

A method for automated temporal knowledge acquisition applied to sleep-related breathing disorders

G. Guimarães¹, J.-H. Peter², T. Penzel², A. Ultsch³

¹CENTRIA (Center of AI), Universidade Nova de Lisboa, Portugal
phone +351-21-2948536, fax +351-21-2948541, guimas@uevora.pt

²Schlaflabor, Philipps-Universität Marburg, Germany
phone +49-6421-2866435, fax +49-6421-2865450, penzel@mailers.uni-marburg.de

³Department of Computer Science, Philipps-Universität Marburg, Germany
phone +49-6421-2822185, fax +49-6421-2828902, ultsch@informatik.uni-marburg.de

Abstract

A method for the discovery of temporal complex patterns (TCPs) in multivariate time series and their conversion into a linguistic knowledge representation is introduced and applied to sleep-related breathing disorders (SRBDs) recorded by polysomnography (PSG). This method enables a simultaneous evaluation of respiration related signals of a PSG in order to identify and classify different types of SRBDs and their severity. The main idea lies in introducing several abstraction levels that allow a stepwise detection of TCPs, in this case SRBDs. First, Artificial Neural Networks with unsupervised learning are used for the detection of elementary patterns in selected time series. At this level, a rule-based description of the detected structures is generated using a machine learning algorithm. In the following levels, temporal relations among them as well as higher level patterns are extracted until the final detection of TCPs using methods from Artificial Intelligence and Statistics. At each level, a conversion into a linguistic representation in form of temporal grammatical rules is performed. This method covers one of the main “bottlenecks” in the design of Knowledge Based Systems, namely the problem of knowledge acquisition. A qualitative evaluation of the knowledge base is made using a questionnaire. For the first time a linguistic description of SRBDs has been generated automatically. Compared to expert knowledge the different types of SRBDs and their total number have been correctly identified. An evaluation of the rules at a lower abstraction level lead to an overall sensitivity of 0,762 and a specificity of 0,758. This is due to the fact that the expert couldn't give an exact answer (maybe or don't know) referring to some of the pattern descriptions. However, none of the pattern descriptions was incorrectly described (answer no).

1. Introduction

The increasing development towards more powerful computers in the last few years has conducted to the measurement and storage of a great amount of data from, for example, industrial processes or medical applications. Often continuous signal data are acquired in order to enhance the diagnosis or control of the underlying process. These signal data are frequently recorded for time periods leading to multi-signal recordings that describe the underlying physiologic or pathologic process.

For example, consider a human being that suffers from daytime sleepiness. Sleepiness is often caused by insufficient sleep, which may result from shift work or poor sleep hygiene. However, the most common medical disorder causing excessive daytime sleepiness appears to be sleep apnea [36]. For the diagnosis of sleep apnea the dynamics of physiological parameters such as respiration and heart rate, have to be recorded and evaluated. Moreover, for an accurate analysis of sleep apnea, a large number of parameters are involved. These are sleep-related signals (EEG, EOG, EMG), signals concerning the respiration (airflow, ribcage and abdominal movements, oxygen saturation, snoring) and circulation-related signals (ECG, blood pressure). The continuous recording of these signals is known as polysomnography (PSG). The identification of the different types of sleep apnea, namely apnea and hypopnea, can be carried out just using signals directly related to respiration [26]. Counting the number of individual apneas and hypopneas per hour of sleep, which is the apneaindex and the hypopneaindex, gives a measure for the extend of these disorders. The respiratory disturbance index (RDI) is the sum of apneas and hypopneas per hour of sleep. It can be seen as pathological, when the RDI is larger than 20 events per hour. After 40 events per hour, the patient has to be referred to therapy.

The diagnosis of sleep apnea, more exactly sleep-related breathing disorders (SRBDs), is made by physicians with the help of technical assistants who perform a visual classification of the different types of apneas using a PSG. An automatic identification of the SRBDs is a very hard task, as on the one side a great amount of signals have to be analysed simultaneously. On the other side, quite different patterns for the same SRBD may occur, even for the same patient. In addition, a strong variation of the duration of each SRBD may also occur [27],[28].

In order to solve these problems, we propose a step-wise identification of SRBDs, such that at different abstraction levels a progressively higher temporal abstraction of the underlying signals can be obtained. The proposed method introduces several abstraction

levels, where Artificial Neural Networks (ANNs) as well as methods from statistics and signal processing are used for the identification of the SRBDs. The advantage in using ANNs lies in their effectiveness in applications dealing with sub-symbolic data, in particular, with noisy or even inconsistent real-world data. Such a sub-symbolic processing seems to be more adequate for pattern recognition problems [5]. Usually supervised ANNs are used to solve these problems. We use unsupervised ANNs, since they are able to learn structures from high dimensional data without any a priori information, i.e. a pre-classification of the patterns [6].

Another main aim of our approach lies in the generation of a temporal knowledge representation of the detected patterns in form of temporal grammatical rules, understandable for a human beings as, for example, medical experts. This includes the use of a machine learning algorithms for the automatic generation of PROLOG rules. Consequently, those rules can be used as knowledge base in traditional artificial intelligence (AI) technologies as, for example, Knowledge Based Systems (KBS). Those systems have been successful in areas like diagnosis, control and planning [19].

The advantages of both technologies, AI-technologies versus ANN-technologies, are wide-ranging. However, both approaches show up some limits, for instance, the incapacity of ANN to explain their behavior and the problem of knowledge acquisition for AI-systems. Recently, there has been an increasing interest in hybrid systems that integrate both technologies [13]. Usually, hybrid systems entail several modules that cooperate with one another. Each module is implemented in a different technology. In contrast, we are mainly interested in hybrid systems that perform a *knowledge conversion*, i.e. a transition between different knowledge representation forms, a sub-symbolic and a symbolic knowledge representation [39]. Previous applications where a knowledge conversion is performed [38], [42] do not consider temporal dependencies among data. In this paper a method will be presented that extends those concepts to multivariate time series enabling a *temporal knowledge conversion* [14], [16].

In section 2 some basic concepts of the Self-organizing Maps (SOMs) for the detection of structures in high-dimensional data are given. The method is introduced in section 3. Section 4 entails a detailed description of the experiments made with data from three patients with SRBDs. The evaluation of the method applied to SRBDs will be given in section 5. Section 6 entails a discussion of the presented method and its application to sleep-related breathing disorders. Conclusions are presented in section 7.

2. Self-organizing maps for the detection of structures in high-dimensional data

Artificial Neural Networks (ANNs) may be classified according to their learning principles in ANNs with supervised and unsupervised learning. The most popular supervised learning algorithm is the Backpropagation algorithm [34]. These kinds of ANNs are often used for pattern recognition problems, where a classification of the data is a priori available, for instance, for the recognition of images representing handwritten characters [5].

In contrast, we are mainly interested in ANNs with unsupervised learning that adapt their internal structures to structural properties (e.g. regularities, similarities, frequencies etc.) of the data. This is due to the fact, that our main aim is the discovery of inherent structures in high-dimensional data without any a priori information. ANNs like Competitive Learning, ART and the Self-organizing Map belong to this type [33]. We are mainly interested in Self-organizing Maps (SOMs), as proposed by Kohonen (1982), since here a topology preserving projection of data in a high-dimensional space to a lower dimensionality (usually the $|\mathbb{R}^2$) is performed, enabling the detection of the structures inherent to the data at a lower dimension. Another advantage of SOMs lies in the possibility in detecting new, not yet known structures in the data.

During learning SOMs adapt their weights such that a high-dimensional input space is projected onto a, usually, two-dimensional map preserving the neighborhood of the input data on the map (see Fig. 1). This means that the result is formed by the properties inherent to the data itself, i.e. no master tells us if the classification is correct or not. Consequently, no previous knowledge is afforded about the structures in the data. The input layer has n units representing the n components of a data vector $v = (v_1, \dots, v_n)$. The output layer is a two dimensional array of units arranged on a map. Each unit in the input layer is connected to every unit in the output layer with a weight $w_r = (w_{r0}, \dots, w_{rn})$ associated. All weights are initialized randomly. They are adjusted according to Kohonen's learning rule [19]

$$\Delta w_r = \eta(t) \cdot h_{rr'}(t) \cdot (v - w_r) \quad (1)$$

that uses a distance measure $\|w_{r'} - v\| = \min \|w_r - v\|$ to determine the bestmatch r' and a neighborhood function

$$h_{rr'}(t) = e^{-\frac{(r-r')^2}{2\sigma(t)^2}} \quad (2)$$

that realizes the lateral inhibition. The learning rate determines the strength of the learning with $0 < \eta(t) < 1$. The radius $\sigma(t)$ determines the set of neurons in a neighborhood of the bestmatch that are included into the learning process. Both functions usually decrease monotonously during learning.

On the map neighboring units form regions that correspond to similar input vectors. These neighborhoods form disjoint regions, thus enabling a classification of the input vectors. However, in order to perform a classification, a visualization of the network structures is needed, since the Kohonen algorithm converges to an equal distribution of the units in the output layer [32]. Therefore, a special visualization algorithm, called U-Matrix method, was developed [37]. This algorithm produces a three dimensional landscape that represents structural properties of the high dimensional input space on the map. The idea is to analyse the weights at each point of the grid with respect to their neighbors and then display the distance between two neighbor units as height into the third dimension. This lead to an interpretable, three-dimensional landscape, called U-Matrix. A U-Matrix has valleys where the vectors on the map are close to each other and represent data that are in the same class. Hills or walls represent larger distances indicating dissimilarities of the input data. Using this visualization technique, clustering of data becomes possible such that the method can be used for exploratory data analysis. In the last years SOMs together with the U-Matrix method have been successfully applied to a wide-ranging number of applications where a clustering of high-dimensional data was afforded [17], [21], [37], [40], [42].

