IN PARTNERSHIP WITH:
CNRS

Institut national des sciences appliquées de Rennes

Université Rennes 1

Activity Report 2011

Project-Team DREAM

Diagnosing, Recommending Actions and Modelling

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

RESEARCH CENTER
Rennes - Bretagne-Atlantique

THEME
Knowledge and Data Representation and Management

Table of contents

Project-Team DREAM

Keywords: Artificial Intelligence, Data Mining, Machine Learning, Diagnosis, Reasoning, Knowledge, Decision Aid

1. Members

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

Faculty Members
Marie-Odile Cordier [Team Leader, Professor, University of Rennes 1, HdR]
Véronique Masson [Associate professor, University of Rennes 1 (80 % time)]
Sophie Robin [Associate professor, University of Rennes 1 (till july 1st 2011)]
Laurence Rozé [Associate professor, Insa, Rennes]

External Collaborators
Philippe Besnard [Research scientist, DR CNRS, IRIT, Toulouse]
Louis Bonneau de Beaufort [Research Engineer, Agrocampus-Ouest, Rennes]
Thomas Guyet [Associate Professor, Agrocampus-Ouest, Rennes]
Christine Largouët [Associate Professor, Agrocampus-Ouest, Rennes (on leave since july 1st 2011)]
Karima Sedki [Associate Professor, Agrocampus-Ouest, Rennes (since september 1st 2011)]

Technical Staff
Florimond Ployette [IR Inria]

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

Administrative Assistant
Marie-Noëlle Georgeault [AA Inria]

2. Overall Objectives

2.1. Introduction

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

In this context, the research questions we are investigating are the following. Even though they are clearly highly related, we have chosen for the sake of clarity to present them in two distinct paragraphs. The first one deals with on-line monitoring issues and the second one with design and model acquisition issues. Finally, we present some application domains.

2.2. On-line monitoring issues

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

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

2.3. Design and model acquisition issues

When designing a dependable and adaptive system, a main point is to formally characterize the intended properties of the system such as the diagnosability (i.e. whether, given the system specifications, it is possible to detect and explain an error in due time), or the repairability (i.e. whether it is possible to get the system back to correctness, in due time). Moreover, these two properties must be joined to get the best compromise for building real self-healing systems. Some of these concepts have been defined, but in a centralized context. We aim at extending the solutions proposed so far for discrete-event systems in the decentralized context.

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

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

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

2.4. Application domains

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

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

3. Scientific Foundations

3.1. Computer assisted monitoring and diagnosis of physical systems

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 [59], [66]. Two main approaches have been proposed then: (i) the consistency-based approach, relying on a model of the expected correct behavior ; (ii) the abductive approach which relies on a model of the failures that might affect the system, and which identifies the failures or the faulty behavior explaining the anomalous observations. See the references [15], [17] for a detailed exposition of these investigations.

Since 1990, the researchers in the field have studied dynamic system monitoring and diagnosis, in a similar way as researchers in control theory do. What characterizes the AI approach is the use of qualitative models instead of quantitative ones and the importance given to the search for the actual source/causes of the faulty behavior. Model-based diagnosis approaches rely on qualitative simulation or on causal graphs in order to look for the causes of the observed deviations. The links between the two communities have been enforced, in particular for what concerns the work about discrete events systems and hybrid systems. Used formalisms are often similar (automata, Petri nets ,...) [22], [21].

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

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

- In the first method, the automaton used as a model is transformed off-line into an automaton adapted to diagnosis. This automaton is called a *diagnoser*. This method has first been proposed by M. Sampath and colleagues [61]. The main drawback of this approach is its centralized nature that requires to explicitly build the global model of the system, which is most of the time unrealistic. It is why we proposed a decentralized approach in [56].
- In the second method, the idea is to associate each failure that we want to detect with a *chronicle* (or a scenario), i.e. a set of observable events interlinked by time constraints. The chronicle recognition approach consists in monitoring and diagnosing dynamic systems by recognizing those chronicles on-line [37], [58], [35].

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

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

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

- The ultimate goal is usually not merely to diagnose, but to put back the system in some acceptable state after the occurrence of a fault. One of our focus is to develop self-healable systems able to self-diagnose and -repair.
- When designing a system and equipping it with diagnosis capabilities, it may be crucial to be able to check off-line that the system will behave correctly, i.e., that the system is actually 'diagnosable'. A lot of techniques have been developed in the past (see Lafortune and colleagues [60]), essentially in automata models. We extended them to deal with temporal patterns. A recent focus has been to study the self-healability of systems (ability to self-diagnose and -repair).

3.2. Machine learning and data mining

The machine learning and data mining techniques investigated in the group aim at acquiring and improving models automatically. They belong to the field of machine or artificial learning [32]. In this domain, the goal is the induction or the discovery of hidden objects characterizations from their descriptions by a set of features or attributes. For several years we investigated Inductive Logic Programming (ILP) but now we are also working on data-mining techniques.

