INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team DREAM

Diagnosis, REcommending Actions and Modeling

Rennes

THEME COG

Activity Report

2005

Table of contents

1. Team

Head of project
Marie-Odile Cordier [Professor, University of Rennes 1]

Administrative assistant
Marie-Noëlle Georgeault [AA Inria]

Inria staff members
Yves Moinard [Research scientist]
René Quiniou [Research scientist]
Ansaf Salleb [Postdoc since september 1st 2004]

Faculty members
Véronique Masson [Associate professor, University of Rennes 1 (50 % time)]
Sophie Robin [Associate professor, University of Rennes 1]
Laurence Rozé [Associate professor, Insa, Rennes]
Pascal Garcia [Associate professor, Insa, Rennes]

Ph. D. candidates
Élisa Fromont [INRIA Ph. D. candidate, since October 2002]
Alban Grastien [MENRT Ph. D. candidate, since October 2002]
François Portet [MENRT Ph. D. candidate , since October 2002]
Ronan Trépos [Inra Ph. D. candidate, since October 2004]
Alexandre Vautier [MENRT Ph. D. candidate, since October 2004]
Xavier LeGuillou [University Ph. D. candidate, since October 2005]

Research scientists (partners)
Philippe Besnard [Research scientist, DR CNRS, IRIT, Toulouse]
Christine Largouët [Associate Professor, Ensar, Rennes (on leave at the University of New-Caledonia since June 2003)]

2. Overall Objectives

2.1. Overall Objectives

Keywords: *diagnosis*, *machine learning*, *supervision*.

The research objectives of the Dream team are about aiding monitoring and diagnosing time evolving systems. The main issue is to help the person in charge of the system by analyzing the observations provided by sensors and giving her/him information about diagnosis hypotheses (potential anomalies or failures) and recommended actions. Qualitative model-based approaches are advocated for at least two main reasons:

- they are "glass-box" approaches and consequently diagnoses and recommended actions can be explained to the user in an explicit and adequate language,
- they are flexible enough and are then adapted to quickly evolving systems such as technological systems (for instance telecommunication components).

We use a model-based approach relying on normal and faulty behavioral models. These models are discrete-event models such as (temporal) communicating automata, temporal causal graphs or chronicles.

In this context, two main research themes are developed:

1. Classical model-based diagnosis methodologies cannot be directly used for complex systems due to the intractable size of the model and the computational complexity of the process. It is especially true when on-line diagnosis is considered. Two solutions are investigated:

  - We propose a decentralized approach which relies on combining local diagnoses built from local models (or local diagnosers). Three problems are currently investigated: Which strategy should be used for an optimal merge of the local diagnoses in order to preserve the efficiency and the completeness of the process? How the process incrementality can be ensured in an on-line diagnosis context where observations are incrementally collected? How to deal with reconfigurable systems the topology of which can be changed at running time.
  - We propose to use model-checking techniques in order to improve the efficiency of the computation and to cut down the combinatorial complexity of the state space. It means using adequate symbolic representations such as BDD for instance and partial order reduction techniques taking advantage of the existing inherent concurrency.

2. It is well recognized that model-based approaches suffer from the difficulty to acquire the model. It is why we focus on automatically acquiring models from data with symbolic learning methods coupled with data mining methods. One of the challenges we tackle with is to extend existing inductive logic programming methods (ILP) to temporal data in order to be able to deal with data coming from signals (as electrocardiograms in the medical domain) or alarm logs (in the telecommunication domain). Two problems are currently investigated: how to adapt the learning process to deal with multiple sources of information (multi-sensor learning)? How to integrate signal processing algorithms to the learning or diagnosis task when this latter relies on a qualitative description of signals?

Our application domains are the following:

- large industrial systems monitoring applications with a focus on telecommunication networks;
- medical applications and especially cardiac monitoring, i.e the on-line analysis of cardiac signals to detect arrhythmias and the development of "intelligent" cardiac devices (pacemakers and defibrillators) having some signal analysis and diagnosis capabilities;
- environmental protection, and more precisely the development of decision support systems to help the management of agricultural plots with the objective of preserving water quality threatened by pesticide pollution.

3. Scientific Foundations

3.1. Physical systems monitoring aid

Keywords: *chronicle acquisition*, *chronicle recognition*, *deep model*, *diagnosis*, *fault model*, *monitoring*, *simulation*, *temporal causal graph*.

Our work on monitoring and diagnosis relies on model-based approaches developed by the Artificial Intelligence community since the founding studies by R. Reiter and J. de Kleer [75], [81]. Our project investigates the on-line monitoring and diagnosis of systems, which are modeled as discrete events systems, focusing more precisely on monitoring by alarms management [63]. Computational efficiency is a crucial issue for real size problems. We are developing two approaches. The first one relies on diagnosers techniques [79], for which we have proposed a decentralized and generic approach. The second one uses chronicle recognition techniques, focusing on learning chronicles.

Early work on model-based diagnosis dates back in the 70-80's by R. Reiter, the reference paper on the logical theory of diagnosis being [75], [81]. In the same years was constituted the community known as DX, named after the *workshop* on the principles of diagnosis. Research in these areas is still very active and the DX workshop gathers about fifty people in the field every year. Contrary to the expert system approach, which has been the leading approach for diagnosis (medical diagnosis for instance) before 1990, the model-based approach lies on a deep model representing the expected correct behavior of the system to be supervised or on a fault model. Instead of acquiring and representing an expertise from experts, the model-based approach uses the design models of industrial systems. The approach has been initially developed for electronic circuits repair [82], focusing on off-line diagnosis of so-called static systems. Two main approaches have been proposed then: (i) the consistency-based approach, relying on a model of the expected correct behavior, which aims at detecting the components responsible for a difference between the expected observations and the really observed ones; (ii) the abductive approach which relies on a model of the failures that can affect the system, and which identifies the failures or the faulty behavior explaining the anomalous observations. See the references [44], [46] for a detailed exposition of these investigations.

Since 1990, the researchers in the field have studied the monitoring and the diagnosis of dynamic systems, which made them closer to the researchers in control theory. What characterizes the IA approach is the use of qualitative models instead of quantitative ones and the importance given to the search for the real origin 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. The used formalisms are often similar (automata, Petri nets ,...) [51], [63].

Our team focuses on monitoring and on-line diagnosis of discrete events systems and in particular on monitoring by alarm management. In this context, a human operator is generally in charge of the system monitoring and receives events (the alarms) which are time-stamped and emitted by the components themselves, in reaction to external events. These observations on the system are discrete informations, corresponding to an instantaneous event or to a property associated to a time interval. The main difficulties for analyzing this flow of alarms are the following:

- the huge number of received alarms: the supervisor may receive till several hundreds of messages per second, many of which being insignificant,
- the alarm overlapping: the order in which alarms are received may be different from the order in which alarms were emitted. Moreover, various sequences of alarms resulting from concurrent failures may overlap. The propagating delays, and sometimes the ways the alarms are transmitted, must be taken into account, not only for event reordering, but also to decide at what time all the useful messages can be considered as being received.
- the redundancy of received alarms: some alarms are only routine consequence of other alarms. This can provoke a phenomenon known as cascading alarms.
- the alarm loss or alarm masking: some alarms can be lost or masked to the supervisor when an intermediate component in charge of the transmission is faulty. The absence of an alarm must be taken into account, since it can give a useful information about the state of the system.

There are two cases focusing on very different issues. In the first one, the alarms must be dealt with, *on-line*, by the operator. In this case, alarm analysis must be done in real time. The operator must react in a very short period of time to keep the system working at the best in spite of the inputs variability and the natural evolution of the processes. Consequently, the natural system damages (components wear, slow modification of the components properties, etc.) are not directly taken into account but are corrected by tuning some parameters.

This *reactive* treatment withstands the treatment of alarms maintenance. In this second case, a deeper *off line* analysis of the system is performed, by foreseeing the possible difficulties, by planning the maintenance operations in order to minimize significantly the failures and interruptions of the system.

The major part of our work focuses on on-line monitoring aid and it is assumed that the correct behavior model or the fault models of the supervised systems are available. However, an on-line use of the models is rarely possible because of its complexity with respect to real time constraints. This is especially true when temporal models are under concern. A way to tackle this problem is to make an off-line transformation (or compilation) of the models and to extract, in an adapted way, the useful elements for diagnosis.

