Mälardalen University Licentiate Thesis
No.28

A Case-Based Approach for Classification of Physiological Time-Series

Markus Nilsson

May 2004

MÄLARDALEN UNIVERSITY

Department of Computer Science and Engineering
Mälardalen University
Västerås, Sweden

ISSN 1651-9256
ISBN 91-88834-47-6
Printed by Arkitektkopia, Västerås, Sweden
Distribution: Mälardalen University Press

"CBR is Just Hype"

– Ian Watson[1]

[1] Chapter 9.2.4. In Applying Case-Based Reasoning: Techniques for Enterprise Systems

Abstract

Complex measurement classification is often difficult, as in the medical domain, and it usually takes a long time to fully master all aspects involved. An automated measurement classification system would ease the diagnostic process for treatment personnel, especially for less experienced clinicians. This thesis contains results from research in the field of Artificial Intelligence (AI) applied to medical measurement classifications. Artificial Intelligence may be described as a variety of computational methods and techniques that exhibit intelligent behaviour. These methods and techniques enable problem solving comparable to humans. The thesis presents a novel approach for multiple time-series analysis based on Case-Based Reasoning (CBR). CBR is an AI method based on a plausible cognitive model of human reasoning. The approach analyses parallel streams of measurements and uses CBR as well as other AI methods for classification and domain reduction. The approach is implemented as a system for classification of Respiratory Sinus Arrhythmia (RSA). The time-series are composed of physiological measurements as the system identifies dysfunctions within the RSA. RSA is identified by analysing the heart and the pulmonary systems of the human body. The developed system, named HR3modul, functions as a decision support tool for treatment personnel in the field of psychophysiological medicine. A classification proposal is presented to the user. The proposal is based on stored knowledge and current physiological time-series.

To Tuula, Leif, and Mikael

Preface

$E = mc^3$

Markus Nilsson
Västerås, March 31, 2004

Contents

I

Thesis

Chapter 1

Introduction

Classification of complex measurements is essential in many diagnostic tasks. Correct classification of measurements may in fact be the most critical part of the diagnostic process. Diagnosing psychophysiological dysfunctions is a medical field where measurement classification is difficult and complex.

Clinicians use sensors to measure physiological parameters in order to diagnose and treat stress related psychophysiological dysfunctions. The clinicians make a manual classification of the measurements before they can make a reliable diagnosis. This is difficult, even for experienced clinicians. A system with an ability to automatically classify these measurements would ease the clinicians work and increase the reliability of the diagnosis, especially for less experienced clinicians. Case-based reasoning (CBR) is a concept that is recognized in medical domains and has been applied in a number of medical diagnostic projects [1, 2, 3], but not for direct classification of complex sensor readings from patients, as in this research. CBR is an Artificial Intelligens (AI) method and is described in section 1.2. Section 1.3 contains further information on psychophysiology.

1.1 Motivation

The motivation can be viewed from both a computational and a medical side. The reader may choose which s/he prefers as both are motivated.

AI point of view

The medical domain is often regarded as one of the most difficult domains for a computer system to analyse. Case-Based Reasoning (CBR) is an interesting AI method for building medical applications. One of the more intuitively and attractive features of CBR in medicine is that the concepts of *patient* and *disease* lend itself naturally to a case representation.

Another attractive feature is the knowledge storage. Knowledge is not lost as it is stored as an individual case. The entire case may be retrieved at any time. This aspect is pleasing for a physician as the AI system may explain its reasoning by saying *I believe this is a good solution, as it is similar to this already solved case I have in my memory.* Hence, no uncertainty exists on how the system came to its conclusion.

This research has provided a novel approach to the classification of complex measurements. The new AI system makes classifications of complex measurements in a domain without the availability of explicit domain knowledge, and the system is able to sugest solutions based on a sparse number of examples. The system handles multiple physiological time-series in order to classify physilogical dysfunctions.

Medical point of view

There is a limited number of experts working in psychophysiological medicine since psychophysiology is a relatiely new area in medical community. Hence, it is not widely known how to make an accurate diagnosis. A general practitioner could possibly make a basic diagnosis in psychophysiology if he or she had a decision support system to aid them.

We have concentrated our research on a decision support system for classification on Respiratory Sinus Arhhythmia (RSA). RSA is one of the more important factors in a psychophysiological diagnosis. RSA is described in section 1.3.2. We have also developed a research tool for the psychophysiological domain. A research tool enables the researchers to generate new knowledge within the domain. The research tool is integrated in the decision support system. We have identified the main requirements a decision support system must have to be sucessfull in psychophysiological medicine. The requirements are:

- The system must handle multiple online physiological parameters, i.e. measurements. The parameters include a continious stream of time-series measurement from sensors. There exist approximately 10 parameters. We concentrate on the carbondioxide and heart rate parameters as we are classifying the RSA.
- Have the ability to identify features. In this case it is to identify the respiration period and notches in the heart rate.
- Be able to classify heart rate patterns.

