming on [29], [30], [31], [32], [33], 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 [36], [47], they were too close to standard logic in order 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 also been translated into answer set programming [57], [58]).

This year, we have extended our formalism in order to deal with more 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. We are on the way to end this task. 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 [34] 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.

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 answer set programming, we have also examined some more difficult examples [16] and participated to a chapter in the to be published "Panorama de l'intelligence artificielle. Ses bases méthodologiques, ses développements" [19].

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Grants 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 using decision trees, the aim being to characterize the faults and consequently update the global knowledge base and its distributed instances. This year, we studied an incremental version of this learning step in order to get an on-line process [20]. This means being able to learn new faults characterizations and add new preventive rules, and also forget no longer needed ones.

7.2. National Initiatives

7.2.1. SACADEAU-APPEAU: Decision-aid to improve streamwater quality

Participants: Marie-Odile Cordier, Véronique Masson.

The SACADEAU project (Système d'Acquisition de Connaissances pour l'Aide à la Décision pour la qualité de l'EAU - Knowledge Acquisition System for Decision-Aid to Improve Streamwater Quality) was funded by INRA (French institute for agronomy research) from October 2002 to October 2005. The main partners were from INRA (SAS from Rennes and BIA from Toulouse) and from IRISA.

We were then involved in a new project, named APPEAU and funded by ANR/ADD, which started in February 2007 and ended in december 2011. The APPEAU project aimed at studying which politics, for which agronomic systems, are best adapted to improve water management. It includes our previous partners as well as new ones, mainly from INRA(http://wwwagir.toulouse.inra.fr/agir/index.php?option=com_content&view=article&id=62&Itemid=134). A synthesis paper has been published in 2012 [5].

7.2.2. 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://www.rennes.inra.fr/umrsas/programmes/acassya_accompagner_l_evolution_agro_ecologique_des_systemes_d_elevage)

7.2.3. PayOTe-Network: characterizing agricultural landscapes via data mining

Participants: Thomas Guyet, René Quiniou.

The PAYOTE project (Paysage Ou Territoire) was initialy funded by AIP INRA/INRIA. It started at the end of 2010 and will end by the end of 2012.. The project is turning into a network mainly funded by INRA. This network still associates Inria Teams (Orpailleur and Dream) with INRA Team (UBIA, MIAJ and SAD-Paysage).

One of the objectives of the PAYOTE network is to provide tools to generate "realistic" agricultural landscapes. This kind of generator are expected by expert to study the impact of the landscape on agro-ecological systems. The main approach of this network is to use data mining to automatically construct a neutral model of a landscape. Then, the model of a landscape may be used to generate new landscapes with same spatial properties.

In this context, the challenge is to develop spatio-temporal data mining algorithms to analyse the spatial organization of agricultural landscapes.

8. Dissemination

8.1. Scientific Animation

8.1.1. Journal editorial board

- *AAI: Applied Artificial Intelligence* (M.-O. Cordier).
- *ARIMA: Revue Africaine de la Recherche en Informatique et en Mathématiques Appliquées* (M.-O. Cordier).
- *Revue I3* (M.-O. Cordier).
- *Revue JESA (Journal Européen des Systèmes automatisés)* (M.-O. Cordier).
- *Interstices webzine* (M.-O. Cordier).
- *Bulletin AFIA* Invited editors of the special issue on "Intelligence Artificial and Agronomy": T. Guyet, M.-O. Cordier

8.1.2. Conference program committees and organizations

- Program committee member of KR'2012, ECAI 2012, RFIA'2012 (M.-O. Cordier)
- Program committee member of RFIA'2012 (T. Guyet).
- Program committee member of ECAI 2012, EGC 2013 (R. Quiniou).
- Program committee members of FOSTA workshop at EGC 2013 (T. Guyet, R. Quiniou).
- Organization committee member of CIDN workshop at EGC 2013 (T. Guyet).

- Organization committee members of EGC 2014 scheduled in Rennes (T. Guyet, R. Quiniou).
- Steering Committee of RFIA'2014 (T. Guyet).