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

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

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

Relational learning

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

Concerning temporal data, our work is more concerned with applying relational learning rather than developing or improving the techniques. Nevertheless, as noticed by Page and Srinivasan [55], the target application domains (such as signal processing in health-care) can benefit from adapting relational learning scheme to the particular features of the application data. Therefore, relational learning makes use of constraint programming to infer numerical values efficiently [62]. Extensions, such as QSIM [43], have also been used for learning a model of the behavior of a dynamic system [38]. Precisely, we investigate how to associate temporal abstraction methods to learning and to chronicle recognition. We are also interested in constraint clause induction, particularly for managing temporal aspects. In this setting, the representation of temporal phenomena uses specific variables managed by a constraint system [57] in order to deal efficiently with the associated computations (such as the covering tests).

For environmental data, we have investigated tree structures where a set of attributes describe nodes. Our goal is to find patterns expressed as sub-trees [30] with attribute selectors associated to nodes.

Data mining

Data mining is an unsupervised learning method which aims at discovering interesting knowledge from data. Association rule extraction is one of the most popular approach and has deserved a lot of interest in the last 10 years. For instance, many enhancements have been proposed to the well-known Apriori algorithm [19]. It is based on a level-wise generation of candidate patterns and on efficient candidate pruning having a sufficient relevance, usually related to the frequency of the candidate pattern in the data-set (i.e., the support): the most frequent patterns should be the most interesting. Later, Agrawal and Srikant proposed a framework for "mining sequential patterns" [20], which extends Apriori by coping with the order of elements in patterns.

In [50], Mannila and Toivonen extended the work of Aggrawal et al. by introducing an algorithm for mining patterns involving temporal episodes with a distinction between parallel and sequential event patterns. Later, in [36], Dousson and Vu Duong introduced an algorithm for mining chronicles. Chronicles are sets of events associated with temporal constraints on their occurrences. They generalize the temporal patterns of Mannila and Toivonen. The candidate generation is an Apriori-like algorithm. The chronicle recognizer CRS [34] is used to compute the support of patterns. Then, the temporal constraints are computed as an interval whose bounds are the minimal and the maximal temporal extent of the delay separating the occurrences of two given events in the data-set. Chronicles are very interesting because they can model a system behavior with sufficient precision to compute fine diagnoses. Their extraction from a data-set is reasonably efficient. They can be efficiently recognized on an input data stream.

Relational data-mining [16] can be seen as generalizing these works to first order patterns. In this field, the work of Dehaspe for extracting first-order association rules have strong links with chronicles. Another interesting research concerns inductive databases which aim at giving a theoretical and logical framework to data-mining [44], [33]. In this view, the mining process means to query a database containing raw data as well as patterns that are implicitly coded in the data. The answer to a query is, either the solution patterns that are already present in the database, or computed by a mining algorithm, e.g., Apriori. The original work concerns sequential patterns only [49]. We have investigated an extension of inductive database where patterns are very close to chronicles [64].

Mining data streams

During the last years, a new challenge has appeared in the data mining community: mining from data streams [18]. Data coming for example from monitoring systems observing patients or from telecommunication systems arrive in such huge volumes that they cannot be stored in totality for further processing: the key feature is that “you get only one look at the data” [40]. Many investigations have been made to adapt existing mining algorithms to this particular context or to propose new solutions: for example, methods for building synopses of past data in the form of or summaries have been proposed, as well as representation models taking advantage of the most recent data. Sequential pattern stream mining is still an issue [51]. At present, research topics such as, sampling, summarizing, clustering and mining data streams are actively investigated.

A major issue in data streams is to take into account the dynamics of process generating data, i.e., the underlying model is evolving and, so, the extracted patterns have to be adapted constantly. This feature, known as *concept drift* [65], [45], occurs within an evolving system when the state of some hidden system variables changes. This is the source of important challenges for data stream mining [39] because it is impossible to store all the data for off-line processing or learning. Thus, changes must be detected on-line and the current mined models must be updated on line as well.

4. Application Domains

4.1. Software components monitoring

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

We have been working on this problem within the WS-DIAMOND project [63], a large European funded project involving eight partners in Italy, France, Austria and Netherlands http://wsdiamond.di.unito.it/. Our own work consisted in two distinct contributions:

The first contribution has been to extend the decentralized component-oriented approach, initially developed for monitoring telecommunication networks [3] to this new domain. To this end we have proposed the concept of distributed chronicles, with synchronization events, and the design of an architecture consisting of distributed CRSs (Chronicle Recognition Systems) communicating their local diagnoses to a broker agent which is in charge of merging them to compute a global diagnosis. During his thesis, X. Le Guillou developped two approaches for solving this problem [46], [47], [48].

The second one was in formally characterizing the intended system properties 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), managing to relate them in order to define a so-called 'self-healability' property which will ensure the system is self-healing, i.e. it can always match a set of observations to an adequate repair procedure, which will resolve any of the faults which are consistent with the observations [31].

