INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team DREAM

Diagnosis, REcommending Actions and Modeling

Rennes

THEME COG

Activity Report

2006

Table of contents

1. Team

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

Administrative assistant
Marie-Noëlle Georgeault [AA Inria]

Inria staff members
Lucie Callens [Associate Engineer, since September 2006]
Yves Moinard [Research Associate (CR) Inria]
Tristan Moreau [Associate Engineer, since November 2006]
René Quiniou [Research Associate (CR) Inria]
Thierry Vidal [Associate Professor on leave from ENIT Tarbes, since September 2006]

Faculty members
Pascal Garcia [Associate professor, Insa, Rennes, until October 2006]
Véronique Masson [Associate professor, University of Rennes 1 (50 % time)]
François Portet [ATER, University of Rennes 1, until August 2006]
Sophie Robin [Associate professor, University of Rennes 1]
Laurence Rozé [Associate professor, Insa, Rennes]

Ph. D. candidates
Élisa Fromont [Inria Ph. D. candidate, since October 2002 until February 2006]
Ronan Trépos [InraPh. 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, back in September 2006)]

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, the research themes which are developed are :

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). The problems which are investigated are : 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? Hox to improve the efficiency of the computation by using adequate symbolic representations such as BDD and partial order reduction techniques. 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.
   - It is important to formally characterizing intended properties of the system such as the diagnosability (i.e. the capability to detect and explain an error in due time) and repairability (i.e. the capability to get the system back to correctness, in due time). We are mainly interested in extending the solutions proposed for discrete-event systems and adapting them in order to develop decentralised approaches not requiring to rely on a global model as it is generally the case.
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). The problems which are investigated are : 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? how to apply learning techniques to spatio-temporal data? Another way is to help an expert to build the model. In this context, visualisation techniques are investigated in order to couple them to learning techniques.
3. We are focussing on computer assisted monitoring which means that we are not only interested in an efficient computation of on-line diagnoses but also in recommending actions in order to restore the functionalities of the system or the expected quality of service.

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. Computer assisted monitoring and diagnosis of physical systems

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 [60], [66]. 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 [49]. Computational efficiency is a crucial issue for real size problems. We are developing two approaches. The first one relies on diagnosers techniques [64], 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 [60], [66]. 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 [67], 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 [31], [33] 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 ,...) [37], [49].

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 [64]. We have extended this method to the communicating automata formalism [63] (see also [61]). We have also developed a more generic method which takes advantage of the symmetries in the architecture of the system [62].

  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 [57]. This approach can be compared to R. Debouk and colleagues [42], [41] and also to P. Baroni and colleagues [36], [35]. 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 [47], [34]. 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 [45], [59], [43] and are based on C. Dousson's thesis [44]. 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 [40]. 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).

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* [32].

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 $\forall n \in N\ B \cup H \not\models n$. 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 [51], another enables the discovery of differential equations from examples describing the behavior of a dynamic system [46].

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 [65] who have shown the value of associating ILP to CLP (Constraint Logic Programming) in order to compute efficiently numerical values. D. Page [55] 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 [55], 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 [58] 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 [56]. 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.

An important issue for telecommunication networks is the security of these networks. We have started a joint work with M. Ducassé from the LANDE (now LIS) IRISA research group and France-Telecom R&D funded by a CRE (external research contract). The main goal is to use chronicles for detecting intrusion or (distributed) attacks. Chronicle acquisition techniques that we developed in the cardiac domain have been proposed to acquire automatically intrusion detection or attack patterns.

Another important issue for telecommunication networks monitoring is to predict the subjective quality of monitoring and diagnosis from collected technical data. If many false alarms are generated or many faults are missed during monitoring, new diagnostic knowledge must be acquired. Furthermore, in the context of data streams new data arrive continously and cannot be stored. Thus, knowledge acquisition must be performed *on the fly*. Mixing data-mining and symbolic learning techniques are investigated to acquire this kind of predictive knowledge.

4.3. Software components

Keywords: *software components; web services*.

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: on-line monitoring, especially diagnosis and repair capabilities, become then a crucial concern.

We are working on this problem in the WS-DIAMOND project, a large European funded project involving eight partners in Italy, France, Austria and Netherlands. Our own work consists firstly in extending the decentralized component-oriented approach, initially developed for monitoring telecommunication networks [4], to this new domain. We are also involved in formally characterizing intended properties of the system such as the diagnosability (i.e. the capability to detect and explain an error in due time) and repairability (i.e. the capability to get the system back to correctness, in due time).