3. A method for automated temporal knowledge acquisition

In this section, the recently developed method for temporal knowledge conversion (TCon) [14] is presented. The method enables an automated temporal knowledge acquisition with extended Self-Organizing Maps (SOMs) for the discovery of temporal complex patterns (TCPs) in multivariate time series. After the discovery process, the results of the SOMs are used for an automatic generation of temporal rules using Machine Learning algorithms. These rules are a linguistic description of the discovered TCPs, intelligible for human beings, and are part of the temporal knowledge base of a KBS. In order to be able to perform such a temporal knowledge conversion, an adequate knowledge representation for multivariate time series has to be found considering the following constraints:

- It has to be simple enough such that both, the discovery and the automatic generation of a rule-based description, becomes feasible using methods from Artificial

Intelligence, such as ANNs with unsupervised learning and Machine Learning algorithms.

- The results must be interpretable and intelligible for domain experts, since exploratory methods are used for the extraction of temporal knowledge from multivariate time series and no a priori knowledge is presumed.
- Not too strong restrictions should be made in order to capture the main structures without a loss of information.

The main idea lies in breaking down this highly structured and complex problem into several sub-problems introducing several abstraction levels (see Fig. 2), where at each level the main structures in multivariate time series as well as temporal relations among those structures are detected successively using SOMs. At each level, the discovered structures are converted into a symbolic knowledge representation in form of temporal grammatical rules. This includes the use of Machine Learning algorithms.

Previous approaches for the generation of a syntactical description of signals in form of grammars or automata used a pre-classification of the signals, for instance, sub-patterns as the QRS complex in ECG signals [22] or used simple waveform recognition operations as local minimum or negative slope [4]. In this paper a new approach is presented, where the main structures in multivariate time series are extracted using clustering algorithms, such as SOMs. As we are dealing with more than one time series, rules are generated at different abstraction levels. In order to generate a temporal knowledge base from the results of the ANN, we propose to use a machine learning algorithm, called sig* [38], that generates PROLOG rules at the first abstraction level. At the following levels temporal grammatical rules are generated. An implementation of those rules could be using definitive clause grammars (DCGs) [15]. DCGs are an extension of context-free grammars introducing conditions and tests in the rules [29] such that temporal conditions, for instance, the duration can be specified. A PROLOG-Interpreter then can be used as a parser.

Fig. 2 shows the different abstraction levels and steps of the proposed method. A multivariate time series $Z = \bar{x}(t_1), \dots, \bar{x}(t_n)$ with $\bar{x}(t_i) \in \mathbb{R}^m, m > 1$ sampled at equal time steps t_1, \dots, t_n from signals of complex processes is the input of the system. The results of the method TCon are the discovered TCPs and a linguistic description of the TCPs in form of temporal grammatical rules, as well. These rules should be interpretable and intelligible for human beings in order to permit an evaluation of the generated temporal knowledge. Next, we will describe the different abstraction levels.

FEATURES: First, all time series have to be transformed into a suitable representation form, such that clustering techniques, for instance, SOMs, can be used at the next level for the extraction of the main structures in multivariate time series. This step is a prerequisite for further processing, specially, when dealing with data provided from medical applications, where recordings from several persons have to be considered. For most practical applications the choice of an adequate preprocessing will be one of the most significant factors in determining the final performance of the system [5]. This means that for each application the feature extraction process may differ strongly. In addition, an improvement of the feature extraction may be achieved using domain knowledge.

For the feature extraction process, methods, for instance, from statistics or signal processing may be used [11], [23]. For the feature extraction one or even more than one time series $\vec{x}_S(t_i) = x_{j_1}(t_i), \dots, x_{j_S}(t_i)^T \in \mathbb{R}^S$ may be selected from the multivariate time series Z with $j_k \in S, k = 1, \dots, s, S \subset \{1, \dots, m\}, s = |S|$. A feature $m_S(t_i, l) = f(\vec{x}_S(t_i), \dots, \vec{x}_S(t_{i+l}))$ then is the value of a function $f: \mathbb{R}^{s \times l} \rightarrow \mathbb{R}$ at time t_i with $l \in \{1, \dots, n-l\}$ from selection S .

PRIMITIVE PATTERNS: At this level, SOMs together with U-Matrices, as described in the previous section, are used to find elementary patterns in temporal data, i.e. multivariate time series. The discovered patterns $p_j, j = 1, \dots, k$ are named as primitive pattern classes. An element of a primitive pattern class is a primitive pattern $p_j(t_i), j = 1, \dots, k$ that belongs to a given primitive patterns class and is associated to a given time point t_i . Regions on a U-Matrix that do not correspond to a specific primitive pattern class are associated to a special group, named *tacet*. We are now able to classify the whole time series with primitive patterns and tacets. This will be called a primitive pattern channel.

Instead of determining primitive pattern classes for all time series simultaneously, several feature selections are made. For each feature selection several, but not all, features are selected. This results in several SOMs and U-Matrices. For example, for the detection of SRBDs, we may consider two feature selections, features concerning mainly the respiration and features concerning mainly the movements (see section 4). The main advantage of this proceeding lies in the possibility of dealing with a great amount of time series, performing a selection of contiguous features. This makes sense, especially if we have in mind to extend our method to a larger number of quite different time series, for instance, EEGs, ECGs, etc.

After the determination of all primitive patterns classes for each feature selection, a symbolic description of the classes is generated. Therefore, Machine Learning algorithms,

for instance, CN2, NewId, CART or C5 [25] may be used. For this application, we used the machine learning algorithm sig* [38], since it generates a rule-based description of each primitive pattern class in form of PROLOG rules and each rule describes a primitive pattern class by its most significant features. Another important issue related to a rule-based description of primitive pattern classes lies in the generation of intelligible names for each primitive pattern class. This enables the generation of intelligible temporal grammatical rules at the next levels that just now can be interpreted by human beings as, for example, domain experts.

For the first time, the machine learning algorithm sig* was applied to time series. It takes a classified data set in a high-dimensional space as input and produces descriptions of the classes in form of decision rules (see Example 1). In order to construct appropriate conditions that characterize each class, in this case primitive pattern classes, this algorithm selects significant features for each class. In particular, the generated rules take the significance of the different structural properties of the classes into account. If only a few properties account for most of the cases of a class, the rules are kept very simple. In order to formulize apt conditions for each selected significant feature, the distribution properties of the features of a class can be used. Assuming a normal distribution for a certain feature, this means that 95% of the feature values are kept in the limits $[\text{mean} - 2 \cdot \text{dev}, \text{mean} + 2 \cdot \text{dev}]$, where dev is the value of the standard deviation of the attribute. If no normal distribution can be assumed, the median and quantiles can be used as limits.

Example 1: Structure of a sig*-rule for primitive pattern classes

```
A Primitive pattern is a 'primitive_pattern_name'
if
    'feature i'  $\in$   $[\text{min}_i, \text{max}_i]$ 
and
    'feature j'  $\in$   $[\text{min}_j, \text{max}_j]$ 
and
    ...
and
    'feature k'  $\in$   $[\text{min}_k, \text{max}_k]$ 
```

Altogether, this abstraction level provides the discovered primitive pattern classes and the rules generated by the machine learning algorithm sig*, as well.

SUCCESSIONS: Succeeding identical primitive patterns $p_j(t_i), \dots, p_j(t_{i+k})$, $i=1, \dots, n-k+1$ obtained from one SOM are a *succession* $s_j(a, e)$, where $a := t_i$ is the

starting point, $e := t_{i+l}$ the end point and $l = e - a$ the duration of succession s_j . Each primitive pattern is represented as a bestmatch on a U-Matrix. Trajectories of temporally succeeding primitive patterns, i.e. bestmatches, on a U-Matrix can be used for the identification of successions (see Fig. 3). If a trajectory remains in the same valley, those bestmatches, i.e. primitive patterns, belong to the same succession. A transition to another valley stands for a transition to a new succession. Such a transition may include regions on a U-Matrix, that do not belong to any primitive pattern class. Those regions represent a special group of neurons on a map, named as *tacets*. If a transition includes one or more tacets, a delay between successions occurs. Since neighboring regions on a SOM represent similar bestmatches, a transition to a quite different primitive pattern can be identified considering the distance on the map.

EVENTS: Since several SOMs are learned from different feature selections, two or more successions $s_1(a_1, e_1), \dots, s_q(a_q, e_q)$ may occur more or less simultaneously. At this level, more or less simultaneously successions that occur more than once, are identified as an *event* $e(l)$. $A = \max(a_1, \dots, a_q)$ is the starting point, $E = \max(e_1, \dots, e_q)$ the end point and $l = E - A$ the duration of the event.

As there may occur differences in the frequency of the occurrence of events, a distinction between *significant events* (very frequently occurring) and *less significant events* (less frequently occurring) will be made. The significance of events is calculated using conditional probabilities. Rare events are omitted in the sense that they are regarded as delays between events, named as *event tacets*. The idea is to select the most significant events and associate less significant events to them. Therefore, similarities among significant and less significant events will be considered counting the number of the same types of successions occurring in both events. This results in a reduced number of events. Each event is described by one significant event and, possibly, one or more than one less significant events. At this level, the whole multivariate time series is described by a sequence of events $F = e_{1j}, \dots, e_{mj}, j = 1, \dots, m$, with m as the number of different types of events in F .

For each event a grammatical rule is generated with the name of the event at the left side of the rule. The right side of the rule entails alternatives of more or less simultaneous successions using the tokens '*more or less simultaneous*' and '*or*'.

Example 2: Structure of a temporal grammatical rule for events

An event is a 'event_name'

```

if
    'succession  $i_1$ ' and ... and 'succession  $l_n$ '
is more or less simultaneous
    'succession  $j_1$ ' and ... and 'succession  $j_m$ '
is more or less simultaneous
...
is more or less simultaneous
    'succession  $k_1$ ' and ... and 'succession  $k_n$ '

```

SEQUENCES: At this level, we determine subsequences of events e_i, \dots, e_k that occur more than once in F . Events may succeed immediately or after a time delay, i.e. an event tacet. A sequence $sq(\min, \max) = e_i(\min_i, \max_i), \dots, e_k(\min_k, \max_k)$ then is a subsequence of events that occurs more than once (see Fig. 4). Each event type e_i in a sequence has a minimal \min_i and a maximal duration \max_i . The minimal and maximal duration of all subsequences e_i, \dots, e_k in F , is given by \min and \max .