A research tool would also benefit from a couple of additional features. These features are:

- The ability to go beyond the imposed limitations of a decision support system. As an example, the ability to laborate with measurements and the classification system.
- Be able to analyse arbitrary measurements at the will of the user.
- Facilitate monotonous but necessary work, such as gather statistics.

An AI system with a feature identification and a pattern classification capability [4, 5, 1] would fit the requirements of an online decision support system in the field of psychophysiology.

1.2 Case-Based Reasoning

Case-Based Reasoning [6, 7] is an AI method. The method is partly based on cognitive psychological research. The CBR approach is psychological plausible as it has been shown in empirical studies that humans use specific past experience to solve new problems [8, 9].

CBR is based on a four step model. The four steps are Retrieve, Reuse, Revise and Retain (see figure 1.1). A CBR systems knowledge is its knowledge base. The knowledge base consists of previously stored

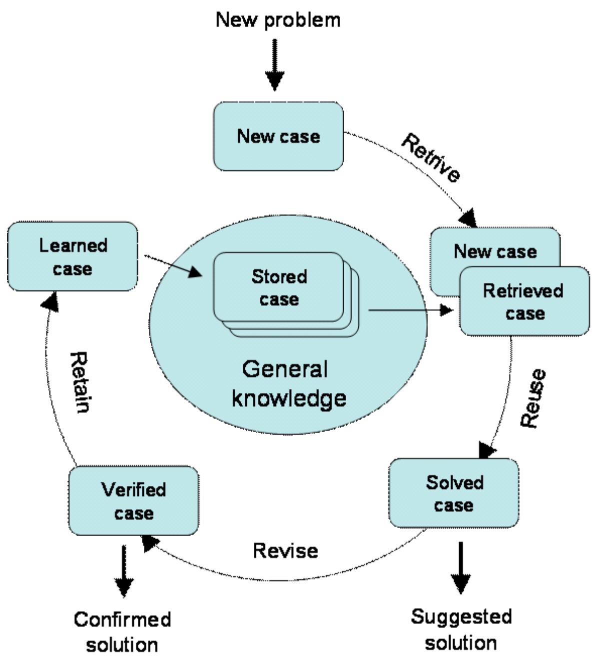

Figure 1.1: The four (Retrieve, Reuse, Retain, Revise) step Case-Based Reasoning model. The figure is adapted from Aamodt-Plaza [8].

experience. Experience is stored as cases, and the cases are stored in a case library. A case represents explicit, specific knowledge, often including a problem, the solution and the result of applying the solution. CBR uses these cases in its reasoning process, hence the name Case-Based Reasoning. Domain knowledge may exist, and that knowledge is stored as general knowledge in form of rules, weights, etc.

The retrieve part tries to find already stored cases that resemble the new case. Matching techniques weight and compare the features in the cases. A search on the matched cases is used to rank and find the cases with the highest similarity to the new case, i.e. the problem. The retrieved cases plus the new case are sent to the reuse part, as seen on the right side in figure 1.1. The reuse part modifies, combines, adapts, etc. (if needed) the retrieved cases in order to find a solution to the problem, i.e. the new case. The CBR system will then suggest the solution for an external evaluation. The evaluator may for instance be a domain expert or another system. The solved case is then sent to a revision part, on the condition the evaluator has not objected to the solution (otherwise, the system returns to the retrieve part for another solution). The solved case is then verified for correctness by the revision. If the solution is valid, it will be presented as a confirmed solution to the problem.

The new case can be added to the case library if the case contain new experience not previosly captured in the case library. This process takes place in the retain step. Cases may also merge, if they for instance cover similar problem areas.

1.2.1 Measuring similarity

Case-Based Reasoning is dependent on a good retrieval. The retrieval is a vital part of CBR because the retrieved cases are supposed to be the best candidates for solving the new problem, i.e. the new case.

A new case may contain any kind of parameters. For instance, ECG measurements, time of day, age of a patient, etc. These parameters are represented by variables. The variables are often arranged in a vector formation. The variables are features describing the vector in an N-dimensional space, hence the name feature vector. The N-dimensional space is variable, the size depends on the number of features the vector is constucted of. Every feature adds a new dimension to the N-dimensional space. A 2-dimensional space is illustrated in figure 1.2.