4.4. 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. Monitoring usually takes place in hospital intensive care unit where an alarm must be immediately treated. But more and more, patients are being monitored at home. In this context alarm processing can be delayed to some extent. 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: huge amounts of data have been and are still recorded in medical databases and these data can be analyzed in order to discover new knowledge. Abstract or behavioral models can be constructed. They are very similar to the old expert systems, but the bottleneck of expert knowledge acquisition is savoided. Processing medical data is a specific research area known as "intelligent data analysis (IDA) in medicine" [54]. An essential feature of the techniques used in IDA is that most are knowledge-based: they can use knowledge about the problem domain. Thus, acquisition techniques such as inductive logic programming or inductive databases are 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. A similar approach is being followed in the agricultural domain for managing big dairy herds (cf. 7.5).

4.5. 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: Lucie Callens, François Portet, René Quiniou [Correspondant].

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. FuzzyTreeGA: induction of fuzzy decision trees

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

Participant: Pascal Garcia.

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

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.

6. New Results

6.1. Diagnosis of large scale discrete event systems

Participants: Marie-Odile Cordier, Pascal Garcia, Christine Largouët, Xavier Le Guillou, Sophie Robin, Laurence Rozé, Thierry Vidal.

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. The diagnostic task consists then 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.

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 [4] only considered partially ordered observations. We proposed to represent the uncertainty on emitted observations by an automaton and extended the decentralized approach to cope with this new representation. This work was part of Alban Grastien's thesis and was presented in [15].

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

A problem, which is rarely considered, is the size of the diagnosis itself. However, it can also be rather huge, especially when dealing with uncertain observations. This is why we proposed to slice the observation flow into temporal windows and to compute the diagnosis in an incremental way from these diagnosis slices. In this context, we defined two independence properties (transition and state independence) and we showed their relevance to get a tractable representation of diagnosis. The diagnosis slices are economically represented by a set of transition-independent diagnoses and its associated set of abstract descriptions, from which the set of final states and the trajectories of the global diagnosis can be easily retrieved. This work was presented in [10] (and accepted to publication at IJCAI07).

6.1.3. Diagnosability of discrete-event systems:

In cooperation with the VERTECS/IRISA and the S4/IRISA research groups, we proposed in [16] a model of supervision patterns corresponding to reachability properties. This allowed 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. This work was extended and presented both to the model-based diagnosis community [17] and to the discrete-event community [18].

In cooperation with the DISCO/LAAS research group, we worked on comparing diagnosability for Continuous Systems (CS) and for Discrete Event Systems (DES) which were developed in two distinct and parallel tracks. The correspondences between the concepts used in CS and DES approaches were clarified and it was shown that the diagnosability problem can be brought back to the same formulation using the concept of signatures. This work was presented both to the model-based diagnosis community [11] and to the control theory community [12]. These results bridge CS and DES diagnosability and open perspectives for hybrid model based diagnosis.

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

Participants: Lucie Callens, Marie-Odile Cordier, Élisa Fromont, Tristan Moreau, François Portet, René Quiniou.

We are interested in diagnosing and monitoring methods which aim at recognizing, on signal streams, temporal patterns that can be related to specific interesting events [1]. Furthermore, we are interested in temporal pattern acquisition either by learning from examples or by mining data streams. Two aspects are particularly studied: chronicle discovery by machine learning and improving event detection on signals.

6.2.1. Chronicle discovery by machine learning

This year we have continued our work on the adaptation of machine learning to multichannel data. Our goal is to exploit the complementary and redundancy nature of data. We have proposed a new multisource learning method which is an extension of inductive logic programming. The process is in two steps: first, learn independently on each channel and, second, instead of merging directly the partial results, use the learned rules to bias automatically a new learning process from data aggregated from the different channels [13]. This way, the learning process is much more focused on what is really interesting on each channel. 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. We are currently working on the automatic translation and fusion of rules into a bias suited to inductive logic programming.