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

4.2. Environmental decision making

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

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

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

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

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

5. Software

5.1. Introduction

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

5.2. QTempIntMiner: quantitative temporal sequence mining

QTEMPINTMINER (Quantitative Temporal Interval Miner) is a software that implements several algorithms presented in [42] and [8].

The software is mainly implemented in Matlab. A standalone application is now avalaible. It uses the Mixmod toolbox [28] to compute multi-dimensional Gaussian distributions. The main features of QTEMPINTMINER are:

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

This year the software has been updated to include two new algorithms: QTIAPRIORI and QTIPREFIXSPAN. The software has been used to compare the efficiency of three algorithms. The software is currently applied to the characterization of cardiac arrhythmias.

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

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

SACADEAU is a software that implements the SACADEAU transfer model presented in section 8.2.1. The SACADEAU simulation model couples two qualitative models, a transfer model describing the pesticide transfer through the catchment and a management model describing the farmer decisions. Giving as inputs a climate file, a topological description of a catchment, and a cadastral repartition of the plots, the SACADEAU model simulates the application of herbicides by the farmers on the maize plots, and the transfer of these pollutants through the catchment until the river. The two main simulated processes are the runoff and the leaching. The output of the model simulation is the quantity of herbicides arriving daily to the stream and its concentration at the outlets. The originality of the model is the representation of water and pesticide runoffs with tree structures where leaves and roots are respectively up-streams and down-streams of the catchment.

The software allows the user to see the relationships between these tree structures and the rules learnt from simulations. A more elaborated version allows to launch simulations and to learn rules on-line. This year, we have developed this new version by enabling access to two recommendation action algorithms (see section 6.3.5). The user can choose different parameters (set of classification rules from which actions will be built, parameters concerning action feasibility, etc) before asking for action recommending process, and then easily visualize the characteristics of situations to improve (polluted ones) compared with the different recommended actions. The software is mainly in Java.

The following website is devoted to the presentation of the SACADEAU: http://www.irisa.fr/dream/SACADEAU/.

5.4. Ecomata

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

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

6. New Results

6.1. Diagnosis of large scale discrete event systems

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

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

More recently, we started research on interacting with large-scale systems in a decision-oriented way. Scenario patterns were defined for exploring complex systems, based on the use of model-checking techniques.

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

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

6.1.2. Scenario patterns for exploring qualitative ecosystems

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

The approach has been experimented on a marine ecosystem under fishing pressure. The model describes the trophodynamic interactions between fish trophic groups as well as interactions with the fishery activities and with an environmental context. A paper has been accepted for publication by the Environmental Modelling Software Journal [4].

We extended the approach to deal with “How to” queries. As before, we rely on a qualitative model in the form of timed automata and use model-checking tools to answer queries. We have recently proposed two approaches to answer questions such as “How to avoid a given situation ?” (safety query). The first one exploits controller synthesis and the second one is a “generate and test” approach. We compared these two approaches in the context of an application that motivates this work, i.e the management of a marine ecosystem and the evaluation of fishery management policies. The results have been accepted for publication in [14].

6.2. Machine learning for model acquisition

Participants: Thomas Guyet, René Quiniou.

Model acquisition is an important issue for model-based diagnosis, especially as modeling dynamic systems. We investigate machine learning methods for temporal data recorded by sensors or spatial data resulting from simulation processes. We also investigate efficient methods for storing and accessing large volume of simulations data. Our main interest is extracting knowledge, especially sequential and temporal patterns or prediction rules, from static or dynamic data (data streams). We are particularly interested in mining temporal patterns with numerical information and in incremental mining from sequences recorded by sensors.

6.2.1. Mining temporal patterns with numerical information

We are interested in mining interval-based temporal patterns from event sequences where each event is associated with a type and time interval. Temporal patterns are sets of constrained interval-based events. This year, we have been working on improving the formal setting of the approach as well as its efficiency [8]. We have introduced the notion of ϵ-covering of temporal patterns over sequences to cope with the dual nature, symbolic and numerical, of temporal patterns. The parameter ϵ specifies the tightness of the similarity used for matching patterns and sequences. It complements the parameter σ representing the minimal support which is used to prune candidate patterns. The ϵ-similar occurrences of some pattern, precisely their associated temporal intervals, are classified to characterize the different classes of numerical temporal intervals that correspond to different patterns sharing the same symbolic part. This process have been embedded in two sequential pattern mining algorithms, GSP and PrefixSpan, and we have compared their performance.

6.2.2. Incremental sequential mining

We investigate the problem of mining and maintaining frequent sequences in a window sliding on a stream of itemsets. We propose in [11] a complete and correct incremental algorithm based on a tree representation of frequent sequences inspired by PSP [52] and a method for counting the minimal occurrences of a sequence. Instead of the frequence, to a node representing a pattern is associated the set of occurrences of this pattern. The algorithm updates efficiently the tree representation of frequent sequences and their occurrences by means of two operations on the tree: deletion of the itemset at the beginning of the window (obsolete data) and addition of an itemset at the end of the window (new data). Experiments were conducted on simulated data and on real data of instantaneous power consumption.

