A Method for Temporal Knowledge Conversion

G. Guimarães[1] and A. Ultsch [2]

[1] CENTRIA, U. Nova de Lisboa and
Department of Mathematics, University of Évora
Portugal

[2] Department of Mathematics and Computer Science
Philipps University of Marburg
Germany

Abstract

In this paper we present a new method for temporal knowledge conversion, called TCon. The main aim of our approach is to perform a transition, i.e. conversion, of temporal complex patterns in multivariate time series to a linguistic, for human beings understandable description of the patterns. The main idea for the detection of those complex patterns lies in breaking down a highly structured and complex problem into several subtasks. Therefore, several abstraction levels have been introduced where at each level temporal complex patterns are detected successively using exploratory methods, namely unsupervised neural networks together with special visualization techniques. At each level, temporal grammatical rules are extracted. The method TCon was applied to a problem from medicine, sleep apnea. It is a hard problem since quite different patterns may occur, even for the same patient, as well as the duration of each pattern may differ strongly. Altogether, all patterns have been detected and a meaningfull description of the patterns was generated. Even some kind of "new" knowledge was found.

1. Introduction

In recent years there has been an increasing development towards more powerfull computers, such that nowadays a great amount of data from, for example, industrial processes or medical applications, is gathered. These measured data are often said to be a starting point for an enhanced diagnosis or control of the underlying process. Particularly interesting for handling noisy or inconsistent data are artificial neural networks (ANN). On the other side, systems with traditional artificial intelligence (AI) technologies have been successful in areas like diagnosis, control and planing. The advantages of both technologies are wide-ranging. However, the limits of these approaches, namely the incapacity of ANN to explain their behaviour and on the other hand, the acquisition of knowledge for AI systems, are important problems to be adressed.

Recently, there has been an increased interest in hybrid systems that integrate AI technologies and ANN to solve this kind of problems [2]. It its worth to remark here that essentially hybrid systems have been developed that entail several modules, each implemented in a different technology, and that cooperate with another. In contrast, we are mainly interested in hybrid systems that perform a *knowledge conversion*, i.e. a transition between distinct knowledge representation forms [14]. A *symbolic knowledge representation* of a subject should always be in a linguistic, for human beings understandable form. Examples for linguistic representation forms are natural languages, as German or English, but also predicate logic, mathematical calculus, etc. In contrast, a *subsymbolic knowledge representation* always entails numerous elements as, for example, data points from a time series or neurons and weights in ANN that cooperate in a shared and distributed representation of a symbol.

Previous approaches that realize a knowledge conversion [10, 11, 15, 16] do not consider data with temporal dependences. Temporal knowledge konversion always assumes the existance of temporal data, i.e. time series sampled from signals that describe some process. All sampled values are a temporal subsymbolic knowledge representation of the time series. A *temporal knowledge konversion* is an, eventually, successive conversion of multivariate time series or temporal complex patterns in time series to a linguistic, for human beings understandable representation of the time series, i.e. a temporal symbolic knowledge representation [3].

In this paper, we will introduce a new method that enables a temporal knowledge conversion, called TCon [3]. In order to handle this complex problem, several abstraction levels have been introduced. We applied our method TCon to sleep apnea, namely sleep-related breathing disorders (SRBD). SRBD claim to be a very hard problem since quite different patterns for the same temporal patterns may occur, even for the same patient, as well as the duration of each temporal pattern may differ strongly [8, 9].

2. A Method for Temporal Knowledge Conversion (TCon)

The method TCon enables a conversion from temporal complex patterns (TCP) in multivariate time series to a linguistic, for human beings understandable temporal symbolic representation in form of temporal grammatical rules (see Fig. 1). The main idea for the detection of TCP in multivariate time series lies in breaking down a highly structured and complex problem into several subproblems. The advantage of such a strategy is the resolution of this highly complex problem into several subtasks, now solvable at a more technical level. Therefore, several abstraction levels have been introduced where at each level TCP are detected successively using exploratory methods, namely unsupervised neural networks. The detection process starts with the identification of primitive patterns, i.e. elementary structures in time series. At the following levels, the time dimension will be introduced smoothly in the detection process until the identification of TCP at the last abstraction level is completed.

