Data, Information and Knowledge for Medical Scenario Construction

Anne-Sophie Silvent
Catherine Garbay

Lab. TIMC, Faculté de Médecine, 38706 La Tronche, France

sophie.silvent@imag.fr
Catherine.Garbay@imag.fr

Pierre-Yves Carry

DAR CHU Lyon-Sud, F 69495 Pierre Bénite Cedex, France

Michel Dojat

INSERM-UJF U594, CHU, 38043 Grenoble Cedex 9, France

Michel.Dojat@ujf-grenoble.fr

Abstract

The automatic recognition of typical pattern sequences (scenarios), as they are developing, is of crucial importance for computer-aided patient supervision. However, the construction of such scenarios directly from medical expertise is unrealistic in practice. Starting from the monitored data and clinical information available, our objective is to extract typical abstracted pattern sequences and then construct scenarios validated by clinical experts as representative of a class of situations to recognize. In this paper, we present a methodology for data abstraction, based on the management of data, information and knowledge, for the extraction of specific events and eventually the construction of such scenarios.

1 Introduction

The automatic recognition of typical pattern sequences (scenarios) is of crucial importance for computer-aided patient supervision. As the construction of scenarios directly from medical expertise is unrealistic in practice, our objective is to extract typical abstracted pattern sequences from the monitored data and clinical information available. Starting from these sequences, we want to construct scenarios, eventually validated by clinical experts as representative of a class of situations to recognize. This construction needs different step and we interest in the first, progressive data abstraction. We insist on the importance of an implicit management of information and knowledge during the process.

The clinical context is the weaning of patients from mechanical ventilation. Weaning is the period when the clinician tries to gradually withdraw the mechanical assistance. This is a crucial period. A too fast decrease of the assistance can cause an excessive fatigue for the respiratory muscles and an additional stress for the patient. On the other side, the maintain under mechanical ventilation has negative physiological, psychological and economical impacts. Clinical data used in this study were acquired based on a specific weaning protocol in three parts. During one hour the patient is classically ventilated. During two hours the patient breathes by himself on a T-piece and then on spontaneous ventilation. At any time during the last two periods, the patient can be replaced on artificial ventilation. For each step, physiological data are recorded and a medical assistant annotates all specific alarming situations. This procedure lasts four hours even in case of weaning failure. At the end, weaning is considered as a success if the patient can breathe without assistance during 24 hours. This protocol was approved by our local ethical committee. Only patients who fulfil specific pre-weaning criteria are included into the protocol.

Based on the weaning protocol, the automatic extraction of specific patterns can lead to 1) a better understanding of the physiological effects due to weaning, and 2) the prediction of weaning failure or success by recognition of specific scenarios as they are developing. In the latter case, the total duration of the weaning could be reduced. This work is part of a global French project called OSSCAR for the optimisation of medical strategies from cardio-respiratory signals in intensive care and anaesthesia [3].

2 Abstraction role in knowledge discovery process

According to researchers such as Fayyad [6], the process of knowledge discovery via data mining can be divided into five basic activities; selection, pre-processing, transformation, data mining, and interpretation, which are necessary to extract useful knowledge from data. Selection allows to focus, according to pre-defined criterions, on data supposed to be "interpretable" and with informative potential. Pre-processing concerns treatment of missing and noisy data. Transformation is useful to define representations and/or data abstractions adapted to knowledge discovery task. Data mining allows to find patterns and interpretation select the more relevant and extract knowledge. The global process is incremental. According to Fayyad, it can contain loops between 2 steps and, in particular, interpretation can guide all the others. Here, we study the 3 first steps that we call "abstraction".In this process, we insist on the importance of three concepts: data, information and knowledge.

These concepts are often confused in the literature. Here, we consider **Data** as the result of an observation or an experiment; **Information** as the result of data abstraction and **Knowledge** defines the way by which data and information are handled [11]. Knowledge can be considered as elements not useful to describe the situation, but as essential elements to construct new information or determine actions to be undertaken [17].

We propose (see Fig. 1) a modified version of the process described by Fayyad, to point up the information concept and the central role of knowledge in the incremental and dynamic process of abstraction in knowledge discovery. Abstraction is a complex process: knowledge is required to transform data into information, and this, as early as pre-processing step [18]. At each step, new information is produced that allows in turn the control of abstraction. Abstraction resumes available data via the creation of information. Knowledge drives the abstraction process. It may emerge from mining process and in turn is used to drive it.