We study two different methods:

- 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*. The transitions of the automaton are only triggered by observable events and the states contain only information on the failures that happened in the system. Diagnosing the system consists in going through all the different states of the diagnoser as observable events become available. This method has been proposed by M. Sampath and colleagues [79]. We have extended this method to the communicating automata formalism [78] (see also [76]). We have also developed a more generic method which takes advantage of the symmetries in the architecture of the system [77].
  The main drawback of centralized approaches is that they require to explicitly build the global model of the system which is unrealistic for large and complex systems as telecommunication networks. It is why our more recent work deals with a decentralized approach [72]. This approach can be compared to R. Debouk and colleagues [56], [55] and also to P. Baroni and colleagues [50], [49]. Our method, unlike R. Debouk et al., relies on local models. We do not need to construct a global model. Indeed, the size of the global model would have been too large in our applications. Even if the methods are very close, P. Baroni et al. are concerned with an *a posteriori* diagnosis (off-line) whereas we propose an on-line diagnosis. Each time an alarm comes, it is analyzed and the diagnosis hypotheses are incrementally computed and given to the operator. Our main theme of study is close to E. Fabre and colleagues [62], [47]. The main difference is that they propose a multi-agent approach where the diagnoses are computed locally at the component level using message exchanges, whereas we construct a global diagnosis which is given to the operator at the supervisor level.
- 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. One way to supervise dynamic systems is to recognize those chronicles on-line. The principle is to follow the possible chronicles corresponding to a set of received failure messages until finding one or several chronicles that satisfy all the constraints. To perform this task, we have to create a chronicle base that contains all the possible chronicles. This base must be updated each time the supervised system evolves physically or structurally. An expert is needed to create the chronicle base. However, this makes the maintenance of the base very expensive. That is why we prefer to use an automatic method to learn the base. Most of the studies on chronicle recognition are french [59], [74], [57] and are based on C. Dousson's thesis [58]. Applications generally deal with system monitoring (telecommunication network) and video-surveillance (underground, bank, etc...). Our research studies do not focus directly on the development of chronicle recognition systems but on the automatic acquisition of the chronicle base. This idea is developed in the next section.

3.2. Machine learning

Keywords: *Inductive Logic Programming (ILP), Machine learning*.

The techniques investigated in the group aim at acquiring and improving models automatically. They belong to the field of machine or artificial learning [54]. In this domain, the goal is the induction or the discovery of objects characterizations from their descriptions by a set of features or attributes. Our work is grounded on Inductive Logic Programming (ILP). %endabstract

A learning method is supervised if samples of objects to be classified are available and labeled by the class they belong to. Such samples are often called *learning examples*. If the examples cannot be classified a priori, the learning method is unsupervised. Kohonen maps, induction of association rules in data mining or reinforcement learning are typical unsupervised learning methods. From another point of view, learning methods can be symbolic, such as inductive rule or decision tree learning, or numerical, such as artificial neural networks.

We are especially interested in structural learning which aims at making explicit relations among data where such links are not known. The temporal dimension is of particular importance in applications we are dealing with, such as process monitoring in health-care, environment or telecommunications. Additionally, we consider that the comprehensibility of the learned results is of crucial importance as domain experts must be able to evaluate and assess these results. ILP is the learning technique that best meets these requirements. We use a supervised version of this technique but also intend to use the unsupervised version which is called *Relational Data Mining* [45].

ILP began in the early 80's, though not under this name, when knowledge representation paradigms coming from logic programming began to be used in the field of machine learning. Such a high-level language meets the needs of relational representations for the description of structured objects or true relations between objects

During the 90's, ILP has become a proper research topic at the intersection of domains such as machine learning, logic programming and automated deduction. The main goal of ILP is the induction of classification or prediction rules from examples and from domain knowledge. The ILP research field has been extended to data mining enabling the discovery of association rules describing the correlations between data descriptors. As ILP relies on first order logic, it provides a very expressive and powerful language for representing hypotheses as well as the domain knowledge, this is its major feature.

Formally, ILP can be described as follows: given a set of positive examples P and a set of negative examples N of some concept to be learned, a logical theory B called the background knowledge and a language L_H specifying which clauses are syntactically and semantically acceptable, the goal is to discover a hypothesis H in the form of a logic program belonging to L_H such that $\forall p \in P \ B \cup H \models p$ and $\forall n \in N \ B \cup H \neg\models n$. This definition can be extended to multi-class learning. From a computational point of view, the learning process consists in searching the hypothesis space, either top-down by refining clauses that are too general (that cover negative examples) by adding literals to clause body or bottom-up by generalizing clauses that are too specific (that do not cover enough positive examples) by deleting literals or transforming constants into variables in literals. An interesting property is that the clause space has a lattice structure which enables an efficient search.

ILP is mainly used for learning classification rules. Similar techniques can also be used for inducing decision trees as well as for first order regression. The goal of regression is to predict the value of a real variable instead of a class value. Some more recent extensions deal with learning dynamic models: one such extension uses a representation coming from the qualitative simulator QSIM [66], another enables the discovery of differential equations from examples describing the behavior of a dynamic system [60].

Nowadays, work in ILP is mainly concerned with improving learning robustness (dealing with noisy or incomplete data) or efficiency (improving the search space exploration by taking structural properties into account, by stochastic techniques or by parallelizing algorithms for massively parallel computers). Another research direction investigates how to associate ILP to other learning methods which are more efficient for particular kind of data or to associate different learning strategies during ILP search. Extending the language to full first-order is also investigated. In this direction, learning from temporal data is of major interest because many application domains, such as telecommunications, health-care or environment, provide huge amounts of such data. This is why we have chosen to rely upon work by C. Rouveirol and M. Sebag [80] who have shown the value of associating ILP to CLP (Constraint Logic Programming) in order to compute efficiently numerical values. D. Page [70] wrote that a final challenge for ILP is to elaborate tight collaboration schemes between experts and ILP systems for knowledge discovery in order to avoid their complexity i) by enabling the evaluation of alternative hypotheses and not only those that maximize some heuristic function, ii) by devising tests and experiments for choosing among several hypotheses, iii) by providing non numerical justifications of the hypotheses such as belief measures or illustrative examples, iv) by consulting the expert when anomalies are detected in the data.

Our work is more concerned with the application of ILP rather than developing or improving the techniques. Nevertheless, as noticed by Page and Srinivasan [70], the target application domains (such as signal processing in health-care) can benefit from the adaptation of ILP to the particular features of the application data. Thus, 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, some variables are devoted to the representation of temporal phenomena and are managed by a constraint system [73] in order to deal efficiently with the associated computations (such as the covering tests, for example).

4. Application Domains

4.1. Panorama

The following application domains are concerned by our work: telecommunication networks, medicine and environment.

4.2. Telecommunication networks

Keywords: *telecommunications*.

Monitoring telecommunication networks is an important task and is one of the conditions to ensure a good quality of service. Given a monitoring system continuously receiving observations (alarms) sent by the system components, our purpose is to help operators to identify failures.

In this context, we developed a decentralized component-oriented approach, able to incrementally compute on-line diagnoses [71]. The efficiency of the algorithm is increased by the use of model-checking techniques as partial order reduction techniques and BDD. Currently, we are extending our research to reconfigurable systems, i.e systems the topology of which is changing along time, due for instance to reconfiguration actions decided to remedy upload problems.

Another important challenge for telecommunication networks is predicting the subjective quality of proposed services (as it could be felt by the user) from collected technical data. Mixing data-mining and symbolic learning techniques is the way we chose to acquire this predictive knowledge.

A last issue is the security of these networks and we are starting a joint work with Lande (M. Ducassé) and France-Telecom R&D funded by a CRE (external research contract). The main idea is to use the chronicle acquisition techniques developed in the cardiac domain in order to acquire automatically intrusion detection patterns.

All this research work on telecommunication networks is done in collaboration and with the support of France-Telecom R&D.

4.3. Decision aiding in medicine and health-care

Keywords: *health-care, medicine*.