6.2.3. Multiscale segmentation of satellite image time series

Satellite images allow the acquisition of large-scale ground vegetation. Images are available along several years with a high acquisition frequency (1 image every two weeks). Such data are called satellite image time series (SITS). In [9], we present a method to segment an image through the characterization of the evolution of a vegetation index (NDVI) on two scales: annual and multi-year. We test this method to segment Senegal SITS and compare our method to a direct classification of time series. The results show that our method using two time scales better differentiates regions in the median zone of Senegal and locates fine interesting areas (cities, forests, agricultural areas).

6.3. Decision aiding with models and simulation data

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

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

6.3.1. Exploring models thanks to scenarios: a generic framework

In the framework of the APPEAU project (see 8.2.1), that ended in December 2010, a paper, describing a generic framework for scenario exercises using models applied to water-resource management, has been written during 2011 in cooperation with all the partners and submitted to *Environmental Modelling and Software*. It is currently under revision.

6.3.2. A datawarehouse for simulation data

The ACASSYA project aims at providing experts or stakeholders or farmers with a tool to evaluate the impact of agricultural practices on water quality. As the simulations of the deep model TNT2 are time-consuming and generate huge data, we have proposed to store these simulation results in a datawarehouse and to extract relevant information, such as prediction rules, from the stored data. We have devised a general architecture for agro-environmental data on top of the framework Pentaho. An article presenting the principles of this architecture as well as a set of realistic scenarios and their transformation into OLAP queries has been submitted to Compag (Computers and Electronics in Agriculture).

6.3.3. Efficient computation of skyline queries in an interactive context

Skyline queries retrieve from a database the objects that optimizes multiple criteria, related to user preferences for example, or objects that are the best compromises satisfying these criteria. When data are huge such objects may shed light on interesting parts of the dataset. However, computing the skylines (i.e. retrieving the skyline points) may be time consuming because of many dominance tests. This is, especially the case in an interactive setting such as querying a data cube in the context of a datawarehouse. We have worked on how to answer efficiently to skyline queries by the materialization of precomputed skyline queries related to dynamic user preferences. An article has been submitted to the conference SIGMOD 2012.

6.3.4. Influence Diagrams for Multi-Criteria Decision-Making

For multi-criteria decision-making problems, we propose in [6] a model based on influence diagrams able to handle uncertainty, represent interdependencies among the different decision variables and facilitate communication between the decision-maker and the analyst. The model makes it possible to take into account the alternatives described by an attribute set, the decision-maker's characteristics and preferences, and other information (e.g., internal or external factors) that influence the decision. Modeling the decision problem in terms of influence diagrams requires a lot of work to gather expert knowledge. However, once the model is built, it can be easily and efficiently used for different instances of the decision problem. In fact, using our model simply requires entering some basic information, such as the values of internal or external factors and the decision-maker's characteristics.

6.3.5. Recommending actions from classification rules

In the framework of the SACADEAU project (see 8.2.1), a paper dedicated to building recommendation actions for a given situation, from the set of classification rules, learnt from simulation results, has been published in the KAIS journal [7].

6.4. Causal reasoning and influence diagrams

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

This work stems on [23], [24], [25], [26], [27] and, for the logic programming translation, on [53], [54]. It is related to diagnosis (observed symptoms explained by faults).

The previously existing proposals were ad-hoc or, as in [29], [41], they were too close to standard logic in order to make a satisfactory diagnosis. Our proposal starts from a restricted first order logic (of the Datalog kind: no function symbols) and introduces *causal formulas*, built on causal atoms such as (α *causes* β) intended to mean: "α causes β". The system is described thanks to these causal formulas, classical formulas, and taxonomy atoms such as (α IS_A β) (α is of kind β).

The system produces *explanation atoms* of the kind (α *explains* β *if_possible* $\{\gamma_1, \cdots, \gamma_n\}$), meaning that β can be explained by α if all the γ_i's are possible together in the context of the given data.

This year, we have improved our logic programming translation in ASP. The aim is to improve efficiency and also reduce the work of the programmer, taking advantage of the declarative aspect of this type of programming. We have applied some of these improvements to two classic riddles, in order to illustrate the power and limitations of current answer set programming systems, and we proposed a few improvements which could make the present systems yet easier to use [12], [13].

We are starting a work with some similarities to automatize the treatment of cognitive maps. The aim is to extract relevant information from these maps, which means: building a graph formalism for representing mixed causal and influence relations, and defining a framework (argumentation theory is a good candidate) to aggregate the graphs and provide inference rules in order to infer new information and relations. This work is done in the framework of the RADE2BREST project, involving Agrocampus Ouest and CNRS (GEOMER/LETG), funded by "Ministère de l'Ecologie"[1]. The goal of this project is to model shellfish fishing

[1] This project is not mentioned in section 8.1 because DREAM is not a partner of this project.

in order to assess the impact of management pollution scenarios on the Rade de Brest. The cognitive maps result from interviews with fishermen.