At the different abstraction levels temporal grammatical rules are generated for a linguistic description of TCP. The advantage of a temporal symbolic knowledge representation in form of temporal grammatical rules is not only the acquisition of a for human beings apropriate representation of the TCP but also the generation of a knowledge representation form that can be processed by a machine engine, like a prolog interpreter. In order to achieve both, the detection of TCP as well as their description at a symbolic level, we suggest a temporal knowledge conversion. Next, we will introduce the different abstraction levels of TCon and give an overview of the tasks.

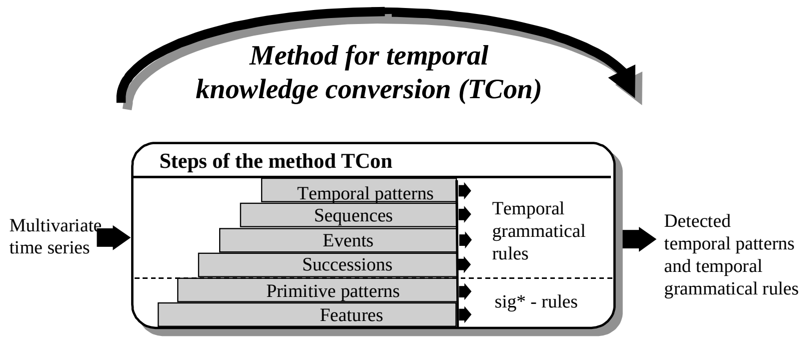

Fig. 1. Abstraction levels and steps of the method Tcon in [3]

Multivariate time series gathered from observed signals of complex processes, as they occur in industrial processes or in medicine, are the input of TCon. We generate a multivariate time series by sampling the observed values at equal time intervals. The result of TCON are the detected TCP as well as a grammatical description of the TCP at the different abstraction levels.

For example, consider a patient with sleep apnea, namely sleep-related breathing disorders (SRBD), where different types of signals, concerning respiratory flow, i.e. ‘*airflow*’, and respiratory effort, are registered during one night [9]. The respiratory effort comprises ‘*chest wall and abdominal wall movements*’. Furthermore, ‘*snoring*’ as well as ‘*oxygen saturation’* are considered for the identification of SRBD. Fig. 2 shows such a registration for a short time period. All time series are sampled at 2.5 Hz. In this paper, we use this example from medicine to illustrate our method.

2.1 Feature extraction and pre-processing

First, an extraction of the main features for all time series is advisable, or even a prerequesite for further processing. Therefore, methods, for instance, from statistics or signal processing are applied to time series in order to find a suitable representation. This process usually includes a pre-processing of the time series such that a clustering with unsupervised neural networks becomes possible. However, 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 [1]. An improvement of the whole performance may be achieved by incorporating prior knowledge, which might be used for the extraction of the features. For each aplication the feature extraction process may differ strongly. Therefore, we will not focus on this issue in this paper. For a detailed description of the feature extraction process see [3]. Nevertheless, it is worth to mention that we considered criteria from the application that usually are applied in sleep laboratories for the identification of SRBD [8].

As multimodal distributions occured in the data for each time series, namely ‘*airflow*’, ‘*chest wall movements’* and *abdominal wall movement’*, fuzzy-membership-functions for *‘no’*, ‘*reduced’* and *‘strong’* averaged amplitude changes have been deduced from histograms. Additionaly, lags between ‘*chest wall movements’* and ‘*abdomen wall movements’* may occur that have a high significance for the identification of SRBD. Therefore, crosscorrelations between ‘*chest and* *abdomen wall movements'* have been calculated. Besides, a rescaling of '*snoring'* was performed. As oxygen saturation is not relevant for the pattern detection process, we just will consider the ocurrence of a decay from at least 4% of the oxygen saturation for the past 10 sec. Altogether, twelve features named as '*strong airflow'* ∈ [0, 1], '*reduced airflow'* ∈ [0, 1], '*no airflow'* ∈ [0, 1], '*strong chest wall movements'* ∈ [0, 1], '*reduced chest wall movements'* ∈ [0, 1], '*no chest wall movements'* ∈ [0, 1], '*strong abdomen wall movements'* ∈ [0, 1], '*reduced abdomen wall movements'* ∈ [0,1], '*no abdomen wall movements'* ∈ [0, 1], '*lag of chest and abdomen wall movements'* ∈ [-1, 1], '*snoring intensity'* ∈ [0, 1] *and* '*oxygen desaturation'* ∈ {0, 0.5, 1} have been extracted (see Fig. 2).