Since the development of expert systems in the 70's, decision aiding tools have been widely studied and used in medicine and health-care. The ultimate goal is to help a physician to establish his diagnosis or prognosis from observations delivered by sensors and the individual patient's data. This involves at least three tasks:

- patient monitoring: processing and abstracting signals recorded by sensors placed on patients, in order to generate alarms when a particular situation has occurred, or is about to occur. The standard context is intensive care units in hospitals where an alarm must be treated within a very short time. With the advent of telemedicine similar situations arise, but the delay to treat an alarm may be much longer. For example, a cardiac or diabetic patient may be surveyed at home and the recorded data are sent every day at some fixed hour to the care unit. If some problem is detected, the patient is urged to consult a doctor, but a long delay may occur between the time at which the problem occurred and the treatment. Time is a major feature of medical data, thus temporal abstraction associated to signal processing techniques must be used for filtering and pre-processing the raw data;
- diagnostic and prognostic reasoning: models, such as causal or probabilistic models, have supplanted expert systems for diagnosis. As the course and outcome of a disease process is dynamic, time plays also an important role in diagnostic and prognostic models. Also, treatment planning or/and the clinical context may interact with these two basic reasoning processes and particular methods have to be studied and implemented to integrate these aspects;

- modeling: though some particular parts of the human body are known very well (e.g. the heart), deep models are generally difficult to build in medicine because of incomplete or too complex knowledge (e.g. the brain). Fortunately, huge amounts of data have been recorded and stored in medical databases. These data can be analyzed in order to discover new knowledge that may be used to construct abstract models or behavioral models, very similar to the old expert systems, but avoiding the bottleneck of expert knowledge acquisition. Processing medical data is a specific research area known as "intelligent data analysis (IDA) in medicine" [69]. An essential feature of the techniques used in IDA is that most are knowledge-based: they can use knowledge about the problem domain. Thus, a learning approach such as inductive logic programming is quite relevant.

These three points are studied in projects involving industrial (ELA medical), medical (University Hospital of Rennes) and academic (LTSI - University of Rennes) partners, especially in the field of cardiology. Particularly, new cardiac devices and monitoring systems are investigated.

4.4. Environmental decision making

Keywords: *environment*.

The need of decision support systems in the environmental domain is now well-recognized. It is especially true in the domain of water quality and a program, named Bretagne Eau Pure (http://www.bretagne-eau-pure.org), was launched a few years ago in order to help regional managers to protect this important resource. 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 quality and quantity of their crops. The difficulty is then to find solutions which satisfy contradictory interests and first to get a better knowledge on pollutant transfer. For instance, it is certainly true that the pesticide transfer through catchments is still not enough analyzed and poorly understood.

In this context, we are developing decision support systems to help regional managers in preserving the river water quality. Two main artificial intelligence techniques are used in this area: multi-agents systems, which are suited to model multi-expert cooperation, and qualitative modeling, to model biophysical processes in an explicative and understandable way. The approach we advocate is the coupling of a qualitative biophysical model, able to simulate the biophysical process, and a management model, able to simulate the farmer decisions.

Two main research themes are investigated in this framework: the use of qualitative spatial modeling to simulate the pollutant transfer through agricultural catchments and the use of learning/data mining techniques to discover, from model simulation results, the discriminant variables and acquire rules relating these variables. In both cases, one of the main challenges is that we are faced with spatio-temporal data.

Our partners are mainly the SAS Inra research group, located in Rennes and other Inra research groups as the BIA group in Toulouse and the LASB group in Montpellier.

5. Software

5.1. Introduction

The pieces of software described in this section are prototypes implemented by members of the project. They are not available through the APP. Any interested person should contact relevant members of the project.

5.2. Calicot: intelligent cardiac monitoring

Keywords: *chronicle recognition*, *diagnosis*, *monitoring*, *signal processing*, *temporal abstraction*.

Participants: François Portet, René Quiniou.

CALICOT (Cardiac Arrhythmias Learning for Intelligent Classification of On-line Tracks) is a software that takes as input several signals coming from sensors and that delivers as output fault states or diseases that were diagnosed by recognizing characteristic temporal patterns on the monitored signals. CALICOT is devoted to monitoring cardiac patients and diagnosing cardiac arrhythmias.The sofware is mainly implemented in Java with a few modules in Prolog and C. The main features of CALICOT are:

- a base of signal processing algorithms for abstracting signals into time-stamped symbolic events,
- a base of chronicles that are used on line by the chronicle recognizer[1]. Chronicles are discriminant temporal patterns related to arrhythmias. They are learned automatically off line from signal extracts by using ILP techniques,
- a pilot that adapts the behavior of the system to the monitoring context: noise on signals, patient's state, relevant arrhythmias, etc,
- a graphical interface that displays the recognized patterns on the signal curves and shows the related diagnoses.

5.3. CausalGraphed: an editor for building and diagnosing temporal causal graphs

Keywords: *temporal causal graphs*, *temporal diagnosis*.

Participants: René Quiniou, Sophie Robin.

CausalGraphed is an advanced editor which aims at aiding a user to build a fault model as a temporal causal graph. The tool has two main functionalities:

- editor: a graphical interface is provided to the user to let him build a temporal causal graph from scratch or edit a temporal causal graph coded in a subset of the dot description language [67],
- diagnoser: given a set of observations and their dates of occurrence, the faults that explain these observations can be searched for. The diagnosis is based on an abduction process which is animated on the graphical representation of the causal graph. Temporal constraints are propagated in order to restrict the set of solutions and to establish the direct temporal relation between a fault occurrence and caused observations. The predicted observations entailed by a set of time-stamped faults can also be computed in order to verify the soundness of the model.

5.4. FuzzyTreeGA: induction of fuzzy decision trees

Keywords: *fuzzy decision tree*, *genetic algorithm*, *induction*.

Participant: Pascal Garcia.

We have developed a fuzzy decision tree induction software with the following set of features:

- automatic induction of the tree based on data,
- automatic optimization of the fuzzy sets using genetic algorithms,
- control of the number of rules during the optimization process. We can make a trade-off between accuracy on the learning set and generalization capabilities (generally, a smaller set of rules have better generalization capabilities).

The software will be available soon on the DREAM web-site.

[1] CRS (Chronicle Recognition System) from France Telecom R & D: http://crs.elibel.tm.fr/.

6. New Results

6.1. Diagnosis of large scale discrete event systems

Participants: Marie-Odile Cordier, Alban Grastien, Christine Largouët, Laurence Rozé.

The problem we deal with is monitoring complex and large discrete-event systems (DES) such as telecommunication networks. Diagnosing dynamical systems represented as DES consists in finding what happened to the system from existing observations. Different terminologies can be found in the literature as histories, scenarios, narratives, consistent paths. They all rely on the idea that 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. The two main difficulties are i) the intractable size of the model and the huge number of states and trajectories to be explored; ii) the on-line change in the system topology and behavior.

To cope with the first difficulty, we proposed to use a decentralized approach which allows us to compute on-line diagnosis without requiring the computation of the global model. Given a decentralized model of the system and a flow of observations, the program computes the diagnosis by combining local diagnoses built from local models (or local diagnosers). In our case, we suppose that a supervisor exists and is in charge of merging these local diagnoses and to display them to an human operator. A paper describing the formal framework and its experimentation on telecommunication networks have been published in [11]. It relies on the work done during Y. Pencolé's thesis.

6.1.1. Dealing with uncertain observations

In real systems, the observations emitted by the system are often transmitted to the supervisor through communication channels. During this transfer, the observations may be reordered, lost, modified. Moreover, the component clocks are not synchronous and emission dates cannot be used to order the observations. The approaches previously developed in our team [11] only considered partially ordered observations. During this year, we extended this work by proposing to represent the uncertainty on emitted observations by an automaton. The diagnosis can then be defined as resulting from the synchronization between the observation automaton and the automaton modeling the system behavior. This work and its advantages are presented in [7].

6.1.2. Using automata chains for an incremental computation of on-line diagnosis:

The observation automaton can be huge when the observations are emitted during a long period. To ensure more flexibility, the concept of automata chain was first presented in [21] and then extended in [22], [23], [20]. With this formalism, the observation automaton is represented by a sequence of automata, each automaton representing the observations emitted during a period slice. It is then possible to compute the behavior for each period slice and to build a diagnosis chain.

The flexibility of the automata chain allows us to consider incremental computation, which is a fundamental step for on-line diagnosis. The incremental computation consists in computing a first diagnosis, and reusing it to efficiently compute a new diagnosis with additional observations. When new observations are known, the observation chain can be augmented by adding a new automaton. The incremental computation consists in computing the behavior during this window (see [21], [22], [23], [20], [7]). The extension to on-line diagnosis is presented in [7].

6.1.3. Extending the decentralized approach to reconfigurable systems:

The decentralized approach as described in [11] has been adapted to the diagnosis based on automata chains (see the previous paragraph) in [24].