6.2.2. Improving event detection on signals

Monitoring systems, like CALICOT, are often faced with signals of varying quality. Taking advantage of context knowledge may generally facilitate signal processing and improve the results which are next exploited by diagnosis. We are studying how to adapt such monitoring systems to the current context. This means adapting signal processing algorithms to, e.g., the noise type and level of the raw signal or the shape of interesting waves, as well as to current failure or disease hypotheses predicted by diagnosis. Furthermore, diagnosis can also be adapted to the knowledge level allowed by signal processing or current possible diagnosis hypotheses. To this end, we have proposed to use a pilot module that operates at three stages [21]:

1. It chooses and tunes the signal processing algorithms that analyze the input signal. Expert rules are used by the pilot to choose the most relevant algorithm according to the current context.
2. It activates or deactivates processing tasks devoted to the extraction of specific events from the input signal.
3. It adapts the diagnosis to the current resolution of the signal analysis.

In our approach, diagnosis is achieved by chronicle recognition. Chronicles are organized in a hierarchy of chronicle bases: more abstracted chronicles use less objects (event descriptions) and fewer attributes; more specific chronicles are more precise and more detailed. 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.

This architecture shares many features with self-adaptive software [52]. We are beginning a study of self-adaptive methods elaborated in different domains to evaluate if they could be suited and adapted to the monitoring context.

6.2.3. Applications

The methods presented above are evaluated on two kinds of data: data recorded on cardiac patients (CEPICA contract, cf. section 7.2) and data recorded on dairy cattle (MOZAE project, cf. section 7.5). Two engineers have been enrolled for this purpose. Lucie Callens is working on the implemention the cardiac monitoring prototype IP-CALICOT. She is introducing or improving different aspects such as self-adaptation or providing better explanation by enhancing the ECG signal with graphical features. Tristan Moreau is working on zootechnic data recorded on dairy cattle. After applying signal processing algorithms in order to improve the data quality he will apply a methodology close to the one we have used on cardiac data for chronicle learning. Next, the results will be evaluated and discussed by veterinary experts.

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 kind of knowledge should 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 and decreasing 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.

More precisely, reducing the size of the dataset without loosing important information is similar to the Minimum Description Length (MDL) principle. Given a sample of data and an effective means to enumerate the relevant alternative theories that can explain the data, the best theory is the one that minimizes the sum of:

- the length, in bits, of the description of the theory,
- the length, in bits, of the data when encoded in the theory.

In addition, a theory has to be readable by a user. Thus, this user has to define the language on which the theory is built. The language is essentially composed of sets and relations between sets. We are currently working on two different domains to validate our method.

Precisely, in the framework of a CRE contract with France Telecom R&D and the LANDE/IRISA research group (cf. section 7.3), we are working on two types of alarm logs: logs from a Virtual Private Network server that generates many alarms due to normal and abnormal connections and alarm logs generated by France-Telecom routers when they detect suspicious flows.

In VPN (Virtual Private Network) abnormal and normal connections are mixed in a unique log. Network experts cannot read the whole log. Thus, we proposed a method to abstract the log into transactions (sets of alarms). Alarm attributes were extracted by matching patterns from a set of attribute signatures. Next, the alarms are built by correlating attributes values [23]. For an expert, reading the extracted transactions is then a lot easier than reading the whole log.

A second aspect of our work is devoted to the extraction of knowledge from a sequence of DDoS records [25]. Distributed Denial of Service (DDoS) attacks aim at overwhelming a target server with a huge volume of useless traffic from distributed and coordinated attack sources. A DDoS record indicates a suspicious Internet flow from sources to targets. France-Telecom uses DDoS records to detect DDoS attacks but the current method is not accurate enough. Thus, we developed a method for visualizing DDoS records in order to extract knowledge. From the visualization, several patterns were extracted that will serve as a basis for building an Internet flow model which to help detecting DDoS attacks.

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

Participants: Marie-Odile Cordier, Pascal Garcia, Véronique Masson, 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 [9]. In collaboration with Inra researchers, three actions are conducted in parallel.

- The first one 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 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. This model is now implemented and used for simulation.
- 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 stream pesticide concentration. Firstly, we used a simplified model to learn rules by Inductive Logic Programing (ILP). ILP aims to get easy-to-read and explicative rules. This first learning step showed the important impact of climate characteristics on streamwater pollution by pesticides. This work was extended on simulation results obtained from the newly implemented model. We take a particular interest in spatial relations between the cultivated plots and the characteristics of crop management practices. The goal is to focus on actionable factors. The idea is to base the learning process on the tree structure which is used by the run-off transfer model. Two approaches were compared: the first one consists in propositionalizing the examples and using a propositional learning process, precisely the decision list learner PART [48]. The second one consists in keeping the examples in the tree structured form and adapting the inductive logic programming (ILP) software for tree structured models.
- The last action consists in automatically analysing propositional 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. This analysis indicates 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 proposed the algorithm DAKAR (Discovery of Actionable Knowledge And Recommendations) [22] which works as follows : starting from an unsatisfactory situation and relying on a set of classification rules, DAKAR discovers a set of action recommendations and proposes them to the domain expert to improve the situation. The actions are built by selecting attributes among those describing the situation and proposing modifications on these attribute values. A heuristic evaluation function is used to evaluate the quality of the actions and rank them.