8.1.3. Scientific and administrative boards

- ECCAI fellow + Honorific 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", "Comité scientifique de l'ISTIC", "Direction scientifique de l'Irisa": M.-O. Cordier
- Member of the "AAAI" award committee: M.-O. Cordier.
- Member of the "Prix de thèse AFIA 2012" 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: T. Guyet.
- Member of the COREGE (Research Committee-COmité de la REcherche du Grand Etablissement) of Agrocampus-Ouest: T. Guyet.
- External evaluator for the portuguese FCT (Portuguese Foundation For Science and Technology): T. Guyet.
- Member of the Payote-Network board: T. Guyet.

8.1.4. Misc

- Expert for specifying "methods and tools for exploring and analyzing simulation results" in the INRA inter-department project "Paysage Virtuel" (Virtual Landscape) (T. Guyet).
- Organization of the "Turing day" in october 2012 (M.-O. Cordier)
- Invited talk at CIDN workshop at EGC 2012 (R. Quiniou).
- Invited talk at Assises du GDR I3, Porquerolles, 2012, May 10 (Y. Moinard).
- Seminar in the Reasoning and Decision theme at IRIT (Toulouse, 2012, june 15) (Y. Moinard).
- Invited lectures at the "Logic Summer School", Canberra, Australie, december 2012 (M.-O. Cordier)

8.2. Teaching - Supervision - Juries

8.2.1. 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: *Module OCI : Interfaces graphiques en C++/GTKMM*, M2, ISTIC University of Rennes (T. Guyet),

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

8.2.2. Supervision

PhD in progress: 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”, january 1st 2010, co-supervisors Marie-Odile Cordier, René Quiniou and Chantal Gascuel, ANR project Acassya grant

PhD in progress: 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.”, october 1st 2010, 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

8.2.3. Juries

- Commitee member of Romain Tavenard’s PhD defence (Univ. Rennes1): M.-O. Cordier
- Commitee member of Nuno Belard’s PhD defence (Univ. Toulouse): M.-O. Cordier
- Commitee member of Annie Foret’s HDR defence (Univ. Rennes1): M.-O. Cordier
- Commitee member of Vincent Armant’s PhD defence (Univ. Paris11): M.-O. Cordier
- Committee member of Christophe Salperwyck’s PhD defence (Université de Lille): R. Quiniou.
- External member of the INRA jury for the recruitment of computer science engineers: T. Guyet

9. 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, p. 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, p. 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, p. 121-170, http://hal.inria.fr/inria-00511104/en/.

Publications of the year

Articles in International Peer-Reviewed Journals

[5] D. LEENHARDT, O. THEROND, M.-O. CORDIER, C. GASCUEL-ODOUX, A. REYNAUD, P. DURAND, J.-E. BERGEZ, L. CLAVEL, V. MASSON, P. MOREAU. *A generic framework for scenario exercises using models applied to water-resource management*, in "Environmental Modelling and Software", 2012, vol. 37, p. 125-133 [*DOI :* 10.1016/J.ENVSOFT.2012.03.010], http://hal.inria.fr/hal-00767926.

[6] P. MOREAU, L. RUIZ, T. RAIMBAULT, F. VERTÈS, M.-O. CORDIER, C. GASCUEL-ODOUX, V. MASSON, J. SALMON-MONVIOLA, P. DURAND. *Modeling the potential benefits of catch-crop introduction in fodder crop rotations in a Western Europe landscape*, in "Science of the Total Environment", 2012, vol. 437, p. 276-284 [*DOI :* 10.1016/J.SCITOTENV.2012.07.091], http://hal.inria.fr/hal-00767910.

[7] K. SEDKI, V. DELCROIX. *A model based on influence diagrams for multi-criteria decision making*, in "International Journal on Artificial Intelligence Tools", August 2012, vol. 21, n° 4, 1250018 [*DOI :* 10.1142/S0218213012500182], http://hal.inria.fr/hal-00757174.

[8] R. TRÉPOS, V. MASSON, M.-O. CORDIER, C. GASCUEL-ODOUX, J. SALMON-MONVIOLA. *Mining simulation data by rule induction to determine critical source areas of stream water pollution by herbicides*, in "Computers and Electronics in Agriculture", 2012, vol. 86, p. 75-88 [*DOI :* 10.1016/J.COMPAG.2012.01.006], http://hal.inria.fr/hal-00767865.

International Conferences with Proceedings

[9] 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, p. 219-233 [*DOI :* 10.1007/978-3-642-32600-4_17], http://hal.inria.fr/hal-00757838.