Moreover, this approach has been extended to deal with reconfigurable systems (Alban Grastien's thesis). By *reconfigurable systems*, we mean systems in which addition, removal or modification of components or connections between components are allowed. These changes result from (on-line) reconfiguration actions. This is for instance the case of telecommunication networks whose topology can be changed to solve overload problems, and whose components can be replaced by new ones when they are defective. In a first step we proposed a formal characterization of reconfigurable systems and of reconfigurable systems diagnosis, based on the decentralized approach [65]. In a second step, the diagnosis by automata chain has been adapted for diagnosing reconfigurable systems [7]. The algorithms were implemented in a prototype called RAG_e and an interface was implemented during a summer training by François-Xavier Payet.

6.1.4. Diagnosability of discrete-event systems :

In cooperation with the VERTECS/IRISA research group, we proposed in [25] a model of supervision patterns corresponding to reachability properties. This allows us to generalize the properties to be diagnosed and to deduce techniques for verification of diagnosability of discrete-event systems. These techniques are general enough to cover a large class of diagnosis problems found in the literature, e.g. diagnosing permanent faults, multiple faults, fault sequences and some problems of intermittent faults.

6.2. Model acquisition from signals and surveillance in a monitoring context

Participants: Marie-Odile Cordier, Élisa Fromont, François Portet, René Quiniou.

Together with the LTSI (Signal Processing Lab - Inserm, University of Rennes 1), we are studying how to use chronicle recognition techniques for cardiac monitoring and diagnosis. Our goal is to analyze the signals coming from several sensors in order to detect and characterize the cardiac arrhythmias a monitored patient is subject to. The nature of the arrhythmias, their features and their frequency can be used to propose convenient therapies, such as specific drugs or cardiac devices (pacemaker or defibrillator). Two aspects have been particularly studied: chronicle discovery by machine learning and improving event detection on signals. These studies are validated by the construction of a cardiac monitoring system prototype called CALICOT [1].

6.2.1. Chronicle discovery by machine learning:

Machine learning techniques have been adapted to deal with multichannel aspects. We are particularly working on symbolic knowledge fusion which appears to be a hot topic in the knowledge acquisition community. This is the subject of Élisa Fromont's thesis [6] which is supported by a grant from the RNTS Cepica project (cf. 7.2). Various control policies have been implemented and assessed:

- global learning and recognition from the whole information provided by all the channels,
- independent learning on each channel and then symbolic knowledge fusion for global recognition,

The data from the different channels can be seen as different views of a common phenomenon and so the information is sometimes redundant and sometimes complementary.

- To exploit redundant data, we have designed appropriate solutions for learning accurate monosource rules from data coming from each source separately. Then, a voting procedure such as boosting can be used to select the relevant classification rules according to the state of the signal (in particular, this method can improve the results when the rules are used to detect specific phenomena on noisy signals).
- To exploit the complementary nature of the data, we propose to learn from the aggregated data with a learning bias, which should be as less restrictive as possible[2], to be able to learn rules that contain relationships between events occurring on the different sources. However, this technique is not efficient and the huge search space associated with the multisource learning problem leads to unsatisfactory solutions when using a relational learner. To solve those dimensionality problems a new multisource learning method has been defined (sketched in [10] and further developed in [17], [18]) that uses learned monosource rules to bias automatically and efficiently a new learning process from the aggregated data.

The results show that the proposed method is much more efficient than learning directly from the aggregated data. Furthermore, it yields rules having better or equal accuracy than rules obtained by monosource learning.

[2]Using a learning bias is mandatory to cope with the size of the hypothesis space.

6.2.2. Improving event detection on signals:

Monitoring systems, like CALICOT, are often faced with signals of varying quality. We are studying how to adapt signal processing algorithms and medical diagnosis to the current context (noise type and level on the ECG signal, shape of ECG beats, current arrhythmia hypotheses, etc.) for monitoring systems like CALICOT. This is the subject of François Portet's thesis [8] . Program supervision techniques have been studied to design a pilot module which has been integrated into CALICOT (the new monitoring system is called IP-CALICOT). To bring adaptation into every part of the system the pilot module operates at three stages [30], [31]:

1. It chooses and tunes the signal processing algorithms that analyze the input signal. Signal processing algorithms used for ECG analysis have been studied extensively in order to assess and to model their performance in different contexts [12]. Expert rules have been devised from this study. They are used by the pilot to choose the most relevant algorithm according to the current context. The performance of the monitoring system have been notably improved [33].
2. It activates or deactivates the different processing tasks devoted to the extraction of specific events from the ECG input signal. Piloting rules enabling temporal abstraction task activation or deactivation have been implemented. They are fired according to the signal noise and error expectation: the pilot module deactivates some task when there is too much noise on the signal to perform this task without errors. The preliminary results of this pilot stage are very encouraging [32].
3. It adapts the medical diagnosis to the current resolution of the signal analysis. In our approach, the diagnosis is achieved by chronicle recognition. In order to adapt chronicle recognition to the context, a hierarchy of chronicle bases has been set up. More abstracted chronicles use less objects (event descriptions) which have fewer attributes. Less abstracted chronicles are more precise and use more objects described by more attributes. The current chronicle base is chosen according to the level of abstraction imposed by the recognition task or the representation of events: on the one hand, a more abstracted chronicle base needs less low-level computation than a more refined one, on the other hand, more abstracted events are more reliable and easier to computed from noisy signals. Using this kind of chronicle hierarchy leads to a smarter use of computational resources.

6.2.3. (No title)

We are still working on this last subject. Another objective is to use multisource ECG to improve the quality of the diagnosis and reduce the effect of noise. Besides, the multisource learned rules have been transformed into chronicles and implemented in the piloted cardiac monitoring system IP-CALICOT and tested on real noisy signals. Preliminary results have been published in [16].

6.3. Learning and mining from temporal data

Participants: Marie-Odile Cordier, René Quiniou, Alexandre Vautier.

One well known definition of data mining is “discovering previously unknown and potentially useful information from large collections of data”. This definition does not take into account *what* knowledge needs to be discovered from the data and *who* will use this discovered knowledge. This is why we prefer to define data mining as the domain of abstracting large collections of data to *decrease their size* in order to make them readable by a “user” *without loosing important information* (the user can be a human or a machine). In data-mining the size of the dataset can be very high and this feature deserves specific algorithms. This is what distinguishes data mining from data analysis: Our work focuses on these two aspects of data mining, specifically *temporal* data mining.

The concept of inductive databases is an attempt to formalize the mining process from a database containing the original data as well as knowledge induced from these data and represented as patterns. The method belongs to the field of relational data-mining and can be seen as an adaptation of relational learning to the specific features of data-mining. The user queries an inductive database in such a way to specify which patterns are interesting for a special task. We are investigating how to enrich the representation of patterns in order to let an inductive database cope with metric temporal information. Until now, only the sequential aspect of data was taken into account but more precise information is often needed in order to supervise patients, plants or networks, for instance. As for our work on temporal relational learning [15], we are interested in the discovery of temporal patterns represented by chronicles.

We have devised a complex generality relation over chronicles which relies on the inclusion of sets of events (itemsets) as well as on the inclusion of temporal constraints stating lower and upper bounds on delays between events. The version space algorithm at the heart of the inductive database formulation has been modified to take the proposed generality relation, especially the notion of maximally specific chronicle, into account [39]. We have introduced the notion of recognition criterion [38], [37] that generalizes the specification of frequent patterns computation on temporal data.

Inductive database disregards the second aspect of data mining: the amount of patterns returned by a query is often very high and an expert of the application domain may encounter great difficulties to understand them. To alleviate these difficulties we are studying the abstraction of very long sequences to make them understandable. Due to poor domain knowledge, general processing techniques are often needed to correlate events in sequences. We have developed some syntactic and semantic techniques to summarize a long sequence as a hierarchy of abstracted and object transactions which are represented by chronicles.

Precisely, in the framework of a CRE contract with France Telecom R&D and the Lande project (cf. section 7.3), we are working on sequences (alarm logs) coming from a Virtual Private Network server that generates many alarms due to abnormal and normal connections [43], [42]. Abstraction techniques using Kolmogorov complexity are studied to summarize sequential events into transactions more easily understandable by an expert.

A second aspect of our work is devoted to the extraction of knowledge from an alarm stream. The goal is to diagnose on line a distributed deny of service (DDOS) by correlating network alarms in a data stream. Due to low computation capabilities of the diagnostic machine on huge volume of streaming data, patterns must be easily recognizable on line (in a computational sense).