7. Contracts and Grants with Industry

7.1. ManageYourSelf: diagnosis and monitoring of embedded platforms

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

ManageYourSelf is a project that deals with the diagnosis and monitoring of embedded platforms, in the framework of a collaboration with Telelogos, a French company expert in mobile management and data synchronization. ManageYourSelf aims to perform diagnostic and repair on a fleet of mobile smartphones and PDAs. The idea is to embed on the mobile devices a rule-based expert system and its set of politics, for example “if memory is full then delete (directory)”. recognition is performed, using the parameters of the phones as the fact base. Of course, it is impossible to foresee all the rules in advance. Upon detection of a non anticipated problem, a report containing all the system’s information prior to the problem is sent to a server and a learning step, using decision trees, aims at updating the global knowledge base and its distributed instances. We are currently working on an incremental version of this learning step. Another current issue is to deal with a dynamic set of attributes. The following website gives many details: http://www.irisa.fr/dream/ManageYourself.

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. CLIMASTER: Changement climatique, systèmes agricoles, ressources naturelles et développement territorial

Participant: Marie-Odile Cordier.

We are involved in a PSDR GO (research program "Pour et Sur le Développement Régional Grand Ouest") project, named CLIMASTER that started in 2009 and will end in november 2011. The goal of the project is to investigate and characterize climate change in the west of France as well as determining the trends, the extrema and the variability with respect to spatial location. For detailed information about the project consult http://www.irisa.fr/dream/PDF/PSDRGO_CLIMASTER-2012.pdf or the CLIMASTER website: http://www.rennes.inra.fr/climaster/.

8.2. National Initiatives

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

Participants: Marie-Odile Cordier, Véronique Masson, Christine Largouët.

The SACADEAU project (Système d’Acquisition de Connaissances pour l’Aide à la Décision pour la qualité de l’EAU - Knowledge Acquisition System for Decision-Aid to Improve Streamwater Quality) was funded by INRA (French institute for agronomy research) from October 2002 to October 2005. The main partners were from INRA (SAS from Rennes and BIA from Toulouse) and from IRISA. We have continued to develop the SACADEAU model with our partners until now and a PdD thesis has been funded by INRA (ASC).

We were then involved in a project, named APPEAU and funded by ANR/ADD, which started in February 2007 and ended in December 2010. The APPEAU project aimed at studying which politics, for which agronomic systems, are best adapted to improve water management. It included our previous partners as well as new ones, mainly from INRA. A paper has been written in 2011 in cooperation with all the partners and submitted to *Environmental Modelling and Software*, currently in revision.

Our work aims at building a decision-aid tool to help specialists in charge of the catchment management to preserve the streamwater quality. The SACADEAU simulation model couples two qualitative models, a transfer model describing the pesticide transfer through the catchment and a management model describing the farmer decisions. The simulation results are analyzed, thanks to classification and symbolic learning techniques, in order to discover rules explaining the pesticide concentration in the stream by the climate, the farmer strategy, the catchment topology, etc., and, finally, in order to build recommendation actions for a given situation. In the APPEAU context, the idea is to study how this kind of model can be used to simulate scenarios in a more generic way and to compare, and possibly unify, our work with what is done by our partners from SAS concerning nitrate transfer. (http://wwwagir.toulouse.inra.fr/agir/index.php?option=com_content&view=article&id=62&Itemid=134)

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

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

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

8.2.3. PayOTe-II: characterizing agricultural landscapes via data mining

Participants: Thomas Guyet, Christine Largouët, René Quiniou.

The PAYOTE-II project (Paysage Ou Territoire) is funded by AIP INRA/INRIA and started at the end of 2010. The project associates INRIA Teams (Orpailleur and Dream) with INRA Team (UBIA, MIAJ and SAD-Paysage).

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

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

8.2.4. PAYTAL: Mining spatial correlations between urban sprawl and landscape

Participant: Thomas Guyet.

The PAYTAL project (Paysage et Etalement Urbain - Landscape and urban sprawl) is funded for 3 years by the french Ministry of Ecology and Sustainable Developement. This project started in september 2011. It involves our colleagues from INRA/SAS and Agrocampus-Ouest.

This project proposes a multidisciplinary approach, firstly, to describe the fine forms of urban sprawl and the dynamics of the landscape and, secondly, to study the links between urban sprawl and landscape evolution. The first aim of the project is to develop a methodology to acquire a spatial description of both the landscape and the urban extent. This spatial information will be acquired by our colleagues from remote sensing images and available official documents (local development plans, landscape repository, etc.) related to several "conurbations" in western France (Rennes, Angers, Lorient, Brest).

Our work aims at using and proposing spatio-temporal data mining tools to extract landscape patterns, i.e. a set of landscape elements linked through spatial relationships. Using symbolic learning techniques, we expect to extract landscape spatial patterns that may explain the urban sprawl (evolution barrier or facilitator).