For the determination of sequences, probabilistic automata [6] as well as the duration of event tacets can be used. Probabilistic automata describe transition probabilities between events such that paths through such an automata describe possible sequences of events. The determination of sequences with the duration of event tacets is made by just considering the occurrence of larger event tacets that are interpreted as transitions between different sequences due to larger delays between succeeding events.

For each sequence a grammatical rule is generated with the name of the sequence and the duration of the sequence at the left side of the rule (see Example 3). The right side of the rule entails a description of the succeeding events and their duration using the tokens *followed by* and *followed after ... by*. We emphasize, that the same event type may appear in different sequences.

Example 3: Structure of a temporal grammatical rule for sequences

```

A sequence is a 'sequence_name' [min, max]
if
    'Event $_i$ ': 'name of event  $i$  [min $_i$ , max $_i$ ]'
followed by / followed after [min $_{tj}$ , max $_{tj}$ ] by
    'Event $_j$ ': 'name of event  $j$  [min $_j$ , max $_j$ ]'
followed by / followed after [min $_{tk}$ , max $_{tk}$ ] by
    'Event $_k$ ': 'name of event  $k$  [min $_k$ , max $_k$ ]'

```

...

followed by / followed after $[\min_{t_1}, \max_{t_1}]$ **by**

`'Event1': 'name of event 1 $[\min_1, \max_1]$ '`

TEMPORAL PATTERNS: At the last level, similarities among sequences are determined. Therefore, string exchange algorithms can be used [1]. Similar sequences $sq_1(\min_i, \max_i) \vee \dots \vee sq_v(\min_v, \max_v)$ will be joined together to a *temporal pattern* $tp(\min, \max)$, where $\min = \min(\min_i, \dots, \min_v)$ and $\max = \max(\max_i, \dots, \max_v)$ is the minimal and maximal duration of the temporal pattern, resp.

At this level alternatives between similar types of sequences are described by a grammatical rule using the token “or” (see Example 4). This step may be dismissed, if a temporal pattern entails just one sequence.

Example 4: Structure of a temporal grammatical rule for temporal patterns

A temporal pattern is a `'temporal_pattern_name' $[\min, \max]$`

if

`'sequencei' $[\min_i, \max_i]$ '`

or

`'Sequencej' $[\min_j, \max_j]$ '`

or

...

or

`'Sequencek' $[\min_v, \max_v]$ '`

4. The Method applied to Sleep-related Breathing Disorders

The method was applied to a sleep disorder with a high prevalence, called sleep apnea. More exactly, sleep-related breathing disorders (SRBDs) consist of various types among which sleep apnea is best known. For the diagnosis of sleep apnea the temporal dynamics of physiological parameters such as respiration and heart rate, have to be recorded and evaluated. Moreover, for an analysis of sleep apnea, a large number of parameters are involved, such as sleep-related signals (EEG, EOG, EMG), signals concerning the respiration (airflow, ribcage and abdominal movements, oxygen saturation, snoring) and circulation related signals (ECG, blood pressure). The parallel recording of those signals during sleep is known as polysomnography (PSG).

One of our main goals was the identification of different types of sleep related breathing disorders, mainly apnea and hypopnea. Therefore, only the signals concerning the respiration

had to be considered [27]. Severity of the disorder is calculated by counting the number of apnea events per hour of sleep. This is called apnea-index. The sum of the index of apneas and hypopneas is a measure for the respiratory disturbance index (RDI). It can be seen as pathological, when the RDI exceeds 20 events per hour of sleep. After having 40 events per hour of sleep, the patient has to be referred to therapy.

Technical assistants usually make the visual classification of the different types of SRBDs based on a PSG recording. An automatic identification of SRBDs is a very hard task, since all signals have to be analysed simultaneously. In addition, quite different patterns for the same SRBD may occur, even for the same patient during the same night, and a strong variation of the duration of each event may occur, as well [27], [28].

SRBDs can be subdivided into SRBDs with and SRBDs without an obstruction of the upper respiratory tracts. The different kinds of SRBDs are identified through the signals 'airflow', 'ribcage movements' and 'abdominal movements', 'snoring' and 'oxygen saturation', where a distinction between amplitude-related and phase-related disturbances is made (see Fig. 5). Concerning the amplitude-related disturbances, we distinguish disturbances with 50% as well as disturbances with 10-20% of the baseline signal amplitude. Phase-related disturbances are characterized by a lag between 'ribcage movements' and 'abdominal movements'. An interruption of 'snoring' is present at most SRBDs as well as a drop in 'oxygen saturation'. We emphasize that only a simultaneous evaluation of all time series enables an identification of the different types of SRBDs [27].

We used 25 Hz sampled data from three patients having the most frequent SRBDs (altogether 27 patterns) and made a filtering to 2,5 Hz in order to reduce the large amount of data. No additional information was provided from the medical experts, since exploratory methods as ANNs with unsupervised learning are used in order to find inherent structures in the time series. Those methods include the possibility of finding new, not yet-known structures. In this section we will present the results for all SRBDs. However, the method will be explained using one of the detected SRBDs.

FEATURES: Altogether, 12 features have been calculated (see Fig. 6 for the extracted features of one of the SRBDs). For the feature extraction we considered the already mentioned characterization of possible occurring disturbances (see Fig. 5). In order to capture changes in the amplitude, the volume of each of the time series 'airflow', 'ribcage movements' and 'abdominal movements' with a lag of $2*d+1$ (this corresponds to a single breath) has been calculated. As multimodal distributions occurred, we deduced fuzzy-membership-functions [43] from the histograms of the volumes of each time series. We

obtained the linguistic variables ‘no airflow’, ‘reduced airflow’ and ‘strong airflow’, resp. for ‘ribcage movements’ and ‘abdominal movements’. For the identification of lags between ‘ribcage movements’ and ‘abdominal movements’ cross-correlations have been calculated. In addition, a rescaling of the measure ‘snoring’ was made. Since ‘oxygen saturation’ is not a relevant feature for the pattern detection process, a switch was implemented, enabling the identification of a drop (decline) of at least 4% of the oxygen saturation for the past 10 sec.

PRIMITIVE PATTERNS: After the feature extraction SOMs together with U-Matrices are used for the detection of *primitive patterns*. As we are searching for structures in time series related mainly to airflow to the movements, two feature selections have been made:

- Features concerning mainly the airflow: ‘strong airflow’ $\in [0, 1]$, ‘reduced airflow’ $\in [0, 1]$, ‘no airflow’ $\in [0, 1]$, ‘lag of respiratory movements’ $\in [-1, 1]$, ‘snoring intensity’ $\in [0, 1]$.
- Features concerning mainly the movements: ‘strong ribcage movements’ $\in [0, 1]$, ‘reduced ribcage movements’ $\in [0, 1]$, ‘no ribcage movements’ $\in [0, 1]$, ‘strong abdominal movements’ $\in [0, 1]$, ‘reduced abdominal movements’ $\in [0, 1]$, ‘no abdominal movements’ $\in [0, 1]$, ‘lag of respiratory movements’ $\in [-1, 1]$, ‘snoring intensity’ $\in [0, 1]$.

The advantage of a feature selection lies in the possibility of detecting elementary structures in time series not having to consider all of them at the same time. We remark here that one feature may occur in several feature selections. For example, consider the feature ‘snoring intensity’ that occurs in both feature selections.

Consequently, two feature selections lead to the generation of two SOMs and U-Matrices. The latter are called ‘AirflowUMX’ and ‘MovementsUMX’ (see Fig. 3). Regions on a U-Matrix that do not correspond to a specific cluster are associated to a special group, called tacet class. All other regions are associated to a primitive pattern class. We now are able to classify the whole time series with primitive patterns and tacets (see Fig. 7). This will be called a primitive pattern channel. Altogether, 6 primitive pattern classes A_1, \dots, A_6 have been identified from ‘AirflowUMX’ and 9 primitive pattern classes B_1, \dots, B_9 from ‘MovementsUMX’.

Although primitive patterns classes have already been identified, until now no meaningful and intelligible names have been given to the classes. This is, however, an important task, as we aim to obtain a symbolic and intelligible description of the discovered patterns at all abstraction levels. Therefore, a machine learning algorithm, called sig* [38], was used to

generate a rule-based description of each primitive pattern class. This means that for each primitive pattern class a characterizing rule is generated with the most significant features. The naming of each primitive pattern class is straightforward, since each rule entails the most significant and characterizing features for each class. Values nearby one mean that this feature occurs with a high probability, while values nearby zero mean that this feature will probably not occur. For example, primitive pattern 'A2' was named as 'no airflow without snoring', since the feature 'no airflow' occurs with a high probability with values in [0.951, 1] and a strong 'snoring intensity' will probably not occur with values in [0, 0.241]. The following example shows some of the sig* rules for two primitive pattern classes from different U-Matrices.

Example 4: sig rules for the primitive patterns 'A2', 'A4' and 'B3':*

```
A primitive pattern is a 'A2'  
if  
    'no airflow' ∈ [0.951, 1]  
and  
    'reduced airflow' = 0  
and  
    'snoring intensity' ∈ [0, 0.241]
```

This primitive pattern class was named as 'no airflow without snoring'.

```
A primitive pattern is a 'A4'  
if  
    'strong airflow' ∈ [0.367, 1]  
and  
    'no airflow' = 0  
and  
    'snoring intensity' ∈ [0.15, 1]
```

Name of the primitive pattern: 'strong airflow with snoring'.

```
A primitive pattern is a 'B3'  
if  
    'no ribcage movements' ∈ [0.772, 1]  
and  
    'no abdominal movements' ∈ [0.641, 1]  
and  
    'reduced ribcage movements' = 0  
and
```

`'snoring intensity' = 0`

Name of the primitive pattern: 'no respiratory movements without snoring'.