6.4. Learning decision rules from simulation data - Application to water resource management domain

Participants: Marie-Odile Cordier, Pascal Garcia, Véronique Masson, Ansaf Salleb, Ronan Trépos.

In the framework of the SACADEAU project, our aim is to build decision support systems to help catchment managers to preserve stream-water quality [14]. In collaboration with Inra researchers, three actions are conducted in parallel.

- The first one has consisted in building a qualitative model to simulate the pesticide transfer through the catchment from the time of its application by the farmers to the arrival at the stream. The model architecture relies on the coupling of two models: a biophysical transfer model and a management model which can simulate the farmer decisions in herbicide application, depending on the climate and the weeding strategy, to cite only some of the decision criteria. Given data on the climate over the year, on the catchment topology and on the farmer strategy, the model outputs the pesticide concentration in the stream along the year. Though Inra is the main contributor, we actively participated to its realization. The model is now implemented and is currently in a validation phase.
- The second action consists in identifying some of the input variables as main pollution factors and in learning rules relating these pollution factors to the temporal distribution of the stream pesticide concentration. We chose to use Inductive Logic Programming (ILP) techniques to get easy-to-read and explicative rules We used the ICL software and obtained first interesting results on a simplified model. Most of learned rules express influences between parameters which were already well-known. But, interestingly enough, some of them underline influences and cause-effect relations the experts were not aware of, and even in some cases, influences partly contradictory with what was believed by experts. We are now beginning to extend our previous work on simulation results

obtained from the recently implemented model which means giving a more important place to spatial and temporal relations.
We have also tested fuzzy decision trees to induce rules in the specific case of propositional rules (in the general case of relational rules, this approach cannot be used). The results are similar in accuracy with the ILP approach but the set of rules is smaller. The fuzzy decision tree software is described in section 5.4.

- The last action consists in automatically analyzing the rules learned in the second step to help the experts in decision making. The aim of this step is to go beyond the simple use of classification rules in prediction by assisting the user to do an intensive post-analysis and exploitation of a large set of rules. Concretely, such analysis indicate what to do in order to reduce pollution whereas learned rules are classification rules predicting if a given farmer strategy or climate leads to a polluted or not polluted situation. We propose the algorithm DAKAR (Discovery of Actionable Knowledge And Recommendations) [35], [36] which works as follows : starting from an unsatisfactory situation that we want to improve and relying on a set of classification rules, DAKAR discovers a set of action recommendations to the expert of the domain in order to improve that situation. Actions are merely modifications in some attribute values describing the situation. Our algorithm explores the search space of actions looking for the "best actions" w.r.t. a quality criterion we have defined using once again the set of classification rules. To be "feasible", actions are also chosen so as to imply little changes in the initial situation. This is achieved in our framework thanks to an appropriate distance. In our experimental evaluation we used the above-cited model that allowed us to measure the relevance of actions discovered by DAKAR. Our experiments suggested also some extensions of our framework like to recommend simple actions ordered in time and to find what are the actions to avoid.

6.5. Diagnostic and causal reasoning

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

6.5.1. Configurations and causal inferences :

We have continued our work on designing a logic for causality. This work stems on [52] and is related to diagnosis where observed symptoms have to be explained by faults. It happens that none of the previously existing proposals is fully satisfactory, since either they are ad-hoc or, as in [53], [64], they are too close to standard logic in order to make a satisfactory diagnosis. The present proposal starts from propositional logic and introduces new *causal formulas*, built on *causal atoms* such as $\alpha \ causes\ < \beta_1, ..., \beta_n >$ intended to mean: "α causes one of the β_i's". The technical definitions are based on "configurations", which are descriptions of possible cases involving α, or β_i. The idea is that, in some context, α will produce β_1, and in another context α will produce β_2. The aim is to get a logic which makes a clear difference between (classical) implication and causal relations. A preliminary version has been published in [13]. It has now been significantly improved since we have now a full description of the logic and its proof system and a paper is currently submitted.

Either deductive reasoning and abductive reasoning can be performed from this causal formalism, which makes this system useful for diagnosis. Notice that the problems encountered by workers on diagnosis have been one of the main motivations for introducing default reasoning and that an important part of the presently active work of causation is an illustration of this long lasting close relationship.

6.5.2. Abductive diagnosis and Answer Set Programming :

We are also investigating practical solutions to temporal abductive diagnosis from causal graphs. Causal graphs provide a graphical representation of a faulty system behavior by means of (causal) relationships between faults and symptoms that can be observed when related faults are present. Nodes represent primary faults, intermediate (i.e. non observable) faults or symptoms. Directed arcs represent causality between two nodes. Arcs can be labeled by metric temporal constraints which set bounds on the delay between two nodes representing the occurrences of a fault and some of its consequences. Answer set programming [48] has been proposed for implementing diagnostic reasoning [61]. However, the computation model of ASP relies on the generation of all the instances of programs among which the models, i.e. the solutions of the problem, are selected. The method proposed so far [61] do not scale well when metric information, such as temporal constraints, is introduced in the representation as the number of instances of a program implementing a temporal causal graph may be huge. We have introduced a new programming method in ASP [41] that restricts notably the amount of instances which can be computed from a program. Also, ASP provides simple means to state preferences among solutions. Several extensions to causal graphs were proposed to model different notions of preferred temporal diagnosis.

6.5.3. Circonscription as default formalism :

We have also continued our work on one of the leading default formalism: circumscription. The idea of circumscription is to deal with the problem of default reasoning as follows. A "rule by default" (R) $if\ \ \alpha\ then\ \beta$ will be transformed into the rule $if\ \alpha\ and\ \neg exception_R\ then\ \beta$. Circumscription will then minimize all the exceptions, letting them to be true only when the overall set of information forces us to do so. This is a natural formalism with great expressive power, but it fails to be tractable. Recently, a new method has been proposed [68], where the old logical notion of forgetting propositional symbols (or reducing the logical vocabulary) has been generalized to the notion of forgetting literals, and this was shown to allow tractable computation of some kinds of circumscription. We have described an extension of this notion, allowing some propositional symbols to vary while forgetting literals, which can be applied to more kinds of circumscription [26], [27].

7. Contracts and Grants with Industry

7.1. Sacadeau: Decision-aid to improve streamwater quality

Participants: Marie-Odile Cordier, Véronique Masson, Ansaf Salleb, Ronan Trépos.

The project SACADEAU (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) has begun in October 2002. It is funded by Inra (French institute for agronomy research) and is ending this year. The project involves the following partners: three INRA research groups (SAS from Rennes, LASB from Montpellier and BIA from Toulouse) and Irisa. It also involves experts belonging to the regional administrative entities. The project aims at building a decision-aid tool to help specialists in charge of the catchment management in order to preserve the streamwater quality. The proposal relies on the building of two coupled qualitative models: a transfer model to simulate the pesticide transfer through the catchment and a management model to simulate the farmer decisions concerning the application of pesticides and the weeding strategy. The final objective was to analyze simulation results by using learning and data mining techniques, to discover the discriminant variables, to acquire rules relating the climate, the farmer strategy, the catchment topology with the pesticide concentration in the stream and to recommend actions.

7.2. Cepica: Conception and Evaluation of an Implantable Cardiac Device

Participants: Marie-Odile Cordier, Élisa Fromont, François Portet, René Quiniou, Ansaf Salleb.

This RNTS (Réseau National Technologies pour la Santé) project has begun at the end of 2003 and will last 3 years. The partners are ELA-Medical, the department of cardiology of the Rennes University Hospital, the LTSI-University of Rennes 1 and IRISA. The project is concerned with the conception of new cardiac devices, the study of which has begun during the instigative concerted action PISE. Its main concerns are: to propose and to evaluate new sensors able to assess the hemodynamic effects of a stimulation; to develop signal processing methods devoted to the specific signals measured by the new sensors and to refine, by using machine learning methods and chronicle recognition, the scenarios that may present some risk for an individual patient; to study different stimulation protocols taking into account the device specificities and constraints; to validate these concepts in clinical situations.

7.3. Using chronicles for network security

Participants: Marie-Odile Cordier, René Quiniou, Alexandre Vautier.