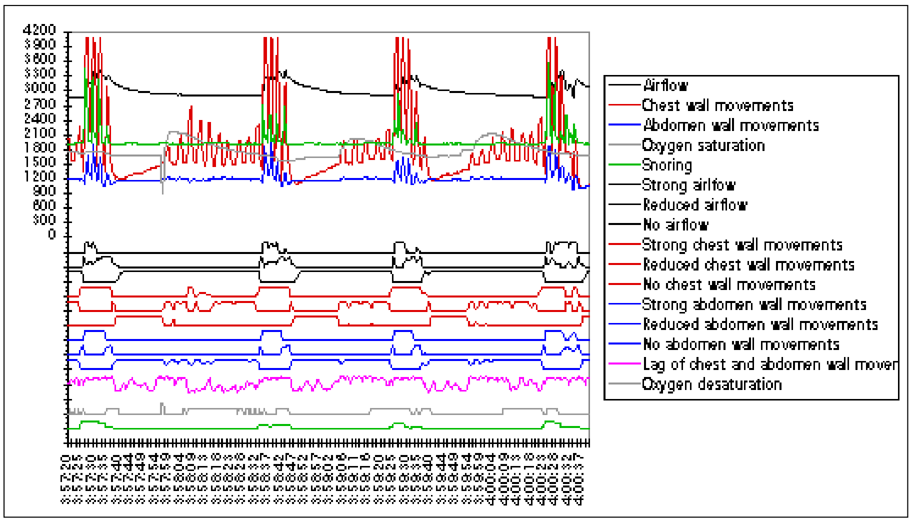

Fig. 2: Small excerpt of multivariate time series and resp. features from a patient with SRBD

2.2 Primitive patterns

At this level, primitive patterns, i.e. elementary structures in the time series, will be determined from the extracted features. Therefore, we propose to use exploratory methods, in particular, unsupervised neural networks like the self-organized neural network (SONN) as proposed by Kohonen [5]. In the last years, SONN enhanced with a special visualization technique, called U-Matrix [12], have been successfully applied to a wide-ranging number of applications where a clustering of high-dimensional data was needed [4, 6, 10, 11, 15, 16].

For the detection of primitive patterns several features have to be selected for the learning process. This means, that we have to identify those features that have a lot in common with regard to criteria from the application. But this also means, that several SONN will be learned to detect primitive patterns from different feature selections. We emphasize that one feature may appear in different feature selections. After the learning process and the identification of the clusters using U-Matrices, we are able to determine primitive patterns. There may appear regions on the U-Matrix that do not correspond to a specified cluster. These regions are regarded as some kind of interruptions, named as tacets. All the other regions are associated to a primitive pattern. As a consequence, we now are able to classify the whole time series with primitive patterns and tacets (see Fig. 3). Successions of primitive patterns from one U-Matrix will be called as a primitive pattern channel.

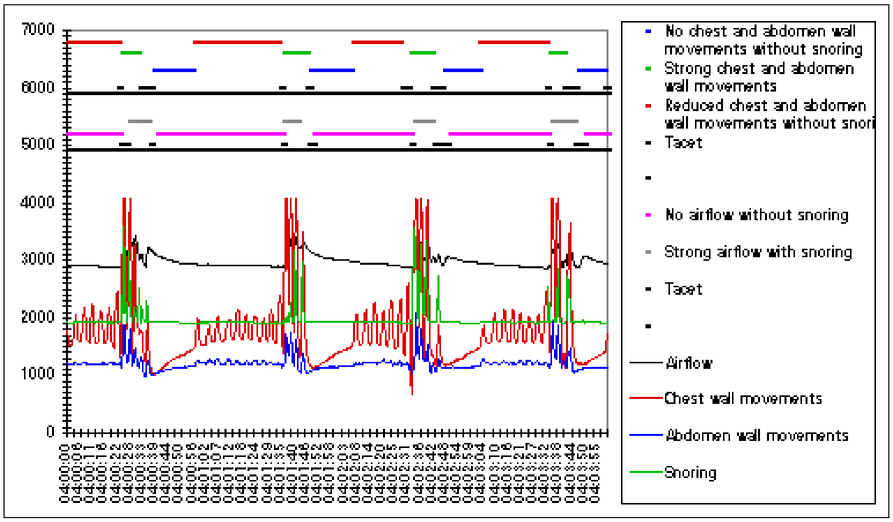

Fig. 3: Multivariate time series and resp. primitive patterns/succession from a patient with SRBD