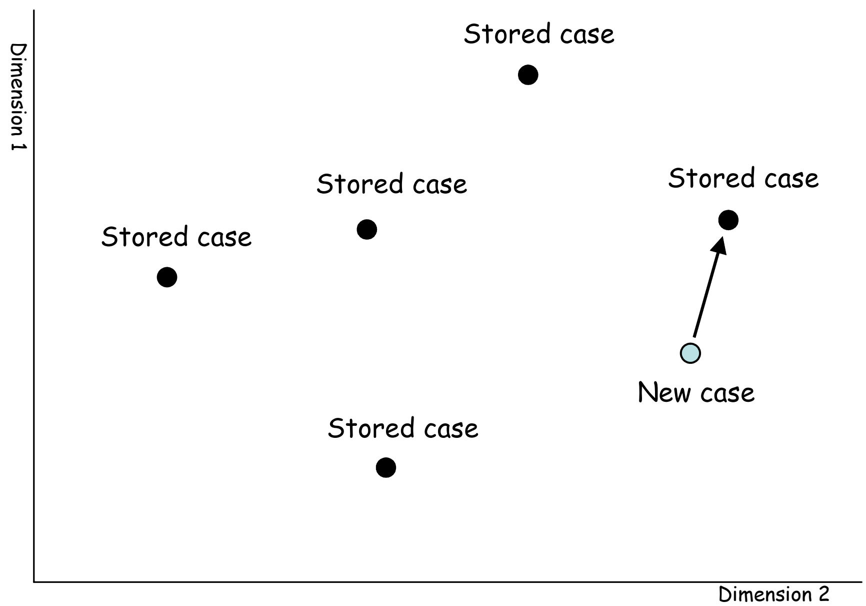

Figure 1.2: The area shows a 2-dimensional space. An arrow marks the euclidian distance between the new case and the closest case in the case library.

A variable is not constrained to a numerical value. The variable may for instance contain subjective knowledge, such as *big*, *brighter* or *bad*. Some sort of pre-determined transformation of the subjective variables is required, if such variables are a part of the feature vector.

The new case is compared with all stored cases in the case library. A similarity measure is calculated between the new case's feature vector and the current case's from the case library. A common approach to the similarity comparison, i.e. the matching is to use a Nearest Neighbour algorithm ($k-NN$). The $k-NN$ tries to find the k closest feature vectors by calculating the euclidian distance between the compared vectors. A feature is often compared with its counterpart in the opposing vector. That is, a feature from the new case is compared with the same feature in the stored case.

$$similarity(sC, nC) = \sum_{i=1}^{n} W_i \times f_i(sFeature_i, nFeature_i) \qquad (1.1)$$

A feature may not be of the same importance as other features in the vector. For instance, the feature *is the heart beating?* is probably more important than *age of the patient*, at least for a heart monitoring system. Domain knowledge may be integrated in order to solve similar situations. The domain knowladge is implemented as weights. Each feature is weighted, W, accoring to its importance. All features in the vectors, $Feature_i$, are weighted and compared. The $f_i()$ function calculates the distance between the features. A similarity value is then calculated for the entire feature vector, $similarity()$. The k closest matching vectors (cases) are then returned as the result of the $k-NN$ algorithm. A typical similarity matching using Nearest Neighbour is illustrated in equation 1.1, where a similarity value is calculated between the stored case, sC, and the new case, nC. n is the size of the feature vector, i.e. number of features used for matching. The matched cases are then ranked in a one dimensional array. The case with the highest similarity to the new case is at position one, the second closest is at the second position, and so on.

$k-NN$ is often a sufficient matching technique in most CBR systems. Another approach to case matching is to use Artificial Neural Networks (ANN) [10, 11]. A large set of training cases is often required when ANNs are used as retrieval techniques in CBR. The integrity of the case base should not be altered too often, as the ANN has to be retrained and validated on every alteration.

1.2.2 Adaptation

A single unaltered case may not be sufficient to propose a viable solution. Adaption of a case, or cases is sometimes required. Cases may be altered or combined in order to find a suitable solution to a problem. The adaptation, i.e. the revision process, identifies prominent differenses between the retrieved case(s) and the new case. Adaptation methods range from null adaptation, i.e. no adaptation at all to complex model-guided repair [12], where the features themselves may be substituted. Adaptation is more often used in systems where the solution is complex and the problem domain is well understood.

1.2.3 Case libary

The case library is a database where all the cases are stored. Cases are retained, i.e. stored, in to the case base in the retain step. A CBR system do not normally store every new case. Retention is often decided by an external supervisor.