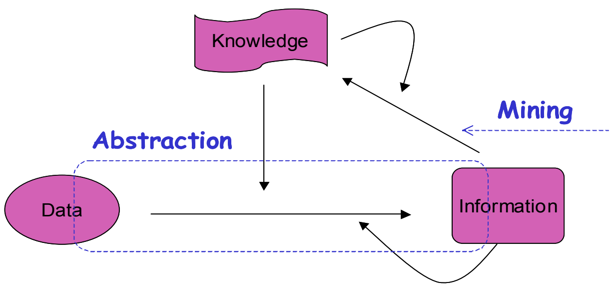


Figure 1: The abstraction and mining processes

3 Our ontology

We define an ontology to refine and instantiate in the medical domain the three concepts: Data, Information and Knowledge.

Data are the observed values of a parameter for instance tidal volume and oxygen saturation.

Information (see Fig. 2) can be supplied or extracted by the process abstraction. For systems supervision, they are essentially represented by temporal objects as states and events [5]:

- An event is classically an instantaneous change of the system. In our problematic, we call event all change or intervention on the system or its environment (examples : settings on devices, cares to patient,...). For the following, we discern 3 types of events: 1) inner events (changes of a parameter values, disconnection from the ventilator), outer events (example: cares) and alarms (overstepping of a threshold for a parameter). Outer events and alarms are provided to us and annotated by the observer in charge during data acquisition process. Causal links exist between events. Local patterns are used for the recognition of specific inner events (see below).
- A state involves the notion of duration and so, persistence. It represents a time period during which an information is valid (example: heart rate steady during 5 minutes). States are generated by the abstraction process. States and events are temporally linked.
- A scenario is the representation of the temporal links between states and events. It can be represented using temporal constraints graph [5].

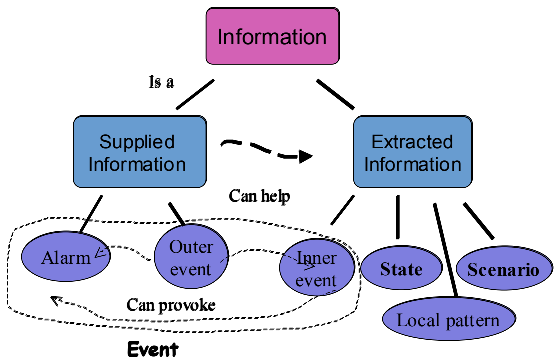


Figure 2: Information ontology

Knowledge: Similarly to information knowledge can be supplied or extracted (see Fig. 3). *A priori* knowledge is essentially in the form of classification rules and causal relations between parameters and is context dependent. Extracted knowledge allows extraction of typical patterns sequences that are eventually used to define scenarios representative of typical situations to recognise.

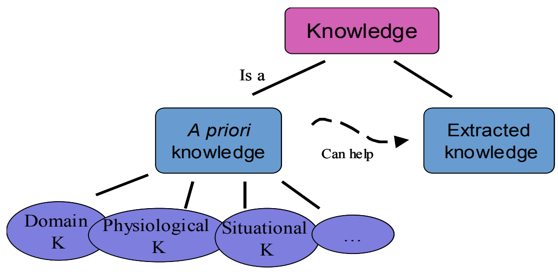


Figure 3: Knowledge ontology

4 Information and knowledge handling through data abstraction

Data abstraction is a knowledge-dependent process that builds abstractions in the domain parameters space at several levels of temporal granularity [18]. Fusion of knowledge and information is central in our approach.
We discern simple and complex temporal abstractions [13]: 1) simple abstraction represents a symbolic data descriptor on a time interval that characterises for instance a state (as stable or unstable), or a trend (as in-

creasing or decreasing), it is constructed directly from raw data; 2) complex abstraction represents temporal relations between intervals constructed from abstractions (simple or complex) [13]. It is also called local patterns and are used to characterize specific events such as disconnection from the ventilator for instance.

Several systems have been proposed for medical data abstraction that can be distinguished by the way they use knowledge and information. In [2], the authors propose a methodology for extraction of local trends from a stream of ICU data. Standard deviation gives a local stability index. Qualitative notions such as "increasing" and "decreasing" or "steady" and "unstable", are obtained directly by trend versus stability plan partitioning. The partitioning requires the introduction of *a priori* knowledge to fix thresholds. Based on the definition of typical trends templates, TrenDx [7] allows their automatic recognition in monitoring patient data. In VIE-VENT [16], an open loop system for the supervision of neonatal mechanical ventilation, knowledge and extracted information are combined. Only Shahar [18] has studied in depth the implication of knowledge in temporal abstraction mechanism.