Without a proper interpretation of the detected structures no meaningfull names can be given to the primitive patterns. As a consequence, we cannot generate for human beings understandable temporal grammatical rules at the next higher levels. In order to achieve a meaningfull description for primitive patterns, we propose the use of machine learning algorithms. For the first time, the rule generation algorithm called sig* [13], was used to generate rules for data in a temporal context. This algorithm selects significant features for each class, here for each primitive pattern, in order to construct appropriate conditions that characterize each class. Example 1 shows sig* rules for two primitive patterns from different feature selections.

Example 1 Consider the primitive patterns '*A2*' and '*B3*' that have been detected from different U-Matrices. The following sig* rules have been generated*:*

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

```
A primitive pattern is a 'B3'
if
      'no chest wall movements'
      ∈ [0.772, 1]
   and
      'no abdomen wall move-
      ments' ∈[0.641, 1]
   and
      'reduced chest wall
      movements' = 0
   and
      'snoring intensity' = 0
```

Values nearby one mean that this feature occurs with a high probabilty, while values nearby zero mean that this feature probably will not occur.

As the sig* algorithm generates rules with the most significant features for each primitive pattern, the naming of the primitive patterns is straightforward. The primitive pattern *'A2'* was named *as 'no airflow without snoring'* and *'B3'* named as '*no chest and abdomen wall movements without snoring'*. These semi-automatically generated names will be used further on for the description of the temporal grammatical rules.

2.3 Successions

At this level, we introduce the dimension time where succeeding identical primitive patterns are regarded as a succession (see Fig. 3). Each succession has a correspondend primitive pattern type. The main diference lies in the fact that a succession additionaly has a start and end point. Successions may be identified by trajectories visualized on U-Matrices. We will not focus on this issue in this paper.

A consequence of several feature selections is that several SONN will be learned and, therefore, several U-Matrices will be generated. This means that two or more successions may occur more or less simultaneously. Two overlapping successions are said to occur more or less simultaneously, if and only if the deviation between their start and end points is small enough, i.e. very small.

2.4 Events

At this level, more or less simultaneous successions are joined together to a new unity, called event. In order to focus on the most significant events, we distinguish between events that occur very frequently and those that occur less frequently. Rare events are omitted in the sense that they are regarded as interruptions. These will be named as event tacets. The idea is to select the most frequent events as the most significant events. Then, less frequent events can be associated to them. Similarities among the successions have been considered to join different types of more or less simultaneous successions, i.e. frequent and less frequent events [3]. This means that the number of events will be extremely reduced and that one event consists of different types of more or less simultaneous successions, i.e frequent and less frequent events, as well.

As a consequence, at this abstraction level temporal grammatical rules not only entail a "more or less simultaneous" but also an "or" for the description of alternations of more or less frequent events. Let us consider the example of the patient with SRBD. In this case, three events have been detected (see Fig. 4).

Names of events can been derived straightforward from the generated grammatical rules, as the names for primitive patterns, i.e. successions, are already known (see Example 2). For a detailed description of the whole detection process and generation of the grammatical rules for the events see [3].