These names will be used at the next level for the description of the temporal grammatical rules of the events. The following names have been generated from the sig*-rules for all primitive pattern classes:

Primitive pattern classes from 'AirflowUMX':

- A1: 'reduced airflow with snoring'
- A2: 'no airflow without snoring'
- A3: 'strongly reduced airflow with snoring'
- A4: 'strong airflow without snoring'
- A5: 'reduced airflow without snoring'
- A6: 'strongly reduced airflow without snoring'

Primitive pattern classes from 'MovementsUMX':

- B1: 'reduced parallel respiratory movements with snoring'
- B2: 'reduced parallel respiratory movements'
- B3: 'no respiratory movements without snoring'
- B4: 'laged respiratory movements'
- B5: 'reduced by strong laged respiratory movements'
- B6: 'strong respiratory movements'
- B7: 'no strong laged respiratory movements'
- B8: 'reduced ribcage movements and no abdominal movements without snoring'
- B9: 'reduced parallel and laged ribcage movements and no abdominal movements without snoring'

SUCCESSIONS: At this level, temporal relations among the primitive patterns in a primitive pattern channel have been considered for the identification of successions. Each succession has a corresponding primitive pattern type. We identified 6 different types of successions 'AirflowUMX' and 9 different types of successions from 'MovementsUMX'. All of them start at a different time point and most probably have different durations. Successions can be identified by trajectories visualized on U-Matrices (see Fig. 3). A transition through a region on the U-Matrix associated to the tacet class means that a delay between two succeeding successions occurs (see Fig. 7).

EVENTS: This level entails the identification of more or less simultaneous successions occurring at the different primitive pattern channels. Therefore, the number of

simultaneously occurring primitive patterns has been calculated, since very small deviations are allowed for more or less simultaneous successions. In order to determine the significance of events (significant, less significant and rare events), conditional probabilities have been calculated for all primitive patterns $A_i, i = 1, \dots, 6$ of ‘AirflowUMX’ with

$$P(A_i|B_j) = \frac{P(A_i \cap B_j)}{P(B_j)} \quad (3)$$

and for all primitive patterns $B_j, j = 1, \dots, 9$ of ‘MovementsUMX’ with

$$P(B_j|A_i) = \frac{P(B_j \cap A_i)}{P(A_i)}. \quad (4)$$

We distinguish between significant and less significant events calculating histograms over the conditional probabilities (see Fig. 8). Most significant events have probabilities $P(A_i|B_j) \geq 0.5$ or $P(B_j|A_i) \geq 0.5$, while less significant events have probabilities $0.12 \leq P(A_i|B_j) < 0.5$ or $0.11 \leq P(B_j|A_i) < 0.5$. The probability for all rare events, i.e. event tacets, lies below these values. Altogether, 6 significant and 11 less significant events have been identified. In order to join less significant and significant events, the number of the different types of successions in different events was calculated. For example, ‘Event5’ was identified as an event that entails one significant and two less significant events. The significant event entails the successions A_6 : ‘strong airflow with snoring’ and B_4 : ‘strong respiratory movements’ that occur more or less simultaneously. The less significant events are characterized by the following more or less simultaneous occurring successions: A_1 : ‘reduced airflow with snoring’ or A_0 : ‘tacets’ occurs more or less simultaneously with B_4 : ‘strong respiratory movements’. See example 6 for a temporal grammatical rule of ‘Event5’.

The occurrence of tacets in a rule means that small interruptions may occur in a succession or that a succession, for example, from one primitive pattern channel occurs simultaneously with irrelevant information at the other channel.

Example 6: Temporal grammatical rules for ‘Event2’, ‘Event3’ and ‘Event5’:

An event is a

‘Event2’

if

‘no airflow without snoring’

is more or less simultaneous


```

        ('no respiratory movements without snoring'
         and
         'tacets')
An event is a
    'Event3'
if
        'no airflow without snoring'
is more or less simultaneous
        ('reduced ribcage and no abdominal movements without
         snoring')
An event is a
    'Event5'
if
        ('strong airflow with snoring'
         and
         'reduced airflow with snoring'
         and
         'tacets')
is more or less simultaneous
        'strong respiratory movements'

```

The following intelligible names for all events have been derived from the temporal grammatical rules. The naming is straightforward, since already intelligible names have been derived from the sig*-rules at a lower abstraction level for the primitive pattern classes and, therefore, also for the successions:

- Event1: 'reduced airflow with snoring'
- Event2: 'no airflow and no respiratory movements without snoring'
- Event3: 'no airflow and reduced ribcage movements and no abdominal movements without snoring'
- Event4: 'no until reduced airflow and reduced parallel and laged respiratory movements without snoring'
- Event5: 'strong breathing with snoring'
- Event6: 'reduced airflow and laged respiratory movements without snoring'

SEQUENCES: At this level, the whole multivariate time series is represented by a sequence of events. TCPs then are subsequences of events that occur repeatedly in this sequence. Now, the main problem lies in the identification of starting and end events, in order to be able to identify relevant sequences. Therefore, probabilistic automata (see Fig. 9) as well as very large event tacets (delays) between two succeeding events (see Fig. 10) have been considered.

The following references to sequences have been found from the probabilistic automata:

- (Event5 \Rightarrow) Event4 \Rightarrow Event5
- (Event5 \Rightarrow) Event2 \Rightarrow Event3 \Rightarrow Event5
- (Event5 \Rightarrow) Event1 \Rightarrow Event4 \Rightarrow Event5
- (Event5 \Rightarrow) Event4 \Rightarrow Event6 \Rightarrow Event5
- (Event5 \Rightarrow) Event4 \Rightarrow Event5

All cycles end up at 'Event5'.

Using the method that considers very large events tacets between two succeeding events, the following results have been obtained: 'Event5' was identified as a very special event, where the most sequences end up. All other events, with exception of 'Event3', are probably starting events. See Fig. 11 for all detected sequences.

For each sequence a grammatical rule was generated. See Example 7 for an example of a temporal grammatical rule. It describes 'Sequence1' with 'Event2', 'Event3' and 'Event5' as well as their durations and a reference to event tacets between succeeding events.

Example 7: Temporal grammatical rule for 'Sequence1'

A sequence is a

'Sequence1' [40 sec, 64 sec]

if

'Event2': 'no airflow and no respiratory movements without snoring'
[13 sec, 18 sec]

followed by

'Event3': 'no airflow and reduced chest and no abdominal movements
without snoring' [20 sec, 39 sec]

followed after [0,5 sec, 5 sec] **by**

'Event5': 'strong breathing with snoring' [6 sec, 12 sec]

TEMPORAL PATTERNS: Finally, similar sequences will be joined together to a temporal pattern. Therefore, string exchange algorithms have been used, where similarities

between occurring events in the sequences and the duration of the events have been considered, as well. Similarities can be calculated using the following operations on strings: insertion of a character (weight = 2), deletion of a character (weight = 2), conversion of two characters (weight = 3) and replacement of a character (weight = 1). Since for this kind of applications the duration plays an important role, the weights will be increased by one for each 5 sec of duration. Fig. 12 entails two examples for the calculation of the similarity between sequences using string exchange algorithms.

For all the sequences we obtained the following distances (see Table. 2) This lead to four temporal patterns, two of them having two similar sequences (see Fig. 11).

If a temporal pattern entails more than one sequence, then it will be described by an alternative of similar sequences (see Example 7).

Example 7: Temporal grammatical rule for ‘Sequence1’

```

A temporal pattern is a
    ‘Temporal Pattern 2’ [19 sec, 37 sec]
if
    ‘Sequence 2’: [36 sec, 37 sec]
or
    ‘Sequence 3’: [19 sec, 35 sec]

```

We remark here, that no names can be given to sequences and temporal patterns automatically by the system. If possible, i.e. the TCPs are well-defined, names will be given by a domain expert.

5. Results

Knowledge acquisition is one of the main “bottlenecks” in the design of KBS [2]. The presented method can be used as a tool of automated temporal knowledge acquisition for KBS. At this stage of the KBS design and development, a qualitative evaluation of the knowledge base is afforded [3], [26]. Therefore, a questionnaire was developed that enables a structured interviewing of a domain expert for the evaluation of the discovered TCPs, including the identification and rule generation process. As we are dealing with exploratory methods, it should also be verified, if some kind of “new” knowledge was found.

The questionnaire provides a structured interviewing of a medical expert, since the rules describe the knowledge at different abstraction levels. Therefore, we used the figures the discovered successions, events, sequences and temporal patterns (see Chapter 4), as well as

the temporal grammatical rules and the rules of the sig*-algorithm. The following questions have been considered in the questionnaire:

- Were the temporal patterns, sequences and events correctly identified?
- Is the duration of the temporal patterns and sequences in a valid range?
- Are the rules of the temporal grammar understandable and intelligible for an expert?
- Could the expert give names from the domain to the temporal patterns and events?
- Was some kind of “new” knowledge detected?

Compared to expert knowledge the different types of SRBDs and their total number have been correctly identified with a sensitivity of 100% and a specificity of 100%.

However, another important part of the knowledge base evaluation concerns the temporal grammatical rules of the events and their descriptions. Therefore, the successions in the events have been examined within the questionnaire. The expert could give the following answers: yes (successions must occur in the event), maybe (successions may occur in the event), don't know (no statement can be made) and no (successions must not occur in the event). Altogether, 56,25% of more or less occurring successions were correctly associated to the events (yes). This corresponds to events that allocate 80,16% of the whole time period. For 37,5% of more or less occurring successions the expert couldn't give an exact answer (don't know). This merely corresponds to events that allocate 18,21% of the whole time period. The latter is mainly due to the occurrence of tacets in the rules, that don't have a meaning in medicine. For 6,25% the expert couldn't give an exact answer (maybe), corresponding to 1,63% of the whole time period. It is remarkable, that none of the events was incorrectly described (answer: no). An evaluation of the rules at this level lead to an overall sensitivity of 0,762 and a specificity 0,758 (see Table. 1). This is due to the fact that the expert couldn't give an exact answer (answer maybe or don't know) referring to some of the event rules.

‘Event5’ was correctly identified as a special event, called ‘hyperpnea’. SRBDs always end up with a ‘hyperpnea’. In some cases the duration of ‘Event5’ was too short. The duration of all other events were in a valid range. The interpretation of the rules was straightforward, since all rules were regarded as well understandable descriptions due to the tokens used in the grammatical rules, such as ‘more or less simultaneous’, ‘or’, ‘followed by’ and ‘followed after ... by’, and due to the intelligible names generated for the primitive patterns and events.