9. Dissemination

9.1. Animation of 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).
- *Revue JAIR* (M.-O. Cordier).
- *Revue JESA: Journal Européen des Systèmes automatisés* (M.-O. Cordier).

9.1.2. Conference program committees and organizations

- Program committee member RFIA'2012 (M.-O. Cordier and T. Guyet).
- President of the organization committee of JIGOT 2011 (T. Guyet).
- Program committee member EGC 2011 (R. Quiniou).

9.1.3. Scientific and administrative boards

- ECCAI fellow + Honorific member of AFIA (Association Française d'Intelligence Artificielle): M.-O. Cordier
- Member of the AFIA board: T. Guyet.
- Member of "Agrocampus-Ouest" scientific board: M.-O. Cordier.
- Member of "Conseil d'administration de l'ISTIC", "Comité scientifique de l'ISTIC", "Direction scientifique de l'Irisa": M.-O. Cordier
- Member of the selection committee for the "Prix de thèse AFIA 2011" (selecting the best French PhD thesis in the Artificial Intelligence domain): M.-O. Cordier.
- Chair of the INRA CSS-MBIA (Commission scientifique spécialisée "Mathématiques, Biologie et Intelligence Artificielle": M.-O. Cordier.
- Member of the COREGE (Research Committee-COmité de la REcherche du Grand Etablissement) of Agrocampus-Ouest: T. Guyet.
- External evaluator for the portuguese FCT (Portuguese Foundation For Science and Technology): T. Guyet.

9.2. Teaching

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

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

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

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

PhD in progress: Tassadit Bouadi, “Analyse interactive de résultats de simulation et découverte de connaissances. Application à l’aide à la décision dans le domaine agroécologique pour l’amélioration de la qualité des eaux des bassins versants”, january 1st 2010, co-supervisors Marie-Odile Cordier, René Quiniou and Chantal Gascuel, ANR project Acassya grant

PhD in progress: Yulong Zhao, “Modélisation d’agroécosystèmes dans un formalisme de type systèmes à événements discrets et simulation de scénarios utilisant des outils de model-checking. Application à l’étude des impacts des changements climatiques et des pratiques agricoles sur les flux de nutriments vers les eaux de surface.”, october 1st 2010, supervisor Marie-Odile Cordier and Chantal Gascuel

PhD in progress: Philippe Rannou [2], “Modèle rationnel pour humanoïdes virtuels”, october 1st 2010, co-supervisors Marie-Odile Cordier and Fabrice Lamarche

10. Bibliography

Major publications by the team in recent years

[1] P. BESNARD, M.-O. CORDIER, Y. MOINARD. *Ontology-based inference for causal explanation*, in "Integrated Computer-Aided Engineering", 2008, vol. 15, n° 4, p. 351-367, http://hal.inria.fr/inria-00476906/en/.

[2] C. GASCUEL ODOUX, P. AUROUSSEAU, M.-O. CORDIER, P. DURAND, F. GARCIA, V. MASSON, J. SALMON-MONVIOLA, F. TORTRAT, R. TRÉPOS. *A decision-oriented model to evaluate the effect of land use and agricultural management on herbicide contamination in stream water*, in "Environmental modelling & software", 2009, vol. 24, p. 1433-1446, http://hal.inria.fr/hal-00544122/en.

[3] Y. PENCOLÉ, M.-O. CORDIER. *A formal framework for the decentralised diagnosis of large scale discrete event systems and its application to telecommunication networks*, in "Artificial Intelligence Journal", 2005, vol. 164, n° 1-2, p. 121-170, http://hal.inria.fr/inria-00511104/en/.

Publications of the year

Articles in International Peer-Reviewed Journal

[4] C. LARGOUËT, M.-O. CORDIER, Y.-M. LE BOZEC, Y. ZHAO, G. FONTENELLE. *Use Of Timed Automata And Model-Checking To Explore Scenarios On Ecosystem Models*, in "Environmental Modelling and Software", November 2011 [*DOI :* 10.1016/J.ENVSOFT.2011.08.005], http://hal.inria.fr/hal-00649275/en/.

[5] J. SALMON-MONVIOLA, C. GASCUEL-ODOUX, F. GARCIA, F. TORTRAT, M.-O. CORDIER, V. MASSON, R. TRÉPOS. *Simulating the effect of techniques and environmental constraints on the spatio-temporal distribution of herbicide applications and stream losses*, in "Agriculture, Ecosystems and Environment", January 2011, vol. 140, p. 382-394, http://hal.inria.fr/hal-00618448/en.

[6] K. SEDKI, V. DELCROIX. *A Model based on Influence Diagrams for Multi-Criteria Decision-Making*, in "International Journal on Artificial Intelligence Tools (IJAIT)", 2011.