6.5. Diagnostic and causal reasoning

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

We have continued our work on designing a logic for causality. This work stems on [38] 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 [39], [50], 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. The full description of the logic and its proof system, which aims at automatizing the computation of the inference at hand,was improved [8]. Then either deductive reasoning and abductive reasoning can be made from this causal formalism. This constitutes our present on-going work which hopefully should make 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 on causation is an illustration of this long lasting close relationship. 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 [53], 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 continued our work on our own extension of this notion, allowing some propositional symbols to vary while forgetting literals, in order to deal with the really useful kinds of circumscription. New results (partially in [20], [19]) consist in an extensive description of a constructive method (given in a few variants) for computing the new notion. This is important since one of the main aim of [53] was to help the effective computation of already known formalisms. Our method proves that this help for computation also applies to the extended version, which is much more expressive than the original one.

7. Contracts and Grants with Industry

7.1. Sacadeau: Decision-aid to improve streamwater quality

Participants: Marie-Odile Cordier, Véronique Masson, 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) lasted from October 2002 to October 2005. It was funded by INRA (French institute for agronomy research). The project involved partners from INRA (mainly (SAS from Rennes and BIA from Toulouse) and IRISA. It also involved experts belonging to the regional administrative entities. Even if this contract is formally ended, the collaboration has been carried on during this year. A book is in preparation. We, with our partners, are involved in a new project, named APPEAU, funded by ANR and starting in 2007.

The SACADEAU 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 first step was to build a simulation model coupling two 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 next step is to rely on this model to build a decision-aid tool. Two studies were conducted: the first one consists in using symbolic learning and classification techniques on the simulation results to discover rules relating the climate, the farmer strategy, the catchment topology with the pesticide concentration in the stream. The second one consists in analysing learned classification rules in order to build recommendation actions for a given situation.

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

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

This RNTS (Réseau National Technologies pour la Santé) project has begun at the end of 2003 and ended in 2006. The partners were 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 instigating 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. CURAR: 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 DREAM research group, the LANDE (LIS) research group (M. Ducassé) and France Telecom R & D on the problem of detecting specific network attacks. This study began 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, Xavier Le Guillou, Sophie Robin, Laurence Rozé, Thierry Vidal.

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. We intend to extend the decentralized diagnosis approaches developed for monitoring reconfigurable systems to this new application area.

The deliverables 1.1, 4.1 and 4.2 and the annual report [28], [27], [29], [26] describe the work done in collaboration with our partners in this project.

7.5. MOZAE: Zootechnic monitoring of dairy herds

Participants: Marie-Odile Cordier, Tristan Moreau, René Quiniou.

The MOZAE (Monitoring Zootechnique des Animaux d'Élevage) project started in march 2006 and will last two years. It is granted by the ministry of Agriculture and the region of Brittany. The partners are the company Medria, the regional chamber of agriculture of Brittany, INRA (National Research Institute of Agronomy), Agrocampus-Rennes, ENVN (National Veterinary Scool of Nantes) and IRISA-INRIA.

The MOZAE project aims at designing and evaluating monitoring and assisting software tools for farmers managing big dairy herds (up to 500 cows). Furthermore, the project aims at discovering veterinary knowledge and at redefining protocols for cattle health-care. The temperature and cardiac rhythm of each individual cow is continuously monitored by a sensor introduced in the cow's stomach. These signals are regularly analyzed in order to detect physiological events that interest farmers, such as oestrus, disease or proximity of birthing.

Dream is in charge of processing signals and discovering the signatures of interesting events. Machine learning and data mining techniques will be used for discovering and refining temporal patterns with the aid of veterinary experts.

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. The Imalaia working group organized a two-days workshop on decentralised approaches for discrete-event systems diagnosis with international attendance.
- GdR I3 working group GT 3.4 (machine-learning, knowledge discovery in databases, data mining) - R. Quiniou, A. Vautier.