Data abstraction objective is 1) to transform numerical values, instant-point based, into symbolic values, time-interval based and 2) to generate several abstracted levels that synthesizes parameters evolution. In our point of view, this process should be carefully controlled to bring reliable information in an efficient way: a low abstraction level, close to the level of data, is too informative to be efficiently used, whereas a high level can lead to an irreversible loss of informative data. Moreover, the abstraction process, to be valid, has to rely on reliable parameter values. It may happen in fact that due to noise in the acquisition process, or external disturbances, the parameter shows erratic and non significant values. We therefore restrict the abstraction process to operate on so-called "valid abstraction domain", eg domains where the considered parameter behaviour is assessed. Our target is to create a reliable abstraction system. This reliability is ensured by:

- modelling the abstraction process as the incremental building of new information;
- bringing knowledge early to control the abstraction process and its domain of validity;
- interacting early with clinical experts to validate each step.

In practice, our data abstraction methodology relies on 4 sub-steps for each step of our incremental process:
Step 1. Definition of the parameters on which abstraction should operate, the way to describe them and the abstraction methods (mean calculation, merging time intervals, etc) to be used ;
Step 2. Identification of the *a priori* knowledge and/or extracted information necessary to drive the abstraction;
Step 3. Definition of the domain where abstraction is valid (abstraction validity domains), in terms of constraints operating on raw data, abstracted data and knowledge;
Step 4. Control by the domain expert of the abstraction process stage by stage via a graphical interface. Note that the introduction of new information contributes to the explicitation of the expertise.

5 Data abstraction at work

Our data abstraction methodology was applied in the context of weaning. Several elements were at our disposal:
Raw physiological data: 7physiological signals were sampled at 100Hz: ECG, heart rate, plethismography, pulse oxymetry, invasive pressure, airway flow and airway-pressure. From these signals, we extracted several numerical parameters, sampled at one Hertz, such as heart rate, respiratory rate, tidal volume, airway pressure maximum, and oxygen saturation.
Extracted information: In order to qualify parameter variation and therefore to better control the data abstraction process, symbolic trends were introduced for each parameter. Trend were computed using a linear regression on a moving temporal window, called "characteristic span" whose the size was determined according to the dynamics of the considered parameter [2]. When computed in such a way, trend is comparable to a derivative. It is then possible to reconstruct filtered data by integration of the trend values. At any time and for a given characteristic span, the value of the parameter and its trend were provided. The standard deviation was used as an index of stability. Following [2], trends and standard deviation were used to qualitatively partition the parameters space and to define the relevant categories: increasing, decreasing, steady and unstable. "Unstable" corresponds to a high standard deviation reflecting extremely noisy signals. "Steady", "increasing" and "decreasing" correspond to a low standard deviation (stable condition) and to a regression coefficient respectively close to zero, positive and close to 1, negative and close to –1 [3].
***A priori* knowledge**: Three kind of *a priori* knowledge were available: 1) Situational knowledge (knowledge about occurring events): at the patient's bedside, an observer qualified each device alarm in terms of false positive, false negative, true positive and true negative, and annotated events occurrences; 2) Physiological (causal knowledge) consists in temporal or causal relations between physiological parameters; 3) Clinical knowledge is used to rely numerical and symbolical values ("normal" or "abnormal" state for instance).

5.1 Simple abstraction for state construction

The objective of simple abstraction process is to obtain parameter state as for example "HR stable for 30 minutes" in the order to get closer to clinician discourse. Starting from physiological parameters and trends, we can develop new abstractions. We restrict the abstraction process validity domain to domains where the parameter behavior is considered as valid. Validity domains are based on the notion of stability: if a parameter is marked as unstable, we consider its numerical values as no valid. Several steps have been necessary to determine validity domain. For each step, parameter abstraction uses different types of knowledge and information. Each step has been validated by the clinician and is described based on the previously proposed methodology.

1) Parameters labeling in terms of stable or unstable

This step qualifies each measured value in term of stable or unstable. Stable is refined in "steady", "increasing" or "decreasing" (see Fig. 4).

Step 1.
Elements : raw measured values.
Description : raw values.
Operation : the counting of the number of moving windows, used to assess stability, that contain a given measured value.

Step 2.
Elements: extracted information: trend, characteristic span.
A priori knowledge: a measured value is considered as stable if it is qualified more times as stable than as unstable

Step 3.
The domain validity abstraction is a succession of measured values considered as stable.