[10] P. RANNOU, F. LAMARCHE, M.-O. CORDIER. *Enhancing the behavior of virtual characters with long term planning, failure anticipation and opportunism*, in "Motion In Games", Rennes, France, Springer, November 2012, http://hal.inria.fr/hal-00763694.

[11] 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, p. 1035-1040 [*DOI :* 10.1109/ICTAI.2012.175], http://hal.inria.fr/hal-00757189.

National Conferences with Proceeding

[12] 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.

[13] 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.

[14] 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, p. 153-162, http://hal.inria.fr/hal-00757120.

[15] V. MASSON, F. PLOYETTE. *Logiciel SACADEAU - Outil d'aide à la gestion de la qualité des eaux d'un bassin versant: simulation, apprentissage de règles de caractérisation et recommandation d'actions*, in "RFIA 2012 (Reconnaissance des Formes et Intelligence Artificielle)", Lyon, France, January 2012, Session "Démo", 978-2-9539515-2-3, http://hal.inria.fr/hal-00660959.

[16] Y. MOINARD. *Utiliser la programmation par ensembles réponses pour de petits problèmes*, in "RFIA 2012 (Reconnaissance des Formes et Intelligence Artificielle)", Lyon, France, January 2012, Session "Posters", 978-2-9539515-2-3, http://hal.inria.fr/hal-00656561.

[17] 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, Session "Posters", 978-2-9539515-2-3, http://hal.inria.fr/hal-00656543.

Scientific Books (or Scientific Book chapters)

[18] 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.

[19] Y. MOINARD, A. LALLOUET, 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.

Other Publications

[20] 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.

References in notes

[21] B. DUBUISSON (editor). *Diagnostic, intelligence artificielle et reconnaissance des formes*, Traité IC2 : Information - Commande - Communication, Hermes, 2001.

[22] S. DZEROSKI, N. LAVRAČ (editors). *Relational Data Mining*, Springer, Berlin, 2001.

[23] W. HAMSCHER, L. CONSOLE, J. DE KLEER (editors). *Readings in Model-Based Diagnosis*, Morgan Kaufmann, San Meteo, CA, Etats-Unis, 1992.

[24] C. AGGARWAL. *Data Streams: Models and Algorithms*, Advances in Database Systems, Springer, 2007.

[25] 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, p. 207–216.

[26] 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, p. 3–14.

[27] G. ALARME. *Monitoring and alarm interpretation in industrial environments*, in "AI Communications", 1998, vol. 11, 3-4, p. 139-173, S. Cauvin, Marie-Odile Cordier, Christophe Dousson, P. Laborie, F. Lévy, J. Montmain, M. Porcheron, I. Servet, L. Travé-Massuyès.

[28] 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.

[29] 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, p. 55-67.

[30] 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, p. 292–304, http://www.irisa.fr/dream/dataFiles/moinard/causeksempub.pdf.

[31] 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, p. 153-164, http://www.irisa.fr/dream/dataFiles/moinard/ksem07bcausesonto.pdf.

[32] 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, p. 723–724.

[33] P. BESNARD, M.-O. CORDIER, Y. MOINARD. *Ontology-based inference for causal explanation*, in "Integrated Computer-Aided Engineering", 2008, vol. 15, n^{o} 4, p. 351-367.

[34] P. BESNARD, A. HUNTER. *Elements of Argumentation*, The MIT Press, http://www-mitpress.mit.edu/, june 2008.

[35] 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, p. 587-600.

[36] 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, p. 33–38.

[37] Y. CHI, S. NIJSSEN, R. R. MUNTZ, J. N. KOK. *Frequent Subtree Mining–An Overview*, in "Fundamenta Informaticae", IOS Press, 2005, vol. 66, p. 161–198.

[38] A. CORNUÉJOLS, L. MICLET. *Apprentissage artificiel : concepts et algorithmes*, Eyrolles, 2002.

[39] L. DE RAEDT. *A perspective on inductive databases*, in "SIGKDD Explor. Newsl.", 2002, vol. 4, n^{o} 2, p. 69–77, http://doi.acm.org/10.1145/772862.772871.

[40] C. DOUSSON. *Chronicle Recognition System*, 1994.