This CRE no 171978 (External Research Contract) is a focused collaboration between the project Dream, the project Lande (M. Ducassé) and France Telecom R & D on the problem of detecting specific network attacks. This study has begun in November 2004 and is planned to last three years. The first objective is to evaluate the use of chronicles, patterns of temporally constrained events, for representing and detecting attack scenarios on telecommunication networks. The second objective is to learn or discover automatically such attack scenarios from network logs, either generated by a simulation process or really observed on active networks.

7.4. WS-Diamond : Web Services : DIAgnosability, MONitoring and Diagnosis

Participants: Marie-Odile Cordier, Pascal Garcia, Laurence Rozé, Xavier LeGuillou.

WS-DIAMOND is a European project (Specific Targeted Research Project or STREP) dedicated to developing a framework for self-healing web services. It started in September 2005 and will last 36 months. It will produce

- an operational framework for self-healing service execution of conversationally complex web services. Monitoring, detection and diagnosis of anomalous situations, due to functional or non-functional errors (as quality of service), is carried on. Repair/reconfiguration is performed, guaranteeing reliability and availability of web services;
- a methodology and tools for service design that guarantee effective and efficient diagnosability/repairability during execution. The results will be demonstrated on real applications.

Our team is mainly involved in the fourth work package (WP4) which is concerned with model-based diagnosis and repair of cooperative web services. The challenge of this work package is to apply recent results and techniques developed for monitoring and reconfiguring complex physical systems, as telecommunication networks, to web services networks. Our aim is to apply decentralized diagnosis approaches developed for monitoring reconfigurable systems to this new application area.

8. Other Grants and Activities

8.1. National projects

Members of the DREAM team are involved in the following national collaboration programs:

- IMALAIA (common working group of the GdR Automatique, GDR- PRC I3 and Afia group) which brings together researchers from automatics and artificial intelligence fields on the subject of dynamic system monitoring. M.-O. Cordier is co-chair with L. Travé-Massuyès.
- RTP “information and intelligence: reasoning and decision” set up by the department STIC of the CNRS (M.-O. Cordier is a member of the steering committee).
- GdR I3 working group GT 3.4 (machine-learning, knowledge discovery in databases, data mining) - R. Quiniou, A. Salleb, A. Vautier.
- A. Salleb has collaborated with C. Vrain (LIFO, University of Orléans) on mining quantitative association rules [28], [29] and on the estimation of the density of datasets [34]. She has also worked with B. Duval (LERIA, Angers) and C.Vrain on mining exception rules [9].

8.2. International networks and working groups

- Monet2 (European network of excellence on Model-based and Qualitative reasoning). DREAM is particularly involved in the Bridge task group, which attempts to integrate AI and automatic methods for diagnosis and monitoring, and the biomedical task group which attempts to forge links between the fields of Biological Research and medical qualitative reasoning research (M.-O. Cordier and R. Quiniou).

8.3. Visits and invitations

- Yuhong Yan, researcher at the National Research Council in Fredericton, Canada, has been invited researcher (CNRS) in our team from may to mid-august 2005. She gave seminars on her work on web services and participated to our annual seminary. We (A. Grastien, M.-O. Cordier) contributed to a paper she was writing on the use of model-based approaches for monitoring web services [40].
- Marina Zanella, from university of Brescia, Italy, has been an invited researcher (University of Rennes1, Ifsic) in our team from June to July 2005. She gave a seminar on her work on diagnosing active systems. It allowed us to discuss new research directions with her and to start a cooperation on our common research themes.

9. Dissemination

9.1. Services to the Scientific Community

9.1.1. Journal editorial board

- *AAI: Applied Artificial Intelligence* (M.-O. Cordier)..
- *JEDAI: Journal Electronique d'Intelligence Artificielle* (M.-O. Cordier, until April 2005).
- *ARIMA: Revue Africaine de la Recherche en Informatique et en Mathématiques Appliquées* (M.-O. Cordier).
- *Revue I3* (M.-O. Cordier).

9.1.2. Conference program committees and organizations

- Program committee member of AI*IA'05 and DX'05; area chair for RFIA'06 (M.-O. Cordier).
- Co-chairs of the workshop "Temporal pattern extraction for on line detection of critical situations" at EGC 2005 (M.-O. Cordier, R. Quiniou).
- Chair (A. Vautier) and Webmaster (F. Portet) of the Majecstic'05 conference.
- Reviewing committee members of the Majectic'05 conference (E. Fromont, A. Grastien, A. Vautier, F. Portet et M.-O. Cordier).
- Co-chairs of the workshop "Temporal data mining" at EGC 2006 (M.-O. Cordier, R. Quiniou).
- Program committee member of "Data and Knowledge Quality" DKQ workshop at EGC 2005 (A. Salleb).
- Program committee member of "Quality Measures in Data Mining" Book. To appear in 2006, Springer in Studies in Computational Intelligence Series. (A. Salleb).

9.1.3. Scientific boards

- ECCAI board member (in charge of the bimonthly bulletin) : M.-O. Cordier

9.2. Academic teaching

Many members of the DREAM team are also faculty members and are actively involved in computer science teaching programs in Ifsic, INSA and ENSAR. Besides these usual teachings Dream is involved in the following programs:

- Master2 in computer science (IFSIC): *RATS module: temporal and spatial reasoning* (M.-O. Cordier, Y. Moinard, R. Quiniou).
- Master2 in computer science (IFSIC): *DIAG module: diagnosis* (M.-O. Cordier, S. Robin, L. Rozé).
- In charge of the DRT (diplôme de recherche technologique) at IFSIC (M.-O. Cordier)

9.3. Participation to workshops, seminars and miscellaneous invitations

- A. Salleb has given a talk entitled "Towards the Discovery of Interesting and Actionable Rules in Real-life Databases" on September 28th at the department of Computer Science, Columbia University, New York.

10. Bibliography

Major publications by the team in recent years

[1] G. CARRAULT, M.-O. CORDIER, R. QUINIOU, F. WANG. *Temporal abstraction and Inductive Logic Programming for arrhyhtmia recognition from electrocardiograms*, in "Artificial Intelligence in Medicine", vol. 28, 2003, p. 231-263.

[2] C. LARGOUËT, M.-O. CORDIER. *Improving the Landcover Classification using Domain Knowledge*, in "AI Communication special issue on Environmental Sciences and Artificial Intelligence", vol. 14, nº 1, 2001, p. 35-43.

[3] Y. MOINARD, R. ROLLAND. *Characterizations of Preferential Entailments*, in "Logic Journal of the IGPL", vol. 10, nº 3, 2002, p. 243-270, http://www3.oup.co.uk/igpl/.

[4] Y. PENCOLÉ, M.-O. CORDIER, L. ROZÉ. *A decentralized model-based diagnostic tool for complex systems.*, in "International Journal on Artificial Intelligence Tools", vol. 11, nº 3, 2002, p. 327-346.

[5] L. ROZÉ, M.-O. CORDIER. *Diagnosing discrete-event systems : extending the "diagnoser approach" to deal with telecommunication networks*, in "Journal on Discrete-Event Dynamic Systems : Theory and Applications (JDEDS)", vol. 12, nº 1, 2002, p. 43-81, http://www.inria.fr/rrrt/rr-4957.html.

Doctoral dissertations and Habilitation theses

[6] E. FROMONT. *Apprentissage multisource par programmation inductive : application à la caractérisation d'arythmies cardiaques*, Ph. D. Thesis, Université de Rennes 1, 2005, http://tel.ccsd.cnrs.fr/tel-00011455.

[7] A. GRASTIEN. *Diagnostic décentralisé et en-ligne de systèmes à événements discrets reconfigurables*, Ph. D. Thesis, Université Rennes 1, 2005.

[8] F. PORTET. *Pilotage d'algorithmes pour la reconnaissance en ligne d'arythmies cardiaques*, Ph. D. Thesis, University of Rennes 1, 2005.

Articles in refereed journals and book chapters

[9] B. DUVAL, A. SALLEB, C. VRAIN. *On the Discovery of Exception Rules: A Survey*, F. GUILLET, H. J. HAMILTON (editors). , Studies in Computational Intelligence, to appear, chap. in Quality Measures in Data Mining Book, Springer in the Series Studies in Computational Intelligence, 2006.

[10] E. FROMONT, M.-O. CORDIER, R. QUINIOU. *Extraction de connaissances provenant de données multisources pour la caractérisation d'arythmies cardiaques*, in "RNTI", to appear, 2005, http://www.irisa.fr/dream/dataFiles/efromont/rnti04_Fromont.pdf.