Step 4.
The obtained results are displayed and validated by the clinician.

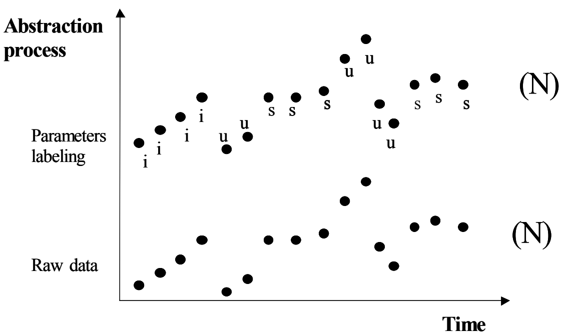


Figure 4 : Abstraction process (1/4)
i : increasing, s : steady, u : unstable, N : numerical

Technical details:
Because trend is calculated on a moving temporal window, each measured value can be used for several trend calculations and then be considered many times as steady, increasing, decreasing or unstable. We consider a value as stable and use it for mean calculation if it is qualified as stable 1.5 times more than as unstable. This empirical value was validated by the clinicians. Stable is eventually refined as "steady", "increasing" or "decreasing" based on the symbol the more frequent.

2) Using trends to control data abstraction: characterization of the stability interval by the mean and the standard deviation of the values labeled as stable

The purpose of this step is to perform a first abstraction, controlled by the trend characteristics. This results in the computation of time intervals called segments, where the mean is computed for each considered parameter (see Fig. 5).

Step 1.
Elements : raw labelled parameter values
Description : raw values.
Operation : mean.

Step 2.
Extracted information: trend, stability factors and characteristic span.
A priori knowledge: abstraction is valid when a parameter remains stable for a period of time.

Step 3.
The mean and the standard deviation are computed for each parameter based on their valid abstraction domains. These domains are identified based on the extracted information (step 2) and *a priori* knowledge.

Step 4.
The obtained results were displayed and validated by the clinicians.

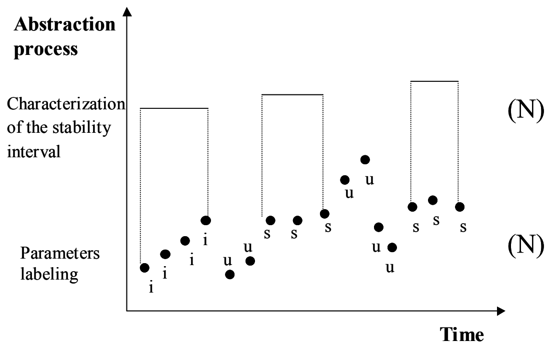


Figure 5. Abstraction process (2/4)
i : increasing, s : steady, u : unstable, N : numerical

3) Using a priori knowledge to control data abstraction: aggregation of the intervals according to temporal relations, mean values or raw data variations

The purpose of this step is to perform a second abstraction, controlled by *a priori* knowledge about parameter temporal variations. This results in the computation of refined time intervals (e.g. refined segments and adjusted mean values) (see Fig. 6).

Step 1.
Elements: parameter segments (e.g. temporal sections where the mean has been previously computed).

Description: mean values, segment length, temporal interval between segments.
Operation: concatenation and mean.
Step 2.
A priori knowledge : two kind of heuristic rules are used in this process. The first kind (a classical heuristic) checks the segment length, inter-segment interval and the gap between the segment means, considering the raw data standard deviation. The second kind controls that the parameter variation is below a given threshold or above this threshold but for a limited time extent corresponding to the characteristic span defined in [2]. The threshold for these rules is fixed by the clinician and depends on the precision of the measured data on which relies the parameter.
Step 3.
The data abstraction process consists in concatenating the segments under interest and computing their mean values.
Step 4.
The obtained results were displayed and validated by the clinicians.

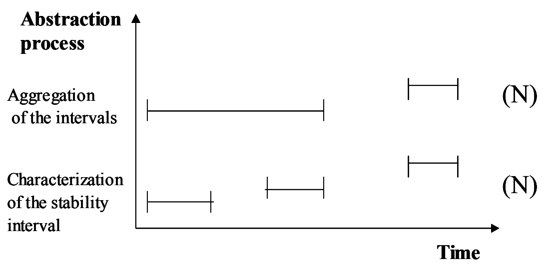


Figure 6: Abstraction process (3/4)
The first 2 segments are aggregated since the gap between them is smaller than their lengths, the difference between their means is low and the variation of the raw data during this gap is below a threshold of precision. These conditions are not satisfied for the following segments.