In order to verify the extracted knowledge, the expert was asked to give a name from the domain to all the temporal patterns and events:

- TemporalPattern1: ‘mixed (obstructive) apnoe’
- TemporalPattern2: ‘obstructive hypopnoe’
- TemporalPattern3: ‘mixed obstructive apnoe’
- TemporalPattern4: ‘obstructive snoring’
- Event1: ‘obstructive snoring’
- Event2: ‘central part (of a mixed obstructive apnoe)’
- Event3: ‘obstructive part (of a mixed obstructive apnoe)’
- Event4: ‘obstructive apnoe’
- Event5: ‘hyperpnoe’
- Event6: ‘hypopnoe’

This means that all discovered TCPs could be associated to important properties of SRBDs from medicine.

The last question was related to the detection of possibly “new” knowledge in the data. A reference was found for ‘*TemporalPattern3: mixed obstructive apnoe*’ that entails two different types of patterns not yet well defined in medicine:

- *mixed obstructive apnoe* with an interruption and snoring having a central and an obstructive part
- *mixed obstructive apnoe* without an interruption and without snoring ending in an hypoventilation

The latter has to be verified with further experiments.

6. Discussion

Systems that integrate several intelligent technologies, for instance, from traditional AI, ANNs and Genetic Algorithms, are called Hybrid Systems (HS) [13]. However, most HS can be regarded as merely cooperative HS, where, for instance, the main component of a technology in the HS is replaced by another (function-replacing HS), several independent modules are implemented in different technologies (cooperative HS) or several technologies are simulated by just one technology (polymorphic HS). In contrast, we are mainly interested in HS, where a knowledge conversion from sub-symbolic knowledge into a symbolic knowledge takes place [39]. HS with a knowledge conversion have already been successfully

applied to multi-dimensional data from medicine with data gathered from patients with blood anaemia, acidosis, and cerebrospinal fluid diseases [42], [41].

This work extends this approach to multivariate time series, being so far much more complex, since several abstraction levels have to be introduced, as proposed in [14]. At each level temporal rules are extracted automatically from the results of the unsupervised ANNs and other exploratory methods, such that a step-wise knowledge conversion of multivariate time series in form of an automated temporal knowledge acquisition becomes feasible.

The acquisition of expert knowledge is a hard task, since a computer-based processing of the knowledge, for instance, through an inference machine, must be achieved and the acquired knowledge should be complete and extensive, as well. For very large data bases such a process might be very prolonged [24]. Usually tools for an interactive knowledge acquisition are used [8]. However, the usually existent lack of time of domain experts often limits those tools for the development of KBS [9]. In addition, expert knowledge and experience is hardly communicable and convertible into rule-based knowledge. Experts often aim to describe exceptions or, on the other side, trivial rules. Frequently, associations to similar cases are searched for. The use of ANNs for such tasks is quite evident, since ANNs have the capacity to learn from examples and operate as an associative memory.

Previous approaches that generate a knowledge acquisition for time series are limited to univariate time series, such as ECGs [22] or carotid pulse waves [4]. Here, the main patterns in the time series are pre-defined, for instance, classifying P-waves or QRS-complex of an ECG signal or using simple waveform operations as local minimum or negative slope. The main difference to our approach lies in the fact that we discover the main patterns in multivariate time series using unsupervised methods.

An analysis and description of multivariate time series not only affords the development of a knowledge representation that describes succeeding events in the time series, but also the introduction of simultaneity. In medicine, usually, overlapping intervals are described using Allen's (1984) interval calculus [10], [12], [18]. Here, 14 temporal relations between intervals are precisely described considering different starting and end points as well as precedence relations between two intervals. However, for pattern recognition tasks this approach would lead to a enormous amount of patterns, such that a more vague representation, as used in Kowalsky and Sergot (1986) event calculus seems more appropriate. Therefore, we introduced a "more or less simultaneity" that enables an extraction and description of the main patterns in SRBDs.

7. Conclusion

This paper presents a recently developed method for temporal knowledge conversion (TCon) [14],[16] applied to sleep-related breathing disorders (SRBDs). Therefore, 25 Hz recorded data of a polysomnography from three patients have been used having the most frequent SRBDs. In order to handle this complex and structured problem, several abstraction levels have been introduced allowing a successive detection of SRBDs, i.e. temporal complex patterns in multivariate time series. It was shown that Self-organized Feature Maps (SOMs), as proposed by Kohonen [19], together with a special visualization technique, called U-Matrix-Method [37], have been successfully applied for the detection of elementary patterns in time series, called primitive patterns. Features obtained from the time series *'airflow'*, *'ribcage movements'* and *'abdominal movements'* and *'snoring'* have been used as input to the SOMs. This step is a prerequisite for further processing, specially, when dealing with data from several patients. The detection of primitive patterns represents the first step in a chain of steps with several abstraction levels until the final detection of SRBDs and can be seen as a step towards a successive higher knowledge abstraction.

Another main aim was the extraction of temporal knowledge, understandable for human beings. At the first level, a machine learning algorithm, called sig* [38], was used to generate a rule-based description of the detected primitive patterns in form of PROLOG rules. One advantage lies in the generation of an intelligible name for each primitive pattern that can be used at the next abstraction levels. Temporal grammatical rules are generated for the description of the patterns at higher abstraction levels, as events, sequences and temporal patterns. An implementation of those rules with definitive clause grammars is straightforward [15]. All rules represent an automatically generated temporal knowledge base of a knowledge base system. In addition, a linguistic representation of the detected patterns at different abstraction levels is generated.

The evaluation of the extracted knowledge in form of decision rules was made in a structured and step-wise form using an especially for this purpose conceptualized questionnaire. Altogether, all four SRBDs have been correctly identified and for the first time a linguistic description of SRBDs was automatically generated using machine learning algorithms. Such a representation allowed an evaluation of the discovered SRBDs and a verification of the extracted temporal knowledge by a domain expert. The presented results show some of the potentials of our approach. Our next goal will be the application and extension of the method to a greater number of time series and to other domains, where

temporal complex patterns are searched. In addition, further research will be needed to verify the possibly “new” kind of knowledge found in one of the SRBDs.

References

- [1] Aho, A.V., Algorithms for Finding Patterns in Strings, in: J.V. Leeuwen, ed., *Algorithms and Complexity: Handbook of Theoretical Computer Science*, Vol. A, (Elsevier Science Publisher, Amsterdam, 1990).
- [2] Andrews, R., Diederich, J., Tickle, A.B. Survey and critique of techniques for extracting rules from trained artificial neural networks, *Knowledge-Based Systems*, Volume 8, Number 6, December, (Elsevier Science Publisher, 1995) 373-389.
- [3] Beach, S.S., Gevarter, W., Standarts for Evaluating Expert Systems Tools, *Expert Systems with Applications*, Vol. 2 (1991) 259 - 267.
- [4] Bezdek, J.C., Hybrid modeling in pattern recognition and control, *Knowledge-Based Systems*, Vol. 8, Number 6 (Elsevier Science Publisher, 1995) 359-371.
- [5] Bishop, C.M., *Neural Networks for Pattern Recognition* (Oxford, Clarendon Press, 1995).
- [6] Bock, H.H., *Clustering and Neural Networks*, in: 6th Conf. of the Intl. Fed. of Classification Societies, Rom (1998).
- [7] Bukharaev, R.G., *Theorie der stochastischen Automaten*. (Teubner Verlag, Stuttgart, 1995).
- [8] Bürsner, S., Spreckelsen, C., Nebe, J., Spitzer, K., Epistemologische Adäquatheit bei der Akquisition und formalen Repräsentation neurologischen Wissens, *Informatik, Biometrie und Epidemiologie in Medizin und Biologie* 26 (39), (Eugen Ulmer Verlag, Stuttgart, 1995) 229-239.
- [9] Cimino, J.J., Barnett, G.O., Automatic Knowledge Acquisition from MEDLINE, *Methods of Information in Medicine* (Schattauer Verlag, 1993) 120-130.
- [10] Combi, C., Pincioli, F., Pozzi, G., Managing Different Time Granularities of Clinical Information by an Interval-based Temporal Data Model, *Methods of Information in Medicine* 34 (Schattauer Verlag, 1995) 458-474.
- [11] Fukunaga, K., *Introduction to Statistical Pattern Recognition*. 2nd Edition (Academic Press, Inc., Boston, 1990).
- [12] Gamper, J., Nejdil, W., Abstract temporal diagnosis in medical domains, *Artificial Intelligence in Medicine* 10 (Elsevier Science Publisher, 1997) 209-234.

- [13] Goonatilake, S., Khebal, S., *Intelligent Hybrid Systems* (Wiley & Sons, New York, 1995).
- [14] Guimarães, G., Eine Methode zur Entdeckung von komplexen Mustern in Zeitreihen mit Neuronalen Netzen and deren Überführung in eine symbolische Wissenrepräsentation. PhD Dissertation, Philipps University of Marburg, Germany, 1998.
- [15] Guimarães, G., Ultsch, A., A Symbolic Description for Patterns using Definitive Clause Grammars, in: Klar, R., Opitz, O., eds., *Classification and Knowledge Organization*, Procs. of the 20th Annual Conf. of the Gesellschaft für Klassifikation, March 6-8 1996, University of Freiburg (1997) 105-111.
- [16] Guimarães, G., Ultsch, A. (1999): A Method for Temporal Knowledge Conversion, in: D.J. Hand, J.N.Kok, M.R. Berthold, eds., *Advances in Intelligent Data Analysis*. Procs. of IDA99, The Third Symposium on Intelligent Data Analysis, August 9-11, Amsterdam, Netherlands, Lecture Notes in Computer Science 1642 (Springer, 1999) 369-380.
- [17] Kaski, S., Kohonen, T., Exploratory Data Analysis by Self-Organizing Map: Structures of Welfare and Poverty in the World, in. A.P.N Refenes, Y. Abu-Mostafa, J. Moody, A. Weigend, eds., *Neural Networks in Financial Engineering*. Proc. of the Intl. Conf. on Neural Networks in the Capital Markets, London, England, 11-13 October, 1995, Singapore (1996) 498 – 507.
- [18] Haimowitz, I.J., Kohane, I.S., Managing temporal worlds for medical trend diagnosis, *Artificial Intelligence in Medicine* 8, (Elsevier Science Publisher, 1996) 299-321
- [19] Jackson, P.: Introduction to expert systems. Addison-Wesley, 1986.
- [20] Kohonen, T., Self-organized formation of topologically correct feature maps, *Biological Cybernetics* 43 (1982)141-152.
- [21] Kohonen, T., *Self-Organizing Maps* (Springer Verlag, New York, 1995).
- [22] Koski, A., Juhola, M. and Meriste, M., Syntactic recognition of ECG signals by attributed finite automata, *Pattern Recognition*, The Journal of the Pattern Recognition Society, ISSN 0031-3203, Volume 28, Issue 12, December (1995) 1927-1940.
- [23] Masters, T., *Neural, Novel & Hybrid Algorithms for Time Series Prediction*, (John Wiley & Sons, New York, 1995).
- [24] McDermott, J., Bachant, J.: R1 Revisited, Four Years in the Trenches, *AI Magazine*, No. 3 (1984) 21-32.
- [25] Michie, D., Spiegelhalter, D.J., Taylor, C.C., *Machine Learning, Neural and Statistical Classification* (Ellis Horwood, 1994).