Example 2 The following grammatical rules have been generated for *'Event1'* and *'Event3'*:

```
An event is a 'Event1'
if
      'no airflow without
      snoring'
 is more or less simultaneous
      ('no chest and abdomen
      wall
      movements without snoring'
   and
      'tacets')
```

```
An event is a 'Event3'
if
      ('strong airflow with
      snoring'
   and
      'reduced airflow with
      snoring'
   and
      'tacets')
 is more or less simultaneous
      'strong chest and
      abdomen wall movements'
```

The ocurrence of tacets in the rules means that small interruptions in successions may occur or that a succession, for example, from one primitve pattern channel occurs simultaneously with irrelevant information at the other channel. A name of an event contains essentially names of the most frequently ocurring successions. Names of rarely occuring successions may be dismissed, since the idea of temporal knowledge conversion also entails an information reduction for the generation of well-understandable rules. If needed, details may then be consulted at lower abstraction levels. The following names have been derived from the rules:

'Event1': *'no airflow and no chest and abdomen wall movements without snoring'*

'Event2': *'no airflow and reduced chest wall movements and no abdomen wall movements without snoring'*

'Event3': *'strong breathing with snoring'*

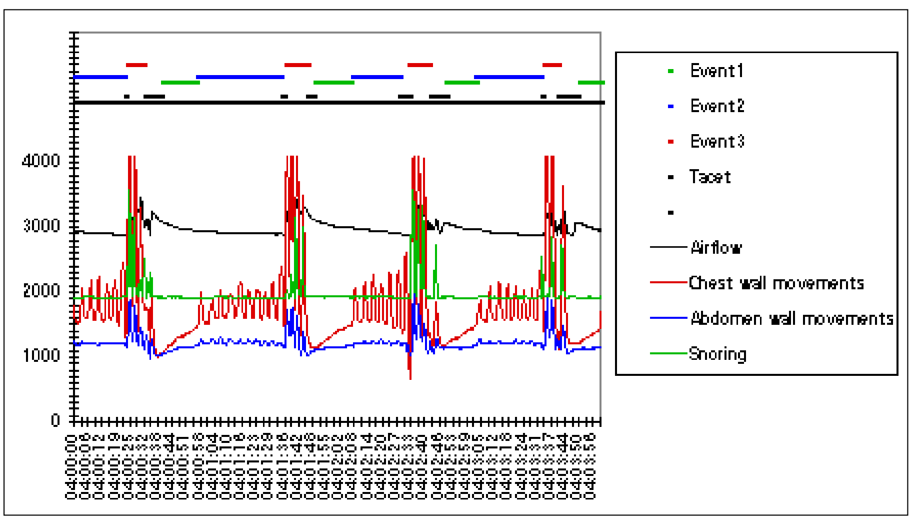

Fig. 4: Multivariate time series and resp. events from a patient with SRBD

2.5 Sequences

At a symbolic representation level, an event may be interpreted as a symbol in a temporal context that cannot be further decomposed. Then, at this abstraction level a multivariate time series can be represented as a sequence of symbols, i.e. events. In order to be able to detect TCP in multivariate time series, we just have to identify repeated subsequences of events. The main problem lies in the identification of start and end events, in particular, when dealing with time series that entail several and distinct TCP. Therefore, we constructed a probabilistic automat as well as considered delays between the ocurrence of two different events. For a details see [3].

A sequence of events together with the multivariate time series from the patient with SRBD is illustrated in fig. 4. Here, we identified the following sequence (see Fig. 5) where '*Event2*' follows immediately after '*Event1*' and '*Event3*' follows '*Event2*' after a small interruption.

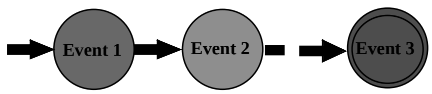

Fig. 5. A detected sequence from the patient with SRBD

For the generation of the grammatical rules we introduced at this level a "followed by" and "followed after" interval "by". As a sequence always occurs more than once in a multivariate time series, lower und upper boundaries for the duration of the events and sequences may be specified (see Example 3).

Example 3 The following gramatical rule has been generated for 'Sequence1'

```
A sequence is a 'Sequence1' [40 sec, 64 sec]
if
      'Event1': 'no airflow and no chest and abdomen wall movements
      without snoring' [13 sec, 18 sec]
  followed by
      'Event2': 'no airflow and reduced chest and no abdomen wall
      movements without snoring' [20 sec, 39 sec]
  followed after [0,5 sec, 5 sec] by
      'Event3': 'strong breathing with snoring' [6 sec, 12 sec]
```

2.6 Temporal patterns

Finally, similar sequences will be joined together to a temporal pattern. Therefore, similarities between the ocurring events in the sequences as well as the duration of the events have been considered [3]. As the example of the patient with SRBD just contains one sequence, the temporal pattern also just has one sequence. Otherwise, the temporal pattern would be described by an alternation of sequences using an "or".