4) Numerical/symbolic transformation of segments.
The purpose of this step is to transform the quantitative values of segments into qualitative values: "normal", "unstable", "normal but high", "abnormal and high", "normal but low", "abnormal and low". All these information are given on temporal intervals (see Fig. 7).
Step 1.
Elements: parameter segments (e.g. temporal sections where the mean has been previously computed).
Description: mean values, segment length, temporal interval between segments.
Operation: classification
Step 2.
A priori knowledge: thresholds given by experts define a normal zone in which the parameter is considered as normal.
Common sense knowledge: Started from these thresholds, we add refinements: "normal but high" and "normal but low". These thresholds are respectively the mean between the sup (resp. inf) value normal and the mean between the 2 thresholds given by clinicians.
Step 3.
The data abstraction process consists in classifying the parameter segments.
Step 4.
The obtained results were displayed and validated by the clinicians.

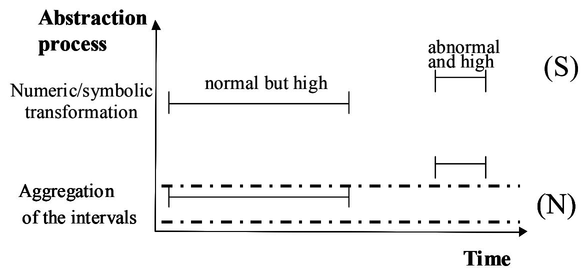
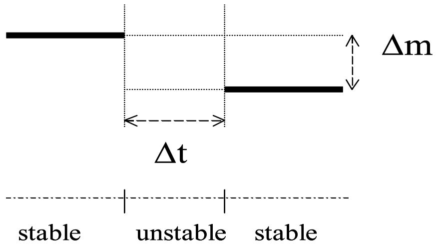


Figure 7 : Abstraction process (4/4)
N : numerical, S : symbolical
The dotted lines represent the zone for the parameter normality. When the parameter is close to the thresholds but into the band, we add the notion of "normal but high" and "normal but low".

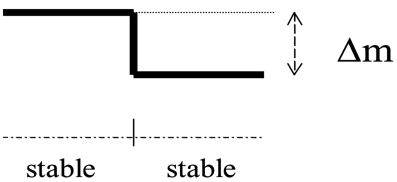
5.2 Complex abstraction for event construction

The objective of complex abstraction process or local patterns is to detect inner events that can be eventually outer events. For example, the event "disconnection of the sensors" is an outer event which can be detected by a change of the behaviour of the parameter. This event is an important cause of false alarms. An important inner event to detect too is the desaturation in blood gases (SpO2). In general, if the patient has a rapid desaturation, it is the sign of an alarming situation. These events have been chosen because of their importance in terms of clinic, frequency and feasibility detection. An another event was studied that we called "spike"; it is a rough variation of a parameter on a short time interval. It is just an inner event that we do not tie to an outer event or a clinical event.
This complex abstraction consists in 4 sub-steps too.
Step 1.
Elements: parameter segments (temporal sections where the mean has been previously computed).
Description: mean values, segment length, temporal interval between segments.
Operation: classification.
Step 2.
A priori knowledge : it consists in expert knowledge about the parameter behaviour for these events.

Common sense knowledge: it checks the segment length, inter-segment interval and the gap between the segment means.
Step 3.
The data abstraction process consists in classifying the patterns according to their thresholds.
Step 4.
The obtained results were displayed and validated by the clinicians and by comparing with the annotations file.
Technical details:
These patterns are based on unstable zones. When the parameter is considered as unstable, we search for some relations between time and variation of the means of the stable zones around the unstable zone. However, we can detect some patterns with no unstable interval (see Fig. 8).



8-a) A temporal gap Δt separates two stable periods (bold lines).



8-b) Meeting of two stable periods

Figure 8 : Two specific patterns. A difference Δm between their mean values can be measured.

Depending on the values of Δt and Δm, specific situations can be detected. Then, when Δt and Δm for SpO2 parameter are high a first desaturation situation can be recognised (see Fig. 8-a). A second desaturation situation can be recognised (see Fig. 8-b) when the decrease is stable. SpO_2 sensor disconnection can be recognised by the pattern shown in Figure 8-a when Δm is low. A disconnection results in a rough fall in SpO_2 leading to an unstable zone. After the reconnection, SpO_2 recovers its anterior value and the difference between values before and after this episode is low. A spike is a high variation of a parameter in a short time interval, occurring before and after a stability period. It is representative of the first pattern (8a) where Δm is high and Δt is small.