[41] C. DOUSSON. *Suivi d'évolutions et reconnaissance de chroniques*, Université Paul Sabatier de Toulouse, LAAS-CNRS, Toulouse, 1994.

[42] 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, p. 620-626.

[43] 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, p. 166-172.

[44] S. DZEROSKI, L. TODOROVSKI. *Discovering dynamics: from inductive logic programming to machine discovery*, in "Journal of Intelligent Information Systems", 1995, vol. 4, p. 89-108.

[45] 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, p. 37-49.

[46] 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.

[47] 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, p. 49–104.

[48] T. GUYET, R. QUINIOU. *Mining temporal patterns with quantitative intervals*, in "4th International Workshop on Mining Complex Data at ICDM 2008", December 2008.

[49] D. T. HAU, E. W. COIERA. *Learning qualitative models of dynamic systems*, in "Machine Learning", 1997, vol. 26, p. 177-211.

[50] T. IMIELINSKI, H. MANNILA. *A database perspective on knowledge discovery*, in "Comm. of The ACM", 1996, vol. 39, p. 58–64.

[51] M. KUBAT, J. GAMA, P. E. UTGOFF. *Incremental learning and concept drift*, in "Intell. Data Anal.", 2004, vol. 8, n^o 3.

[52] 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, p. 123-138, On-line publication : 26 November 2011 [*DOI :* 10.1016/J.ENVSOFT.2011.08.005], http://hal.inria.fr/hal-00649275.

[53] S. D. LEE, L. DE RAEDT. *Constraint Based Mining of First Order Sequences in SeqLog*, LNCS, Springer-Verlag, 2004, vol. 2682, p. 154-173.

[54] 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, p. 259–289.

[55] A. MARASCU, F. MASSEGLIA. *Mining sequential patterns from data streams: a centroid approach*, in "J. Intell. Inf. Syst.", 2006, vol. 27, n^o 3, p. 291–307, http://dx.doi.org/10.1007/s10844-006-9954-6.

[56] 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, p. 176–184.

[57] 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.

[58] 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.

[59] D. PAGE. *ILP: Just Do It*, in "Proceedings of ILP'2000", J. CUSSENS, A. FRISCH (editors), LNAI, Springer, 2000, vol. 1866, p. 3-18.

[60] 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.

[61] 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, p. 220-227, http://www.irisa.fr/dream/dataFiles/quiniou/ilp01.pdf.

[62] 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.

[63] R. REITER. *A theory of diagnosis from first principles*, in "Artificial Intelligence", 1987, vol. 32, n^o 1, p. 57-96.

[64] 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, p. 1555-1575.

[65] 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, p. 908-929.

[66] M. SEBAG, C. ROUVEIROL. *Constraint Inductive Logic Programming*, in "Advances in Inductive Logic Programming", L. DE RAEDT (editor), IOS Press, 1996, p. 277-294.

[67] 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, p. 693-736.

[68] THE WS-DIAMOND TEAM. *2*, in "WS-DIAMOND: Web Services DIAgnosability, MONitoring and Diagnosis", J. MYLOPOULOS, M. PAPAZOGLOU (editors), MIT Press Series on Information Systems, 2009.

[69] 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. KAELBLING, A. SAFFIOTTI (editors), 2005, p. 1640-1641, Poster.

[70] Q. WANG, V. MEGALOOIKONOMOU, C. FALOUTSOS. *Time series analysis with multiple resolutions*, in "Inf. Syst.", 2010, vol. 35, n^{o} 1, p. 56-74.

[71] G. WIDMER, M. KUBAT. *Learning in the Presence of Concept Drift and Hidden Contexts*, in "Machine Learning", 1996, vol. 23, n^{o} 1, p. 69-101.

[72] J. WOJTUSIAK, R. MICHALSKI, T. SIMANIVANH, A. BARANOVA. *The Natural Induction System AQ21 and its Application to Data Describing Patients with Metabolic Syndrome: Initial Results*, in "Machine Learning and Applications, 2007. ICMLA 2007. Sixth International Conference on", dec. 2007, p. 518 -523, http://dx.doi.org/10.1109/ICMLA.2007.98.

[73] 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, p. 721–.

[74] J. DE KLEER, A. MACKWORTH, R. REITER. *Characterizing diagnoses and systems*, in "Artificial Intelligence", 1992, vol. 56, n^{o} 2-3, p. 197-222.