- [26] Ohmann, C.,Belenky, G., Leitfaden zur Evaluation von Wissensbasen. Bericht des MEDWIS-Arbeitskreis "Evaluation",1996.
- [27] Penzel, T.: Zur Pathophysiologie der Interaktion von Schlaf, Atmung und Kreislauf – Konzepte der Kardiorespiratorischen Polysomnographie. Habilitationsschrift, Philipps Universität Marburg , 1995.
- [28] Penzel, T., Peter, J.H., Design of an Ambulatory Sleep Apnea Recorder, in: H.T. Nagle, W.J. Tompkins, eds., *Case Studies in Medical Instrument Design*, IEEE, New York (1992) 171-179.
- [29] Pereira, F.C.N., Warren, D., Definitive Clause Grammars for Language Analysis - A Survey of the Formalism and a Comparison with Augmented Transition Networks, *Artificial Intelligence* 13 (1980) 231-278.
- [30] Peter, J.H., Chronobiologie und Schlaf, *Internist* 32 (Springer Verlag,1991) 363-379.
- [31] Peter, J.H.: Störungen der Atmungsregulation, in: H. Hornbostel, W. Kaufmann, W. Siegenthaler, eds., *Innere Medizin in Praxis und Klinik*, Band I: Herz, Gefäße, Atmungsorgane, Endokrines System,4 Aufl. (Georg Thieme Verlag, New York, 1991) 262-280.
- [32] Ritter, H.J., Martinetz, T., Schulten, K., *Neural Computation and Self-Organizing Maps: An Introduction* (Addison-Wesley, 1992).
- [33] Rojas, R., *Neural Networks – a Systematic Introduction* (Springer, 1996).
- [34] Rumelhart, D.E., McClelland, J.L., *Parallel Distributed Processing* (MIT Press, Cambridge, 1986).
- [35] Schweizer, M., Föhn, P., Schweizer, J., Ultsch, A., A Hybrid Expert System for Avalanche Forecasting. In: W. Schertler, B. Schmid, a.M. Tjoa, H. Werther, eds., *Informations and Communications Technologies in Tourism* (Springer Verlag, Wien 1994) 148-153.
- [36] Sleep Apnea, Sleepiness, and Driving Risk, America Thoracic Society, Medical Section of the American Lung Association. *American Journal of Respiratory and Critical Care Medicine*,Vol. 150 (1994) 1463-1473.
- [37] Ultsch, A., Self-Organizing Neural Networks for Visualization and Classification. Proc. Conf. Soc. *Information and Classification*, Dortmund, April (1992).
- [38] Ultsch, A., Knowledge Extraction from Self-organizing Neural Networks, in: O. Opitz, B. Lausen and R. Klar, eds., *Information and Classification* (Berlin, Springer Verlag, 1993) 301-306.

- [39] Ultsch, A., The Integration of Neural Networks with Symbolic Knowledge Processing, in: Diday et al., eds., *New Approaches in Classification and Data Analysis* (Springer, 1994) 445-454.
- [40] Ultsch, A., Korus, D., Wehrmann, A., Neural Networks and their Rules Classification in Marine Geology. Raum und Zeit in Umweltinformationssystemen. 9th Intl *Symposium on Computer Science for Environmental Protection CSEP'95*, Vol. I, GI-Fachauschuß 4.6 "Informatik im Umweltschutz", Band 7, Metropolis-Verlag, Marburg (1995) 676-693.
- [41] Ultsch, A., Farsch, S., Li, H., Automatic Acquisition of medical Knowledge from Data Sets with Neural Networks, *Advances in Artificial Intelligence* (Springer, 1995) 285-260.
- [42] Ultsch, A., Kleine, T.O, Korus, D., Farsch, S., Guimarães, G.: Pietzuch, W., Simon, J., Evaluation of Automatic and Manual Knowledge Acquisition for Cerebrospinal Fluid (CSF), in: E. Keravnu et al., eds., *Artificial Intelligence in Medicine*, Lecture Notes in Artificial Intelligence 1211, Vol. 934 (Springer Verlag, 1997) 110-121.
- [43] Zadeh, L.A., Fuzzy Sets. *Journal of Information and Control*, Vol. 8 (1965) 338-353.

Tables

	Sensitivity	Specifity
Event1	0,909	0,719
Event2	0,933	0,717
Event3	1	0,709
Event4	0,509	0,822
Event5	0,467	0,836
Event6	0,754	0,741
All	0,762	0,758

Table. 1. Sensitivity and Specifity for the temporal grammatical rules of the events

Distances	S1	S2	S3	S4	S5	S6
Sequence1 (S1)		14	17	14	16	>20
Sequence2 (S2)			2	6	9	>20
Sequence3 (S3)				8	6	>20

Sequence4 (S4)					2	>20
Sequence5 (S5)						>20
Sequence6 (S6)						

Table. 2. Calculated distances between all the sequences using string exchange algorithms

Figures

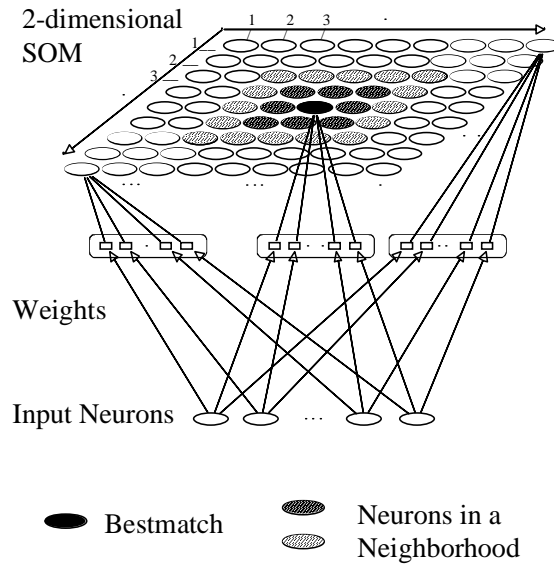


Fig. 1. Two-dimensional Self-organizing Feature Map (SOM) as proposed by Kohonen [19]

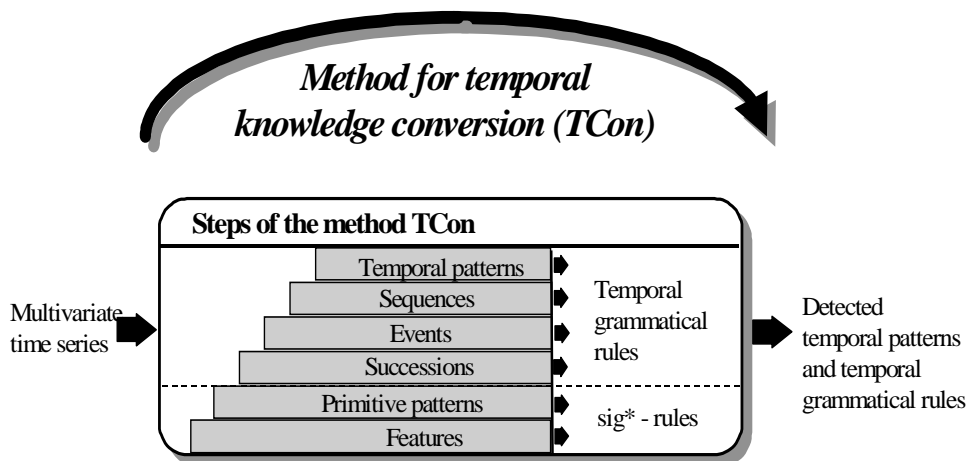
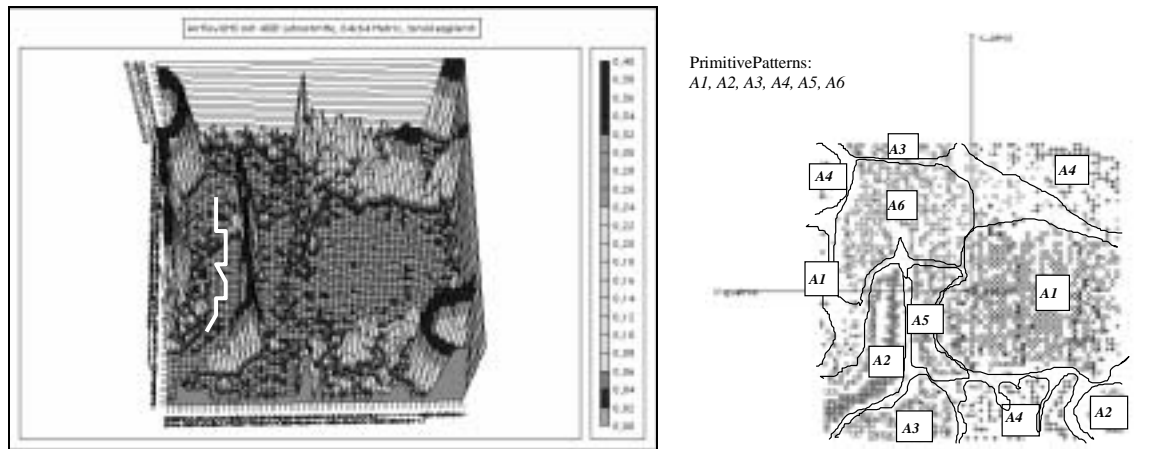
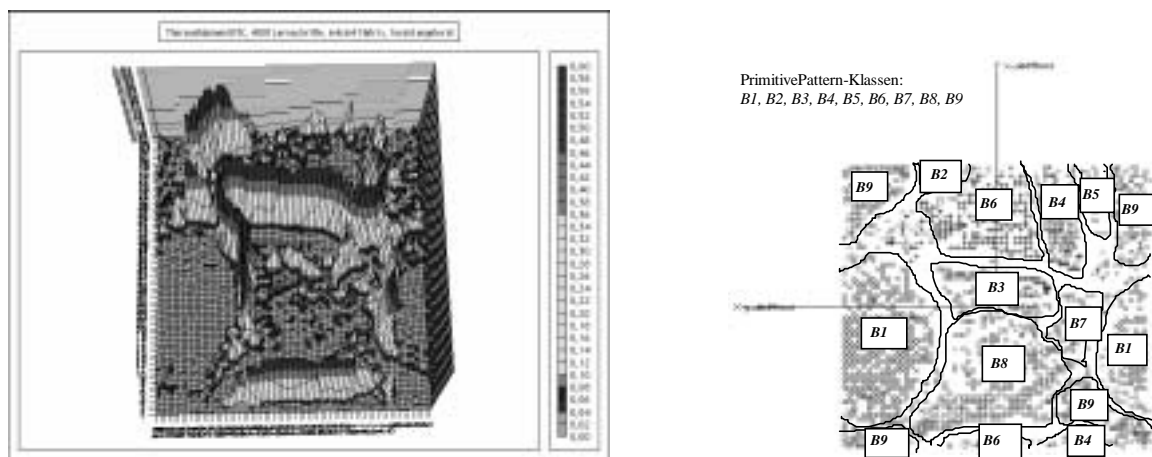