8.2. Visits and invitations

- Hendrik Blockeel, from Leuwen University (Katholieke Universiteit Leuwen) , has been an invited researcher (University of Rennes1, Ifsic) in our team during one month (may 2006). He gave a seminar on his work and participated to the annual seminar of our group. He had discussions with many of us on new research directions in the machine learning and data mining domain. We intend to start a cooperation on learning/mining data streams. Elisa Fromont, PhD student in 2005, has currently a post-doctoral position in H. Blockeel's research group in Leuwen.

9. Dissemination

9.1. Services to the Scientific Community

9.1.1. 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).

9.1.2. Conference program committees and organizations

- Program committee member for DX'06 and senior program committee member for IJCAI'07; area chair for RFIA'06 (M.-O. Cordier).
- Reviewers for IJCAI'07 (R. Quiniou and T. Vidal).
- Program committee member of the workshop "Temporal data mining" at EGC 2006 (M.-O. Cordier, R. Quiniou).
- Organizer of the workshop "Temporal data mining" at EGC 2006 (R. Quiniou with G. Hébrail, ENST, Paris).
- Program committee member of the workshop on "Data Streams" at EGC 2007 (R. Quiniou).
- Programme committee member of the IASTED International Conference on Artificial Intelligence and Applications - AIA 2007 (T. Vidal).

9.1.3. Scientific boards

- ECCAI (European Coordinating Committed for Artificial Intelligence) board member: M.-O. Cordier
- Member of "Agrocampus Rennes (ENSAR)" scientific board: 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

- M.-O. Cordier gave an invited talk at the ECAI Workshop on "Planning, Learning and Monitoring with Uncertainty and Dynamic Worlds".

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 arrhytmia 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, n^{o} 1, 2001, p. 35-43.

[3] Y. MOINARD, R. ROLLAND. *Characterizations of Preferential Entailments*, in "Logic Journal of the IGPL", vol. 10, n^{o} 3, 2002, p. 243-270, http://www3.oup.co.uk/igpl/.

[4] Y. PENCOLÉ, M.-O. CORDIER. *A formal framework for the decentralized 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.

[5] 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, n^o 3, 2002, p. 327-346.

[6] 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^o 1, 2002, p. 43-81, http://hal.inria.fr/inria-00071622.

Year Publications

Articles in refereed journals and book chapters

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

Publications in Conferences and Workshops

[8] 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, n^o 4092, Springer, aug 2006, p. 292–304, http://www.irisa.fr/dream/dataFiles/moinard/causeksempub.pdf.

[9] M.-O. CORDIER, F. GARCIA, C. GASCUEL, V. MASSON, A. SALLEB, R. TREPOS. *SACADEAU project : recommending actions from simulation results*, in "BESAI 2006", abstract and poster, August 2006.

[10] M.-O. CORDIER, A. GRASTIEN. *Exploiting independence in a decentralized and incremental approach of diagnosis*, in "DX'06 (17th International workshop on principles of Diagnosis), Spain", 2006, p. 61–69, http://www.irisa.fr/dream/dataFiles/cordier/dx06-cord-alban.pdf.

[11] M.-O. CORDIER, L. TRAVÉ-MASSUYÈS, X. PUCEL. *Comparing diagnosability in continuous and discrete-events systems*, in "DX'06 (17th International workshop on principles of Diagnosis), Spain", 2006, p. 55–60, http://www.irisa.fr/dream/dataFiles/cordier/dx06-cordetal.pdf.

[12] M.-O. CORDIER, L. TRAVÉ-MASSUYÈS, X. PUCEL. *Comparing diagnosability in continuous and discrete-events systems*, in "Safeprocess'2006, Beijing, China", 2006, http://www.irisa.fr/dream/dataFiles/cordier/safe06-cordetal.pdf.

[13] 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", vol. 2, Presses universitaires François-Rabelais, january 2006, 115, http://hal.inria.fr/inria-00001050.

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

[15] A. GRASTIEN. *Utilisation des chaînes d'automates pour le diagnostic décentralisé*, in "RFIA 2006 (Congrès Reconnaissance des Formes et Intelligence Artificielle), Tours, France", 2006, http://hal.inria.fr/inria-00000608.

[16] T. JÉRON, 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", january 2006, http://www.irisa.fr/dream/dataFiles/cordier/jeronetal-rfia05.pdf.