5.3 First results

An illustration of the proposed methodology is given in Figure 9. The graph at the top shows simple and complex abstractions for several cardio-respiratory parameters (Heart Rate, Respiratory Rate, Tidal Volume, Minute Volume, Oxygen Saturation (derived from pulse oxymetry SpO_2)and Systolic Arterial Pressure). The graph at the bottom corresponds to some events annotated by the observer.

A posteriori, the comments by three physicians were: at time 2080 seconds, there is a transient modification that can be related to a tracheal suction. Our abstraction process detects a SpO_2 disconnection (event a in Figure 9). The same event was detected at time 7000 seconds lasting longer than the previous one, followed by a true episode of desaturation starting around 7600 and ending 8600. Our abstraction process detects desaturation (event b in Figure 9). However, disconnection at 7000 is not identified. It is aggregated with the start of desaturation. The end of the desaturation episode is detected as a disconnection (see c) [3]. Two others episodes of disconnection around 10600 and 11200 are detected (d and e). We observe that "spikes" on respiratory parameters (RR, TV and MV) can correspond to suctioning (b, f), a more precise study of this relation is in progress.

6 Perspectives and conclusion

A reliable parameter abstraction process is a prerequisite for scenario construction. We have proposed in this paper a methodology to drive this process. This methodology is defined as incremental, tying data, information and knowledge in numerous different ways, and aimed at the definition of valid abstraction domains, e.g. domains where the abstraction process is performed under valid conditions. Results obtained on ICU data are encouraging. Applying this methodology on a large data set will allow its refinement and strengthen its validation (in course).

Mining is the next step in the process of knowledge discovery; "it consists of applying data analysis and discovery algorithms that, under acceptable computational efficiency limitations, produce a particular enumeration of patterns (or models) over the data" [6]. We search for knowledge we can represent as "If parameter1 unstable during 20 minutes BEFORE parameter2 unstable THEN a desaturation occurs" which lies state and events which temporal constraints. Two problems can be emphasized : 1) most of the time, data mining algorithms generate a set of association rules. How can we select the most informative ones? 2) How can we represent temporal relationships? How can we qualify and quantify associative rules containing temporal relationships?