Fig. 2. Abstraction levels and steps of the method TCon enabling a conversion of temporal structures in multivariate time series into a symbolic knowledge representation form [4]



(a)



(b)

Fig. 3. (a) AirflowUMX: U-Matrix generated from features related mainly to the airflow. The trajectory entails a transition from succession A2 to succession A6. A visual analysis of the U-Matrix enabled the detection of 6 primitive pattern classes A_1, \dots, A_6 . (b) ChestAbdomenWallUMX: U-Matrix generated from features related mainly to the respiratory movements. A visual analysis of the U-Matrix enabled the detection of 9 primitive pattern classes B_1, \dots, B_9 .

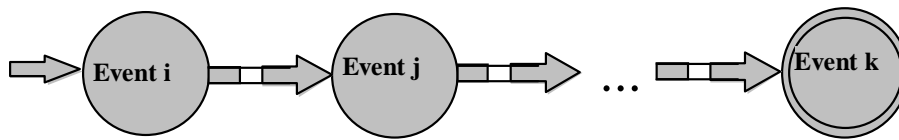


Fig. 4. A possible sequence of events, whereby a transition between events may include a delay.

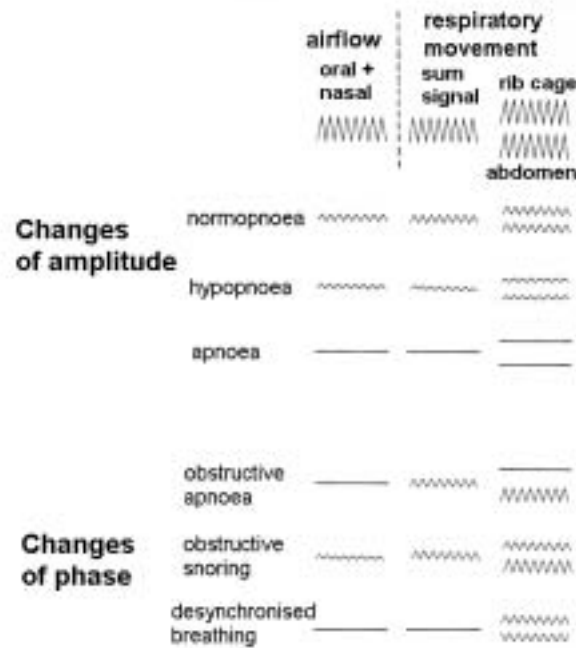


Fig. 5. Disturbances of the signals related to both, amplitude and phase.

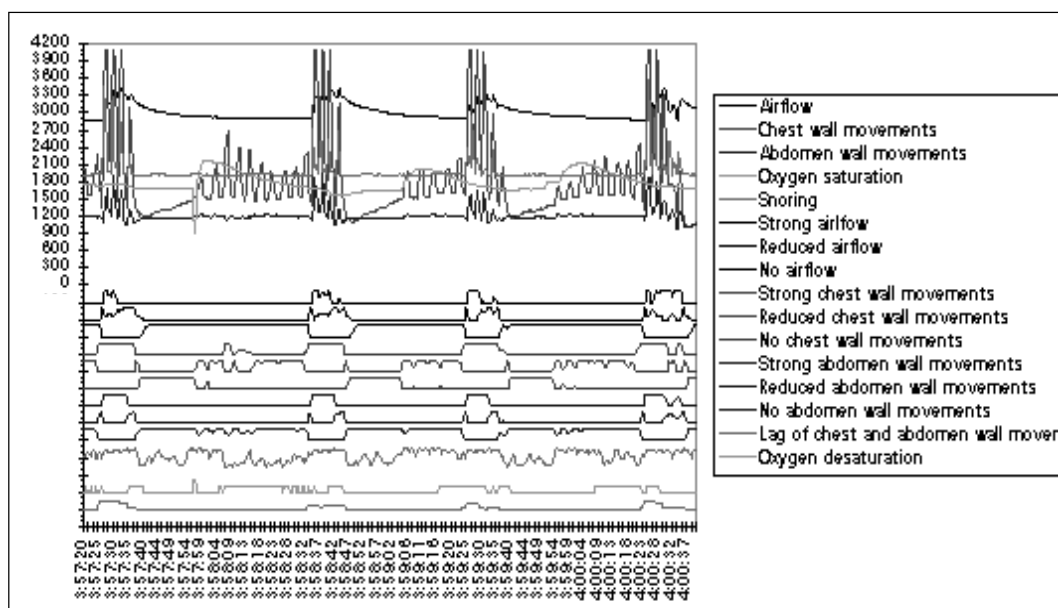


Fig. 6. Small excerpt of multivariate time series and resp. features of a SRBD

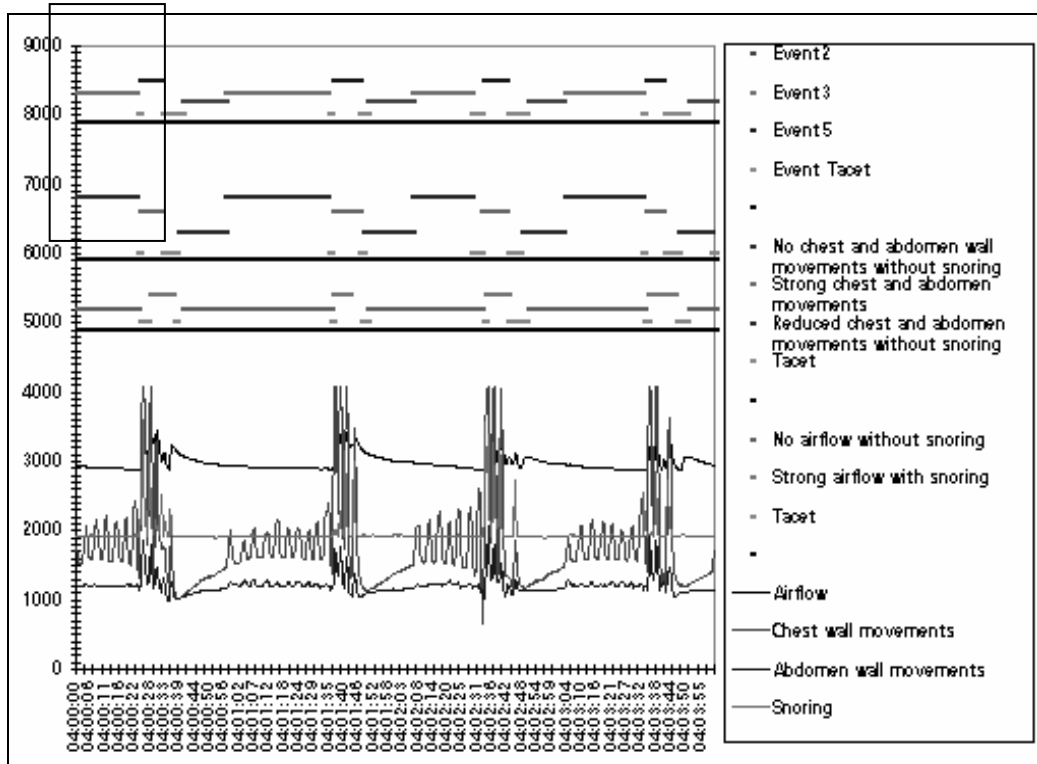


Fig. 7. Multivariate time series and resp. primitive patterns/successions from a patient with SRBDs

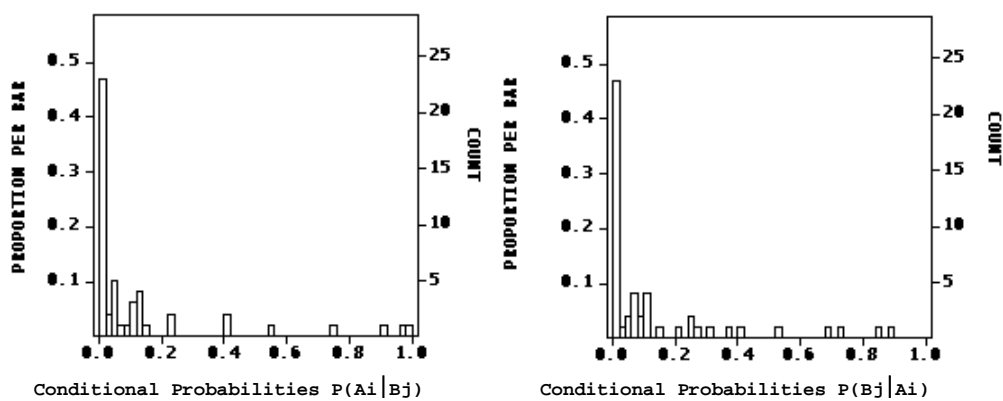


Fig. 8. Conditional Probabilities for the determination of significant and less significant events