[11] 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", vol. 164, n^{o} 1-2, 2005, p. 121-170.

[12] F. PORTET, A. I. HERNANDEZ, G. CARRAULT. *Evaluation of real-time QRS detection algorithms in variable contexts*, in "Medical & Biological Engineering & Computing", vol. 43, n^{o} 3, 2005, p. 381-387.

Publications in Conferences and Workshops

[13] P. BESNARD, M.-O. CORDIER, Y. MOINARD. *Configurations for Inference from Causal Statements: Preliminary Report*, in "AI*IA 2005 (9th Congress of the Italian Association for Artificial Intelligence)", S. BANDINI, S. MANZONI (editors). , LNAI, n^{o} 3673, Springer, 2005, p. 282-285.

[14] M.-O. CORDIER, F. GARCIA, C. GASCUEL, V. MASSON, J. SALMON-MONVIOLA, F. TORTRAT, R. TREPOS. *A machine learning approach for evaluating the impact of land use and management practices on streamwater pollution by pesticides*, in "MODSIM'05 (International Congress on Modelling and Simulation)", MODELLING AND SIMULATION SOCIETY OF AUSTRALIA AND NEW ZEALAND (editor). , to appear, 2005.

[15] M.-O. CORDIER, R. QUINIOU. *Apprentissage relationnel de motifs temporels*, in "Atelier Extraction de motifs temporels pour la détection en ligne de situations critiques à EGC 2005", RNTI, Cépaduès, 2005.

[16] E. FROMONT, F. PORTET. *Pilotage d'un système de monitoring cardiaque multisource*, in "Actes de conférence de MajecStic'2005 (MAnifestation des JEunes Chercheurs STIC), Rennes, France", 2005.

[17] E. FROMONT, R. QUINIOU, M.-O. CORDIER. *Learning rules from multisource data for cardiac monitoring*, in "AIME'05 (Artificial Intelligence in Medicine), Aberdeen, Scotland", S. MIKSCH, J. HUNTER, E. KERAVNOU (editors). , Springer, 2005, p. 484-493, http://hal.inria.fr/docs/00/03/46/29/PDF/aime05_fromont.pdf.

[18] E. FROMONT, R. QUINIOU, M.-O. CORDIER. *Apprentissage multisource par programmation logique inductive*, in "RFIA 2006 (Congrès Reconnaissance des Formes et Intelligence Artificielle), Tours, France", to appear, 2006.

[19] P. GARCIA. *Induction de macro-actions en apprentissage par renforcement*, in "RFIA 2006 (Congrès Reconnaissance des Formes et Intelligence Artificielle), Tours, France", to appear, 2006.

[20] A. GRASTIEN, M.-O. CORDIER, C. LARGOUËT. *Automata Slicing for Diagnosing Discrete-Event Systems with Partially Ordered Observations*, in "AIIA'05 (Congress of the Italian Association for Artificial Intelligence), Milan, Italy", 2005, p. 270–281, http://hal.inria.fr/inria-00000531.

[21] A. GRASTIEN, M.-O. CORDIER, C. LARGOUËT. *First Steps Towards Incremental Diagnosis of Discrete-Event Systems*, in "Canadian AI'05, Victoria", B. KÉGL, G. LAPALME (editors). , Springer-Verlag GmbH, 2005, p. 170–181, http://www.springerlink.com/link.asp?id=ttpq7x9wutfqx6mq.

[22] A. GRASTIEN, M.-O. CORDIER, C. LARGOUËT. *Incremental Diagnosis of Discrete-Event Systems*, in "DX'05 (International workshop on principles of Diagnosis), Pacific Grove, California, USA", 2005, p. 119–124, http://hal.inria.fr/inria-00000121.

[23] A. GRASTIEN, M.-O. CORDIER, C. LARGOUËT. *Incremental Diagnosis of Discrete-Event Systems*, in "IJ-CAI'05 (International Joint Conference on Artificial Intelligence), Edinburgh, Scotland", L. P. KAELBLING, A. SAFFIOTTI (editors). , Poster, Professionnal Bok Center, 2005, p. 1564–1565.

[24] A. GRASTIEN. *Utilisation des chaînes d'automates pour le diagnostic décentralisé*, in "RFIA 2006 (Congrès Reconnaissance des Formes et Intelligence Artificielle), Tours, France", to appear, 2006.

[25] T. JERON, H. MARCHAND, M.-O. CORDIER. *Motifs de surveillance pour le diagnostic de systèmes à événements discrets*, in "RFIA 2006 (Congrès Reconnaissance des formes et Intelligence Artificielle), Tours, France", to appear, 2006, http://www.irisa.fr/dream/dataFiles/cordier/jeronetal-rfia05.pdf.

[26] Y. MOINARD. *Forgetting literals with varying propositional symbols*, in "Common Sense 2005 (7th Int. Symp. on Logical Formalization of Common Sense Reasoning), Corfu, Greece", S. MCILRAITH, P. PEPPAS, M. THIELSCHER (editors). , Technische Universitat Dresden, 2005, p. 169-176, http://www.iccl.tu-dresden.de/announce/CommonSense-2005/.

[27] Y. MOINARD. *Oubli de littéraux avec symboles propositionnels variables*, in "RFIA 2006 (Congrès Reconnaissance des Formes et Intelligence Artificielle)", P. BOUTHEMY, Y. DEMAZEAU (editors). , to appear, AFRIF-AFIA, 2006, xxx.

[28] C. NORTET, A. SALLEB, T. TURMEAUX, C. VRAIN. *Extraction de Règles d'Association Quantitatives Application à des Données Médicales*, in "EGC 2005 (Cinquièmes Journées sur Extraction et Gestion des Connaissances)", Cépaduès éditions, 2005, http://www.irisa.fr/dream/dataFiles/asalleb/egc05.pdf.

[29] C. NORTET, A. SALLEB, T. TURMEAUX, C. VRAIN. *Le rôle de l'utilisateur dans un processus d'extraction de règles d'association*, in "Atelier Qualité des Données et des Connaissances. Associé à ECG 2005", Cépaduès éditions, 2005.

[30] F. PORTET, G. CARRAULT, M.-O. CORDIER, R. QUINIOU. *Pilotage en ligne d'algorithmes de traitement du signal guidé par le contexte courant*, in "RJCIA'05 (Rencontres Jeunes Chercheurs en Intelligence Artificielle), Nice", 2005, p. 1-14.

[31] F. PORTET, G. CARRAULT, M.-O. CORDIER, R. QUINIOU. *Piloting Signal Processing Algorithms in a Cardiac Monitoring Context*, in "First Doctoral Consortium of the 10th Conference on Artificial Intelligence in Medicine (AIME 05), Aberdeen, Scotland", 2005.

[32] F. PORTET, G. CARRAULT, M.-O. CORDIER, R. QUINIOU. *Pilotage d'algorithmes pour un diagnostic médical robuste en cardiologie*, in "RFIA 2006 (Congrès Reconnaissance des Formes et Intelligence Artificielle), Tours, France", to appear, 2006.

[33] F. PORTET, G. CARRAULT. *Piloting Real-Time QRS Detection Algorithms in Variable Contexts*, in "3rd European Medical and Biological Engineering Conference (EMBEC'05), Prague", 2005.

[34] A. SALLEB, C. VRAIN. *Estimation of the Density of Datasets with Decision Diagrams*, in "ISMIS 2005, Saratoga Springs, New York", M.-S. HACID, N. MURRAY, Z. RAS, S. TSUMOTO (editors). , LNAI 3488, Springer, 2005, p. 688-697, http://www.irisa.fr/dream/dataFiles/asalleb/ismis05.pdf.

[35] R. TREPOS, A. SALLEB, M.-O. CORDIER, V. MASSON, C. GASCUEL. *A Distance Based Approach for Action Recommendation*, in "ECML 05 (European Conference on Machine Learning), Porto, Portugal", Lecture Notes in Artificial Intelligence, Springer, 2005.

[36] R. TREPOS, A. SALLEB, M.-O. CORDIER, V. MASSON, C. GASCUEL. *Une approche fondée sur une distance pour la recommandation d'actions*, in "RFIA'2006 (Congrès Reconnaissance des Formes et Intelligence Artificielle), Tours, France", to appear, 2006.

[37] A. VAUTIER, M.-O. CORDIER, R. QUINIOU. *An Inductive Database for Mining Temporal Patterns in Event Sequences (long version)*, in "PKDD (Principles and Practice of Knowledge Discovery in Databases) - Workshop mining spatio-temporal data, Porto, Portugal", 2005.