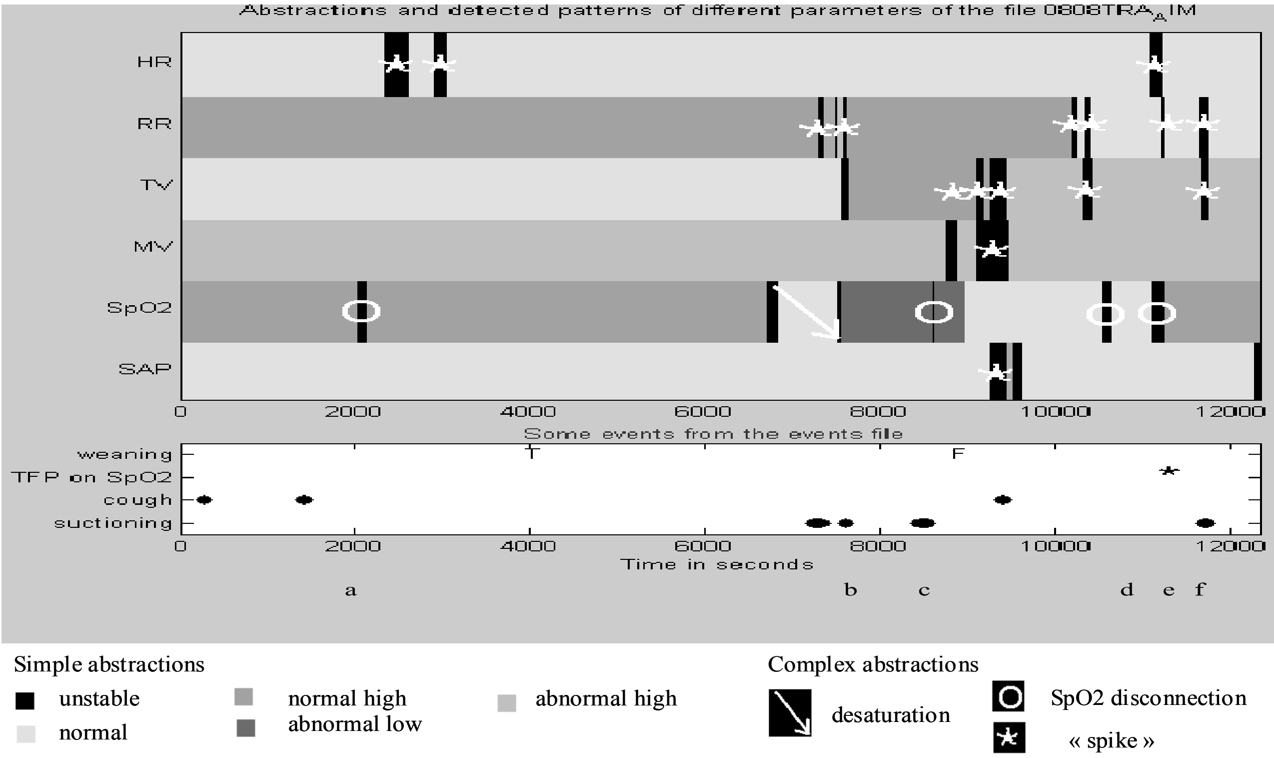


Figure 9 : Abstractions and real events.
Events file : on the row "weaning", "T" corresponds to the moment when the patient started to breath by himself on a T-piece and "F" corresponds to "failure", the patient was replaced on mechanical ventilation. "TFP on SpO2" signifies a technical false alarm on SpO2, as a general rule, it corresponds to a disconnection. Notice the events file is not exhaustive. a, b, c, d, e and f indicate the events detected by our abstraction process.

6.1 Informative power of association rules

Generally, 5 measures are used for ranking the usefulness and utility of discovered patterns. The measures for association rules such as "X→Y" are :

- support: supp(X→Y) = P(X,Y)
- confidence : conf(X→Y) = P(X,Y) / P(X)
- interest : intr(X→Y) = P(X,Y) / (P(X)P(Y))
- conviction : conv(X→Y) = P(X)P(¬Y) / P(X,¬Y)
- reliability : rel(X→Y) = P(X,Y) / P(X) – P(Y)

The following example illustrates the limitations in the usefulness of these measures:
<u>Example:</u>
We search for rule as « X→Y », with the co-occurrence of X et Y, conf=supp(X,Y)/supp(A) and supp(A)=interval length. (in temporal domain, support is not a probability but the length of the intervals [1, 10]). We obtain :"RR abnormal low → SpO2 unstable" conf=0.64 and supp=720 (1)
"RR normal → SpO2 normal high" conf=1 and supp=4342 (2)
"SpO2 normal high → RR unstable" with conf=1 and supp=1146 (3)
Based to medical experts, only the first rule is correct, the second has no interest and the third is false !

In fact, these measures are not sufficient to catch the informative power of associative rules.
Because the causal relation (X→Y) is similar to (¬Y→¬X) (e.g., rule 1 above), and the descriptive relation (X→Y) is the contrary of (X→ ¬Y) (e.g., an important decrease of SpO2 is a desaturation), Kodratoff proposes two additional measures [12]:

- for a descriptive relation : conf(X→Y) = P(Y/X) – P(¬Y/X)
- for a causal relation : conf(X→Y) =1/2(P(Y/X) + P(¬X/¬Y)) – P(¬Y/X).

Note that possible delay between the relation attributes are not taken into account.
<u>Example</u>
For the same rules than the previous example, we obtain as causal confidence :
"RR abnormal low → SpO2 unstable" causal conf=0.4541
"RR normal → SpO2 normal high" causal conf=0.5
"SpO2 normal high → RR unstable" causal conf=-0.4091
Note that these measures are comprise between –1 and 1. The obtained values seem to be more representative, the third rule has the lowest measure. But the 2 first have similar measures. In [8], the authors index 17 measures for ranking knowledge discovery. To be selective, they

propose in [9], 5 principles that a measure has to fulfil for being relevant. However, time is not explicitly represented and these measures do not appear appropriate for the discovery of association rules based on temporal sequences.