[7] R. TRÉPOS, A. SALLEB-AOUISSI, M.-O. CORDIER, V. MASSON, C. GASCUEL-ODOUX. *Building Actions From Classification Rules*, in "Knowledge and Information Systems (KAIS) journal", June 2011, http://hal.inria.fr/hal-00649388/en/.

[2] Philippe Rannou is not a member of DREAM but a member of the IRISA research group MIMETIC

International Conferences with Proceedings

[8] T. GUYET, R. QUINIOU. *Extracting temporal patterns from interval-based sequences*, in "International Joint Conference on Artificial Intelligence (IJCAI)", Barcelone, Spain, July 2011, http://hal.inria.fr/inria-00618444/en.

National Conferences with Proceeding

[9] T. GUYET, H. NICOLAS, A. DIOUCK. *Segmentation multi-échelle de séries temporelles d'images satellite : Application à l'étude d'une période de sécheresse au Sénégal.*, in "Reconnaissance de Forme et Intelligence Artificielle (RFIA)", Lyon, France, January 2012, http://hal.inria.fr/hal-00646158/en.

[10] T. GUYET, R. QUINIOU. *Extraction de motifs temporels à partir de séquences d'événements avec intervalles temporels*, in "Extraction et Gestion de Connaissances", Brest, France, January 2011, http://hal.inria.fr/inria-00547629/en.

[11] T. GUYET, R. QUINIOU. *Extraction incrémentale de séquences fréquentes dans un flux d'itemsets*, in "Extraction et Gestion de Connaissances (EGC'2012)", Bordeaux, France, B. PINAUD, G. MELANÇON, Y. LECHEVALLIER (editors), RNTI, Hermann, February 2012, http://hal.inria.fr/hal-00648893/en/.

[12] Y. MOINARD. *Utilisation de la programmation par ensembles réponses (Answer Set Programming) sur de "petits" problèmes*, in "Cinquièmes Journées de l'Intelligence Artificielle Fondamentale (JIAF)", Lyon, France, L. CHOLVY, ET AL. (editors), LIRIS, June 2011, http://hal.inria.fr/inria-00619527/en.

[13] Y. MOINARD. *Utiliser la programmation par ensembles réponses pour de petits problèmes*, in "Reconnaissance de Forme et Intelligence Artificielle (RFIA)", Lyon, France, January 2012.

[14] Y. ZHAO, M.-O. CORDIER, C. LARGOUËT. *Répondre aux questions "Que faire pour" par synthèse de contrôleur sur des automates temporisés - Application à la gestion de la pêche*, in "Reconnaissance des Formes et Intelligence Artificielle (RFIA)", Lyon, France, January 2012, http://hal.inria.fr/hal-00649182/en/.

References in notes

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

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

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

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

[19] R. AGRAWAL, T. IMIELINSKI, A. N. SWAMI. *Mining Association Rules between Sets of Items in Large Databases*, in "Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data", Washington, D.C., P. BUNEMAN, S. JAJODIA (editors), 26–28 1993, p. 207–216.

[20] R. AGRAWAL, R. SRIKANT. *Mining sequential patterns*, in "Eleventh International Conference on Data Engineering", Taipei, Taiwan, P. S. YU, A. S. P. CHEN (editors), IEEE Computer Society Press, 1995, p. 3–14.

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

[22] M. BASSEVILLE, M.-O. CORDIER. *Surveillance et diagnostic de systèmes dynamiques : approches complémentaires du traitement de signal et de l'intelligence artificielle*, Irisa, 1996, n^{o} 1004, ftp://ftp.irisa.fr/techreports/1996/PI-1004.ps.gz.

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

[24] P. BESNARD, M.-O. CORDIER, Y. MOINARD. *Configurations for Inference between Causal Statements*, in "KSEM 2006 (First Int. Conf. on Knowledge Science, Engineering and Management)", J. LANG, F. LIN, J. WANG (editors), LNAI, Springer, aug 2006, n^{o} 4092, p. 292–304, http://www.irisa.fr/dream/dataFiles/moinard/causeksempub.pdf.

[25] P. BESNARD, M.-O. CORDIER, Y. MOINARD. *Ontology-based inference for causal explanation*, in "KSEM07 (Second International Conference on Knowledge Science, Engineering and Management)", Z. ZHANG, J. SIEKMANN (editors), LNAI, Springer, nov 2007, n^{o} 4798, p. 153-164, http://www.irisa.fr/dream/dataFiles/moinard/ksem07bcausesonto.pdf.

[26] P. BESNARD, M.-O. CORDIER, Y. MOINARD. *Deriving explanations from causal information*, in "ECAI 2008 (18th European Conference on Artificial Intelligence)", Patras, Greece, M. GHALLAB, C. D. SPYTOPOULOS, N. FAKOTAKIS, N. AVOURIS (editors), IOS Press, jul 2008, p. 723–724.

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

[28] C. BIERNACKI, G. CELEUX, G. GOVAERT, F. LANGROGNET. *Model-Based Cluster and Discriminant Analysis with the MIXMOD Software*, in "Computational Statistics and Data Analysis", 2006, vol. 51, n^{o} 2, p. 587-600.