[38] 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). , Poster, 2005, p. 1640-1641.

[39] A. VAUTIER, M.-O. CORDIER, R. QUINIOU. *Extension des bases de données inductives pour la découverte de chroniques avec contraintes temporelles*, in "Proceedings of EGC 2005 (Extraction et Gestion de Connaissances)", N. VINCENT, S. PINSON (editors). , Cepadues editions, 2005.

[40] Y. YAN, Y. PENCOLÉ, M.-O. CORDIER, A. GRASTIEN. *Monitoring Web Service Networks in a Model-based Approach*, in "ECOWS'05 (European Conference on Web Services), Växjö, Sweden", 2005.

Miscellaneous

[41] S. LELOUETTE. *Raisonnement temporel pour le diagnostic par abduction*, Master dissertation, 2005.

[42] A. VAUTIER, M.-O. CORDIER, M. DUCASSÉ, R. QUINIOU. *Rapport d'avancement N°2 - contrat CRE/FT R & D n°171938*, 2005.

[43] A. VAUTIER, M.-O. CORDIER, R. QUINIOU. *Rapport d'avancement N°1 - contrat CRE/FT R & D n°171938*, 2005.

Bibliography in notes

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

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

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

[47] A. AGHASARYAN, E. FABRE, A. BENVENISTE, R. BOUBOUR, C. JARD. *Fault detection and diagnosis in distributed systems : an approach by partially stochastic Petri nets*, in "Discrete Event Dynamic Systems", Special issue on Hybrid Systems, vol. 8, n° 2, Juin 1998, p. 203-231.

[48] C. BARAL. *Knowledge Representation, Reasoning and Declarative Problem Solving*, Cambridge University Press, 2003.

[49] P. BARONI, G. LAMPERTI, P. POGLIANO, M. ZANELLA. *Diagnosis of a Class of Distributed Discrete-Event Systems*, in "IEEE Transactions on systems, man, and cybernetics", vol. 30, n° 6, Novembre 2000, p. 731-752.

[50] P. BARONI, G. LAMPERTI, P. POGLIANO, M. ZANELLA. *Diagnosis of large active systems*, in "Artificial Intelligence", vol. 110, 1999, p. 135-183.

[51] M. BASSEVILLE, M.-O. CORDIER. *Surveillance et diagnostic de systèmes dynamiques : approches complémentaires du traitement de signal et de l'intelligence artificielle*, Technical report, n° 1004, Irisa, 1996, http://www.irisa.fr/bibli/publi/pi/1996/1004/1004.html.

[52] 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, vol. 1638, Springer-Verlag, 1999, p. 55-67.

[53] 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.

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

[55] R. DEBOUK, S. LAFORTUNE, D. TENEKETZIS. *Coordinated Decentralized Protocols for Failure Diagnosis of Discrete Event Systems*, in "Discrete Event Dynamic Systems", vol. 10, n° 1-2, 2000, p. 33-86.

[56] R. DEBOUK, S. LAFORTUNE, D. TENEKETZIS. *A coordinated decentralized protocol for failure diagnosis of discrete event systems*, in "Proceedings of the Workshop on Discrete Event Systems (WODES'98), Cagliari, Italy", 1998, p. 138-143.

[57] C. DOUSSON. *Extending and unifying chronicle representation with event counters*, in "Proc. of ECAI 2002", F. VAN HARMELEN (editor). , IOS Press, 2002.

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

[59] 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.

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

[61] T. EITER, W. FABER, N. LEONE, G. PFEIFER. *The Diagnosis Frontend of the DLV System*, in "AI Communications", 1999.

[62] E. FABRE, A. BENVENISTE, C. JARD, L. RICKER, M. SMITH. *Distributed State Reconstruction for Discrete Event Systems.*, in "Proc. of the 2000 IEEE Control and Decision Conference (CDC), Sydney, Australie", 2000.

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

[64] E. GIUNCHIGLIA, J. LEE, V. LIFSCHITZ, N. MCCAIN, H. TURNER. *Nonmonotonic causal theories*, in "Artificial Intelligence Journal", vol. 153, n^o 1–2, March 2004, p. 49–104.

[65] A. GRASTIEN, M.-O. CORDIER, C. LARGOUËT. *Extending decentralized discrete-event modelling to diagnose reconfigurable systems*, in "DX'04 (International workshop on principles of Diagnosis), Carcassonne", 2004, p. 75–80, http://www.irisa.fr/dream/dataFiles/agrastie/dx04.ps.

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

[67] E. KOUTSOFIOS, S. C. NORTH. *Drawing graphs with dot*, Technical report, ATT Bell Laboratories, Murray Hill, NJ, 1996, http://citeseer.ist.psu.edu/article/koutsofios96drawing.html.

[68] J. LANG, P. LIBERATORE, P. MARQUIS. *Propositional Independence - Formula-Variable Independence and Forgetting*, in "Journal of Artificial Intelligence Research", vol. 18, 2003, p. 391–443.

[69] N. LAVRAČ, E. KERAVNOU, B. ZUPAN. *Intelligent data analysis in medicine*, in "Encyclopedia of computer science and technology", Marcel Dekker, Inc., New York, vol. 42, n^o 27, 2000, p. 113-157.

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

[71] Y. PENCOLÉ, M.-O. CORDIER, L. ROZÉ. *A decentralized model-based diagnostic tool for complex systems*, in "ICTAI'01 (IEEE international conference on tools with artificial intelligence)", IEEE computer society, 2001, p. 95-102.

[72] 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.

[73] 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, vol. 2157, Springer-Verlag, 2001, p. 220-227, http://www.irisa.fr/dream/dataFiles/quiniou/ilp01.pdf.

[74] 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.

[75] R. REITER. *A theory of diagnosis from first principles*, in "Artificial Intelligence", vol. 32, n° 1, 1987 (repris dans Readings in Nonmonotonic Reasoning, M. L. Ginsberg (dir.), Morgan Kaufmann, 1987 ; aussi dans Readings in Model-Based Diagnosis, W. Hamscher, L. Console, J. de Kleer (dir.), Morgan Kaufmann,p. 29-48, 1992), p. 57-96.

[76] L. ROZÉ, M.-O. CORDIER. *Diagnosing discrete-event systems : extending the "diagnoser approach" to deal with telecommunication networks*, in "Journal on Discrete-Event Dynamic Systems : Theory and Applications (JDEDS)", vol. 12, n° 1, 2002, p. 43-81, http://www.irisa.fr/dream/dataFiles/cordier/jeds.pdf.

[77] L. ROZÉ, M.-O. CORDIER. *Diagnosing Discrete Event Sytems : An experiment in Telecommunication Networks*, in "WODES98, Fourth Workshop on Discrete Event Systems, Cagliari, Italy", 1998, p. 130-137.

[78] L. ROZÉ. *Supervision de réseaux de télécommunication : une approche à base de modèles*, Ph. D. Thesis, Université de Rennes 1, 1997.

[79] M. SAMPATH, R. SENGUPTA, S. LAFORTUNE, K. SINNAMOHIDEEN, D. TENEKETZIS. *Active Diagnosis of Discrete-Event Systems*, in "IEEE Transactions on Automatic Control", vol. 43, n° 7, 1998, p. 908-929.

[80] M. SEBAG, C. ROUVEIROL. *Constraint Inductive Logic Programming*, in "Advances in Inductive Logic Programming", L. DE RAEDT (editor). , IOS Press, 1996, p. 277-294, http://citeseer.ifi.unizh.ch/sebag96constraint.html.

[81] J. DE KLEER, A. MACKWORTH, R. REITER. *Characterizing diagnoses and systems*, in "Artificial Intelligence", vol. 56, n° 2-3, 1992 (repris dans Readings in Model-Based Diagnosis, W. Hamscher, L. Console, J. de Kleer (dir.), Morgan Kaufmann,p. 54-65, 1992), p. 197-222.

[82] J. DE KLEER, B. C. WILLIAMS. *Diagnosis with behavioral modes*, in "Proceedings of the 11th International Joint Conference on Artificial Intelligence IJCAI'89, Detroit, MI, Etats-Unis", 1989 (repris dans Readings in Model-Based Diagnosis, W. Hamscher, L. Console, J. de Kleer (dir.), Morgan Kaufmann,p. 124-130, 1992), p. 1324-1330.