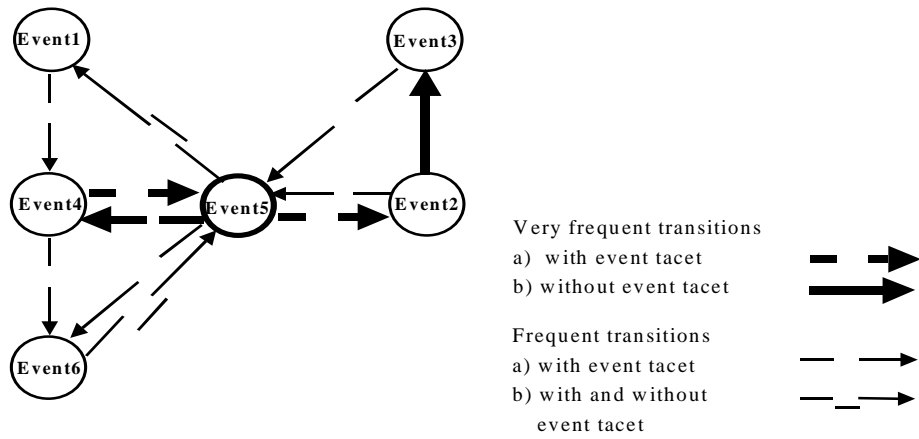


Fig. 9. Probabilistic automata for the determination of starting and ending events

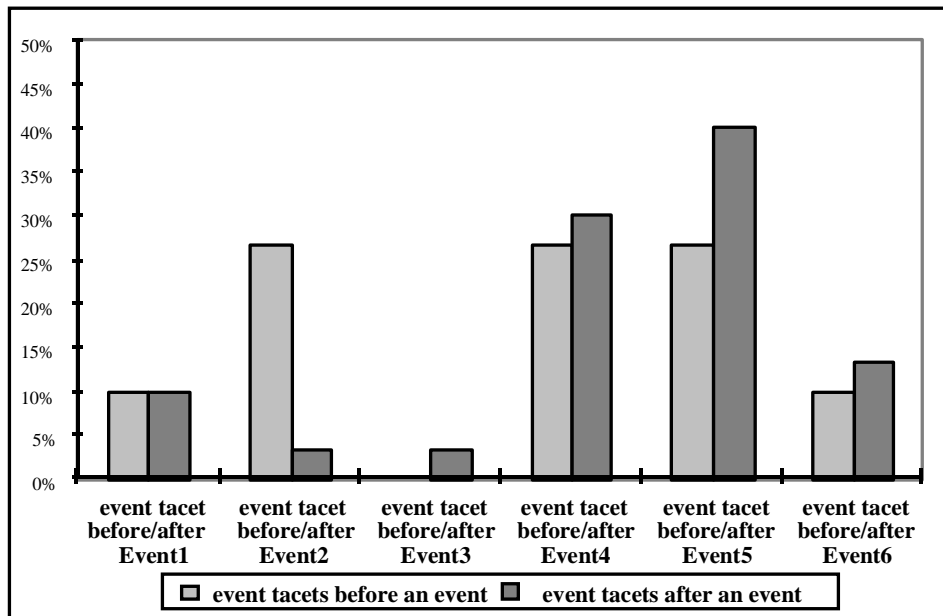


Fig. 10. Number of events before and after a longer event tacet in percents

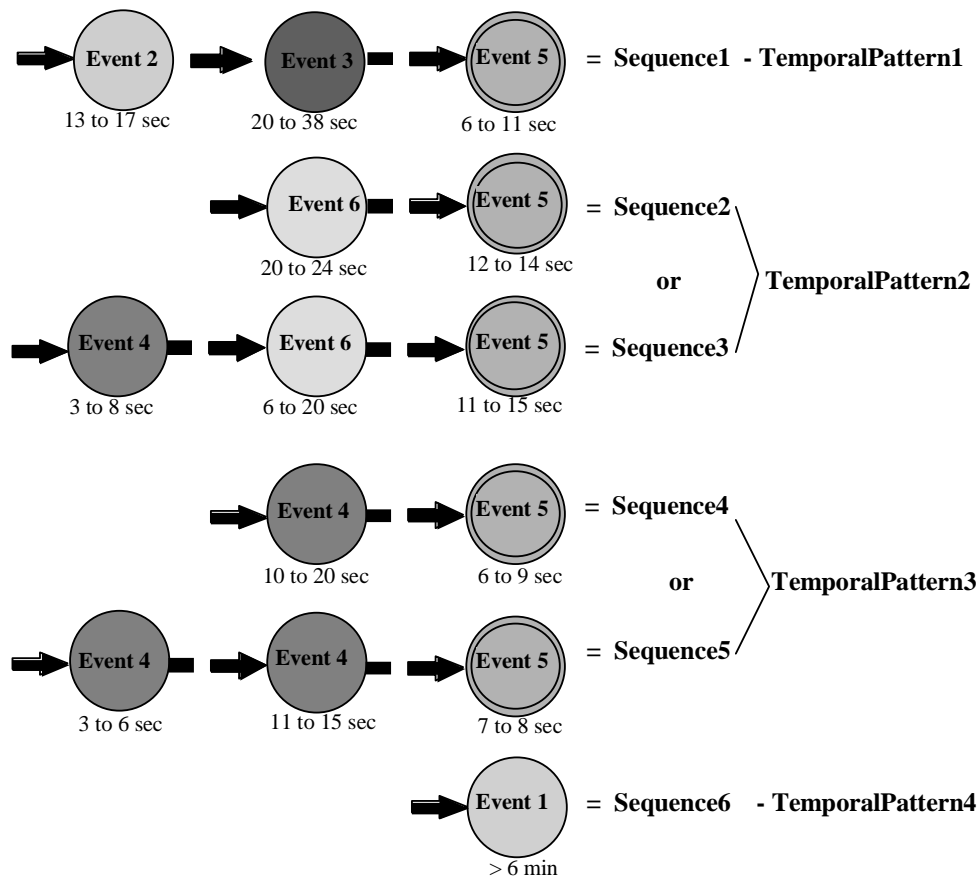
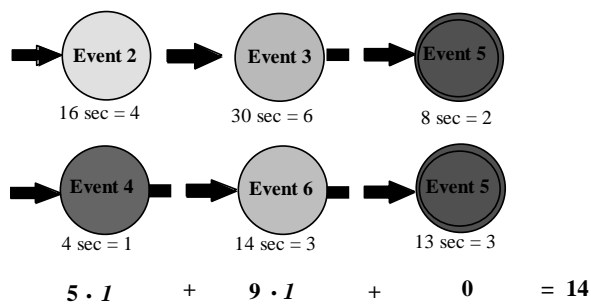


Fig. 11. Temporal Patterns with corresponding sequences and events for all SRBDs



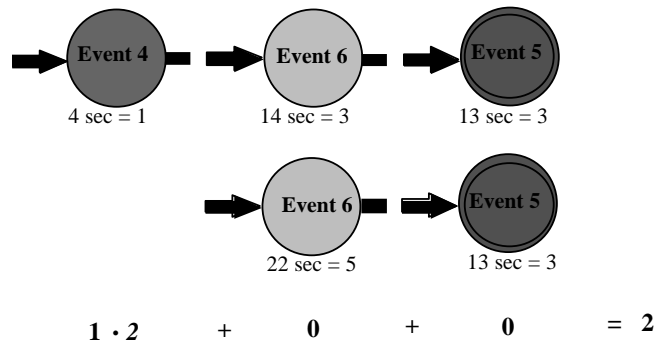


Fig. 12. Two examples for the calculation of the similarity between sequences using string exchange algorithms

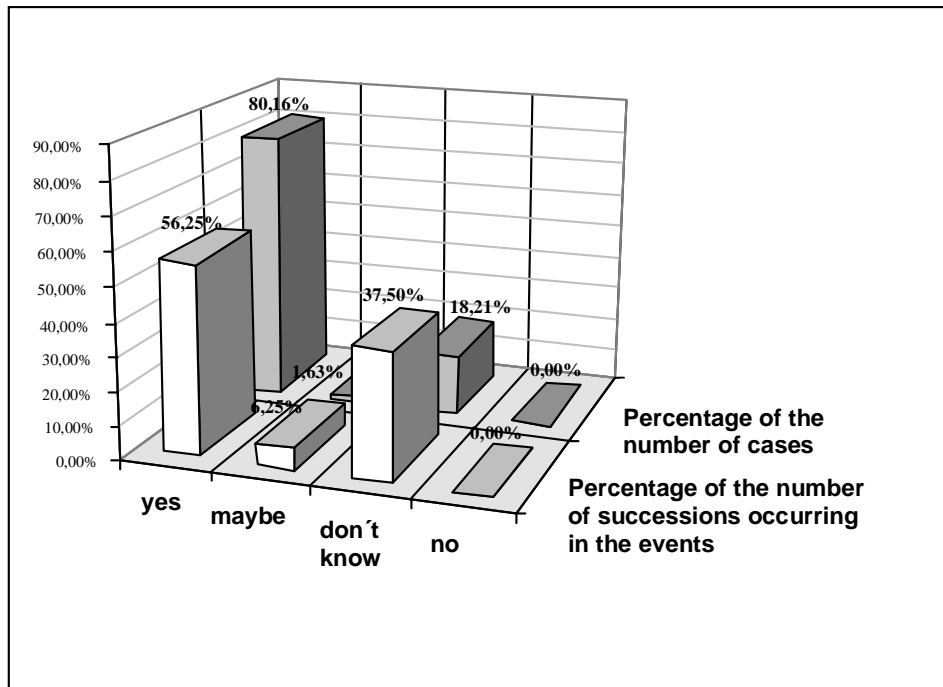


Fig. 13. Results of the questionnaire for the events