[29] 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.

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

[31] M.-O. CORDIER, Y. PENCOLÉ, L. TRAVÉ-MASSUYÈS, T. VIDAL. *Characterizing and checking self-healibility*, in "ECAI 2008 (18th European Conference on Artificial Intelligence)", Patras, Greece, M. GHALLAB, C. D. SPYTOPOULOS, N. FAKOTAKIS, N. AVOURIS (editors), IOS Press, 2008, p. 789–790.

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

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

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

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

[36] C. DOUSSON, T. V. DUONG. *Discovering Chronicles with Numerical Time Constraints from Alarm Logs for Monitoring Dynamic Systems.*, in "IJCAI", T. DEAN (editor), Morgan Kaufmann, 1999, p. 620-626.

[37] 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.

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

[39] J. GAO, B. DING, W. FAN, J. HAN, P. S. YU. *Classifying Data Streams with Skewed Class Distributions and Concept Drifts*, in "IEEE Internet Computing", 2008, vol. 12, n° 6, p. 37-49.

[40] M. GAROFALAKIS, J. GEHRKE, R. RASTOGI. *Querying and Mining Data Streams: You Only Get One Look. Tutorial notes*, in "ACM Int. Conf. on Management of Data", 2002.

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

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

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

[44] T. IMIELINSKI, H. MANNILA. *A database perspective on knowledge discovery*, in "Comm. of The ACM", 1996, vol. 39, p. 58–64, http://courses.cs.ut.ee/datamining/2003/dm-seminar-fall/uploads/Main/arm-1.pdf.

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

[46] X. LE GUILLOU, M.-O. CORDIER, S. ROBIN, L. ROZÉ. *Chronicles for On-line Diagnosis of Distributed Systems*, IRISA, Rennes, France, may 2008, n° 1890.

[47] X. LE GUILLOU, M.-O. CORDIER, S. ROBIN, L. ROZÉ. *Chronicles for On-line Diagnosis of Distributed Systems*, in "ECAI'08 (18th European Conference on Artificial Intelligence)", Patras, Greece, M. GHALLAB, C. D. SPYTOPOULOS, N. FAKOTAKIS, N. AVOURIS (editors), IOS Press, july 2008.

[48] X. LE GUILLOU. *Chronicles for On-line Diagnosis of Distributed Systems*, in "CAISE-dc'08 (Conference on Advanced Information Systems Engineering - doctoral consortium)", Montpellier, France, june 2008.

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

[50] H. MANNILA, H. TOIVONEN, A. I. VERKAMO. *Discovery of Frequent Episodes in Event Sequences*, in "Data Mining and Knowledge Discovery", 1997, vol. 1, n° 3, p. 259–289.

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

[52] F. MASSEGLIA, F. CATHALA, P. PONCELET. *The PSP Approach for Mining Sequential Patterns*, in "Proceedings of the Second European Symposium on Principles of Data Mining and Knowledge Discovery", 1998, p. 176–184.

[53] Y. MOINARD. *A formalism for causal explanations with an Answer Set Programming translation*, in "4th International Conference on Knowledge Science, Engineering & Management (KSEM 2010)", United Kingdom Belfast, B. SCOTNEY, Z. JIN (editors), LNCS, Springer-Verlag, Aug 2010, http://hal.inria.fr/inria-00511093/en.

[54] Y. MOINARD. *Using ASP with recent extensions for causal explanations*, in "ASPOCP10, Answer Set Programming and Other Computing Paradigms Workshop, associated with ICLP", United Kingdom Edinburgh, M. BALDUCCINI, S. WOLTRAN (editors), Jul 2010, http://hal.inria.fr/inria-00542880/en.

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

[56] 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.

[57] R. QUINIOU, M.-O. CORDIER, G. CARRAULT, F. WANG. *Application of ILP to cardiac arrhythmia characterization for chronicle recognition*, in "ILP'2001", C. ROUVEIROL, M. SEBAG (editors), LNAI, Springer-Verlag, 2001, vol. 2157, p. 220-227, http://www.irisa.fr/dream/dataFiles/quiniou/ilp01.pdf.

[58] 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.

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

[60] M. SAMPATH, R. SENGUPTA, S. LAFORTUNE, K. SINNAMOHIDEEN, D. TENEKETZIS. *Diagnosability of discrete event systems*, in "Proceedings of the International Conference on Analysis and Optimization of Systems", 1995, vol. 40, p. 1555-1575.

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

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

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

[64] A. VAUTIER, M.-O. CORDIER, R. QUINIOU. *An Inductive Database for Mining Temporal Patterns in Event Sequences (short version)*, in "Proceedings of IJCAI-05 (International Joint Conference on Artificial Intelligence)", Edinburgh, L. P. KAELBLING, A. SAFFIOTTI (editors), 2005, p. 1640-1641, Poster.

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

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