6.2 Temporal relationships in association rules

Discovering frequent episodes in temporal sequences is largely studied in the literature (see [15] for a review). Generally, the temporal sequencing is defined *a priori*. Recently, in [10], the problem of learning rules about temporal relationships between labelled time intervals have been studied. The authors apply the J-measure to rules with a modified semantics about the support of the pattern and add additional attributes of the state intervals. They obtain good results about the quality of the extracted rules. However, the methodology was only tested on simulated data. In [4,14], the authors describe a method for transforming real time series values into symbolic representations. This new representation is used for searching motifs. Presently, only a mono-parameter approach is proposed. In [1], the authors search for multi-parameter rules. But, searching for interval temporal rules is "very hard from both the theoretical and computational viewpoints". Starting from hemodialysis data, they try to discover rules such as "IF Trend of parameter1 is decreasing BEFORE State of parameter2 is High THEN dialyse fails". Presently, they can only find rules about co-occurrences of states or trends such as "IF Trend of parameter1 is decreasing AND State of parameter2 is High THEN dialyse fails". Again, standard confidence measure is used.

Future work should be concentrated on the extraction of patient-dependent knowledge and on the development of patient-adapted methods. This will imply the introduction of a learning phase in order to extract temporal scenarios tying information about parameter abstractions and alarms. Research about finding informative power measures for qualifying temporal interval association rules is under progress.

Acknowledgements
This work has been conducted under the OSSCAR RNTS project, under the auspice of the French minister of health. We thank the anonymous referees for their relevant remarks and helpful suggestions.

References
[1] R. Bellazzi, C. Larizza, P. Magni, R. Bellazzi, Quality assessment of hemodyalisis services through temporal data mining, *AIME* 2003.

[2] D. Calvelo, M.-C. Chambrin., D. Pomorsky, P. Ravaux, Towards symbolisation using data-driven extraction of local trends for ICU monitoring, *Arti. Intell Med* 19(3), 2000, pp. 203-223.

[3] M.-C. Chambrin & al, OSSCAR : A collaborative project for intelligent patient monitoring in intensive care and anesthesia, *workshop IDAMAP & KBIM-AC*, Chyprus, 2003.

[4] B. Chiu, E. Keogh, S. Lonardi, Probabilistic discovery of time series motifs, *9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, USA, 2003.

[5] M. Dojat, N. Ramaux, D. Fontaine, Scenario recognition for temporal reasoning in medical domains, *Artif Intell Med* 14(1,2), 1998, pp.139-155.

[6] U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, From data mining to knowledge discovery in databases, *AI magazine,* 37-54, 1996.

[7] I. J. Haimowitz, I. S. Kohane, Managing temporal worlds for medical trend diagnosis, *Artif Intell Med* 8(3), 1996, pp. 299-321.

[8] R. J. Hilderman, H. J. Hamilton, Knowledge discovery and interestingness measures : a survey , *Technical report CS-99-04, Department of Computer Science*, University of Regina, 1999.

[9] R. J. Hilderman, H. J. Hamilton, Principles for mining summaries using objectives measures of interestingness, *Proc. of the 12th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'00)*, pp72-81, Canada, 2000.

[10] F. Höppner, F. Klawonn, Finding informative rules in interval sequences, *Advances in Intelligent Data Analysis, Proc. of the fourth International Symposium, Lectures notes in Computer Sciences 2189,* Springer, Portugal, 2001.

[11] D. Kayser, *La représentation des connaissances*, Eds Hermès, France, 1997.

[12] Y. Kodratoff, Comparing Machine Learning and Knowledge Discovery in DataBases : An Application to Knowledge Discovery in Texts, *LNAI 2049*, pp1-21, 2000.

[13] C. Larizza, G. Bernuzzi, M. Stefanelli, A general framework for building patient monitoring systems, *Artif Intell Med* (5), 1995, pp. 91-102.

[14] J. Lin, E. Keogh, S. Lonardi, B. Chiu, A symbolic representation of time series, with implications for streaming algorithms, *8th ACM SIGMOD workshop on Research Issues in Data Mining and Knowledge Discovery*, San-Diego, CA, 2003.

[15] H. Mannila, H. Toivonen, A. Inkeri Verkamo, Discovering frequent episodes in sequences, *Proc. of the t conference on Knowledge Discovery and Data Mining (KDD-95)*, pp210-215, Canada, 1995.

[16] S. Miksch., W. Horn, C. Popow, F. Paky, Utilising temporal data abstraction for data validation and therapy planning for artificially ventilated newborn infants, *Artif Intell Med* 8(6), 1996, pp.543-576.

[17] J. Pitrat, *Métaconnaissance. Futur de l'intelligence artificielle*, Eds Hermès, France 1990.

[18] Y. Shahar, A framework for knowledge-based temporal abstraction, *Artif. Intell.*, 90 (1-2), 79-133, 1997.