[17] T. JÉRON, H. MARCHAND, S. PINCHINAT, M.-O. CORDIER. *Supervision patterns in discrete event systems diagnosis*, in "DX'06 (17th International Worshop On Principles of Diagnosis), Spain", 2006, p. 117-124, http://www.irisa.fr/dream/dataFiles/cordier/dx06-jeronetal.pdf.

[18] T. JÉRON, H. MARCHAND, S. PINCHINAT, M.-O. CORDIER. *Supervision patterns in discrete event systems diagnosis*, in "WODES'06 (8th International Workshop on Discrete Event Systems), Ann Arbor, Michigan, USA", (extended version in Irisa Technical Report #1784), IEEE (ISBN 1-4244-0053-8), 2006, http://www.irisa.fr/vertecs/Publis/Ps/PI-1784.pdf.

[19] Y. MOINARD. *About the computation of forgetting symbols and literals*, in "NMR 2006 (11th Workshop on Non-monotonic Reasoning)", Institut für Informatik, Technische Universität Clausthal, Germany (http://cig.in.tu-clausthal.de/), june 2006, p. 209–217, http://www.irisa.fr/dream/dataFiles/moinard/forgetLakepub.pdf.

[20] Y. MOINARD. *Oubli de littéraux avec symboles propositionnels variables*, in "RFIA 2006 (Congrès Reconnaissance des Formes et Intelligence Artificielle)", P. BOUTHEMY, Y. DEMAZEAU (editors). , AFRIF-AFIA, jan 2006, p. 120 (1-10), http://www.irisa.fr/dream/dataFiles/moinard/forgetRFIApub.pdf.

[21] 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", january 2006.

[22] R. TRÉPOS, 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", january 2006.

[23] A. VAUTIER, M.-O. CORDIER, M. DUCASSÉ, R. QUINIOU. *Agrégation d'alarmes faiblement structurées*, in "Atelier fouille de données temporelles à EGC 06", january 2006, http://www.irisa.fr/dream/Pages_Pros/Alexandre.Vautier/publis/egc_06_vpn.pdf.

Internal Reports

[24] T. JÉRON, H. MARCHAND, S. PINCHINAT, M.-O. CORDIER. *Supervision Patterns in Discrete Event Systems Diagnosis*, Technical report, n^{o} 1784, Irisa, 2006.

[25] A. VAUTIER, M.-O. CORDIER, M. DUCASSÉ, R. QUINIOU. *Visualization of Internet Flow Records*, Rapport Avancement projet CURAR - CRE France Telecom R & D, Technical report, n^{o} 1828, IRISA, november 2006.

Miscellaneous

[26] WS-DIAMOND PROJECT. *Annual activity and management report*, september 2006.

[27] WS-DIAMOND PROJECT. *Characterization of diagnosis and repair for web services*, june 2006.

[28] WS-DIAMOND PROJECT. *Deliverable D1.1*, march 2006.

[29] WS-DIAMOND PROJECT. *Specification of diagnosis algorithms for web services - Phase 1*, october 2006.

[30] WS-DIAMOND PROJECT. *Specification of repair/reconfiguration algorithms for web services - Phase 1*, october 2006.

References in notes

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

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

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

[34] 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.

[35] 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.

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

[37] 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.

[38] P. BESNARD, M.-O. CORDIER. *Inferring causal explanations*, in "Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU'99)", A. HUNTER, S. PARSONS (editors). , Lecture Notes in Artificial Intelligence, vol. 1638, Springer-Verlag, 1999, p. 55-67.

[39] 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.

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

[41] 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.

[42] 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.

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

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

[45] 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.

[46] 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.

[47] 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.

[48] E. FRANK, I. H. WITTEN. *Generating Accurate Rule Sets Without Global Optimization*, in "Proc 15th International Conference on Machine Learning", Madison, Wisconsin, Morgan Kaufmann, 1998, p. 144-151.

[49] 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.

[50] 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.

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

[52] R. LADDAGA, P. ROBERTSON. *Self Adaptive Software: A Position Paper*, in "SELF-STAR: International Workshop on Self-* Properties in Complex Information Systems, Bertinoro, Italy", June 2004.

[53] 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.

[54] 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.

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

[56] 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.

[57] 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.

[58] 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.

[59] 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.

[60] 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.

[61] 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.

[62] 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.

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

[64] 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.

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

[66] 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.

[67] 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.