3. Conclusion

Recently, different kinds of hybrid systems that integrate AI technologies and neural networks have been developed [2]. We emphasize that above all "cooperative" hybrid systems have been developed, i.e. a cooperation between several modules implemented in different technologies exists. The main difference to our approach is that in cooperative hybrid system no transition between different knowledge representation form takes place [14]. The hybrid system WINA [7] is an example for a hybrid system where a knowledge conversion is realized. This work was the starting point for the recently developed method for temporal knowledge conversion (TCON) [3].

The main issue of the present paper was to give a brief description of the different abstraction levels introduced by the method TCON. This approach enables a successively and, even, smoothly conversion of temporal complex patterns in multivariate time series to a linguistic, for human beings understandable temporal symbolic knowledge representation in form of temporal grammatical rules. In order to detect elementary structures in the time series, self-organized neural networks, as proposed by Kohonen [5], together with special visualization techniques, called U-Matrices [12], have been used. The realization of the tasks at each level as well as the generation of the temporal grammatical rules was illustrated through an example from medicine, namely sleep-related breathing disorders (SRBD) [9]. SRBD claim to be a very hard problem since quite different patterns for the same temporal patterns may occur, even for one patient. Additionally, the duration of each temporal pattern can differ a lot.

For a lack of space we could just give an overview of the method and present a small example of our experiments with SRBD. We used a much larger data base with the most significant, i.e. most frequently ocurring, SRBD. For details see [3]. Altogether, we detected all temporal patterns with our method TCON and were able to give a, for an expert of SRBD, meaningfull description of the temporal patterns with the temporal grammatical rules. Additionally, some kind of “new” knowledge for one temporal pattern, i.e. some not yet well-described SRBD in medicine, have been found.

Aknowledgements

We thank Prof. Dr. J. H. Peter and Dr. T. Penzel, Medizinische Poliklinik, Philipps University of Marburg for making us the data available.

References

1. Bishop, C.M.: Neural Networks for Pattern Recognition. Oxford, Clarendon Press, 1995.
2. Goonatilake, S., Khebal, S. (Eds.): Intelligent Hybrid Systems, Wiley & Sons, New York, 1995.
3. 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, Marburg, Germany,1998.
4. 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: 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, pp. 498 – 507.

5. Kohonen, T.: Self-Organization and Associative Memory. Springer Series in Information Sciences 8, Springer Verlag, New York, 1984.

6. Kohonen, T.: Self-Organizing Maps, Springer Verlag, New York, 1995.

7. Palm, G., Ultsch, A., Goser, K., Rückert, U.: Knowledge Processing in Neural Architecture. In: VLSI for Neural Networks and Artificial Intelligence, New York, 1994, pp. 207 – 216.

8. Penzel, T.: Zur Pathophysiologie der Interaktion von Schlaf, Atmung und Kreislauf – Konzepte der Kardiorespiratorischen Polysomnographie. Habilitationsschrift, Philipps Universität Marburg, 1995.

9. 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, pp. 171-179.

10. 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, pp. 148-153.

11. Ultsch, A.: Konnektionistische Modelle und ihre Integration mit wissensbasierten Systemen. Forschungsbericht Nr. 396, Institut für Informatik, Universität Dortmund, Februar, Dortmund, Habilitationsschrift, 1991.

12. Ultsch, A.: Self-Organizing Neural Networks for Visualization and Classification. In: Proc. Conf. Soc. Information and Classification, Dortmund, April, 1992.

13. 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.

14. Ultsch, A.: The Integration of Neural Networks with Symbolic Knowledge Processing. In: Diday et al.: New Approaches in Classification and Data Analysis, Springer, 1994, pp. 445-454.

15. Ultsch, A., Korus, D., Wehrmann, A.: Neural Networks and their Rules Classification in Marine Geology. In: Raum und Zeit in Umweltinformationssystemen, 9th Intl Symposium on Computer Science for Environmental Protection CSEP'95, Vol. I, GI-Fachausschuß 4.6 "Informatik im Umweltschutz", Band 7, Metropolis-Verlag, Marburg, 1995, pp.676-693.

16. 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, pp. 110-121.