The computational performance of a CBR system is connected to the number of cases the case base contain in retrieval methods such as $k-NN$. Methods like inductive retrieval index the entire case base offline for better retrieval performance [12]. A position in the N-dimensional space is calculated for each case and a decision tree is constructed with the indexed cases as a basis. A retrieval is fast since it is only necessary to traverse the decision tree to find matching cases. The drawback with inductive retrieval is that the decision tree has to be rebuilt every time the case library is changed.

Case library maintenance is sometimes necessary. The case base may for instance grow too large to handle efficiently without some sort of case maintenance. A common case library maintenance method is to cluster cases into stereotypical classes, or prototypes. A prototype is a generalisation of the cases it represents.

1.3 Psychophysiology

Stress is a fuzzy and widely used word which triggers many different associations. This section defines stress within the terms of psychophysiological dysfunctions. Psychophysiological dysfunctions range from Burn-Out syndrome to involuntary anxiety attacks. An extreme example of a patient who was suffering from a psychophysiological dysfunction was a torture victim. The patient had nightmares and violent anxiety attacks long after the apparent physiological scars healed. The patient was diagnosed and treated with methodologies within psychophysiology. The methods are almost always non invasive. Treatment personnel do sometimes need, for example arterial blood samples for an accurate diagnosis [13]. Exdermal sensors is otherwise the standard method for gathering physiological data.

Clinicians concentrate on two out of three systems within the Autonomic Nervous System (ANS) when they diagnose psychophysiological dysfunctions. These two systems are the sympathetic and the parasympathetic systems. Clinicians may for instance study the the balance between the sympathetic and the parasympathetic systems as a part of a diagnosis. The third part of the ANS is the enteric nervous system, which is not of immediate interest for a diagnosing clinician.

1.3.1 Heart rate variability

The pulse is not constant. It is well known that the pulse increases during physiological exercises. What is not widely known outside the medical community is that the pulse constantly oscillates, with or without physiological strain. As en example, the pulse oscillates even during sleep. The pulse is often measured in the terms of the interval between two consecutive heart beats. The intervals is mean-valued to reflect the number of beats per minute. Heart rate variability (HRV) [14, 15] is this natural oscillation of the heart beat. HRV is the basis for measuring the balance between the sympathetic and the parasympathetic systems [16]. The oscillation of the heart rate is illustrated in figure 1.3.

A major component in the occurance of HRV is the state of the vagus nerve. The vagus nerve enervates the heart, larynx and the gut. The vagal tone (signal strength) decides the activity level of the sympathetic and the parasympathetic levels, which controls the heart rate. A high vagal tone inhibits the sympathetic system, which leads to a slower heart rate.

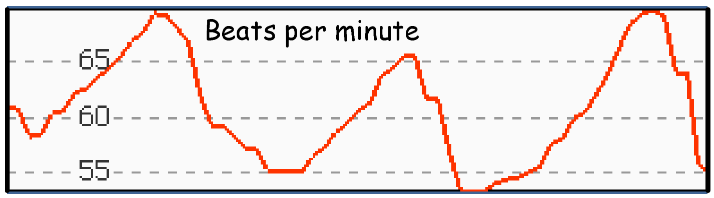

Figure 1.3: An example of heart rate measurements. This example illustratates the heart rate variability, i.e. the oscillating effect of the heart beat. The large oscillation of the heart beat mainly caused by the pulmonary system. The heart rate increases during inhalation and decreases during exhalation.

1.3.2 Respiratory sinus arrhythmia

The definition of Respiratory Sinus Arrhythmia (RSA) is somewhat difficult for an layman to absorb. We will open with a definition of RSA, followed by an explanation.

Respiratory Sinus Arrhythmia (RSA) [17, 18] is defined as centrally modulated cardiac vagal and sympathetic efferent activities associated with respiration [19]. RSA is also often referred to as a non-invasive index of parasympathetic cardiac control [20]. RSA is a form of variability found in the HRV. HRV is as described an effect of the ever changing state of the vagal tone. The pulmonary system has an important influence on the vagal tone, as the heart rate increses remarkably during an inhalation and consequently decreases during an exhalation. This is RSA. The effect the breathing has on the heart rate is clearly visible in figure 1.3, where oscillating heart rate measurements are shown. A common method of observing the RSA, as well as other components in the HRV, is to make a frequency spectrum analysis of the HRV.

1.3.3 Quantification methods

RSA is as mentioned one of the components of the HRV. RSA is quantified in the same manner and with the same methods as the HRV, as RSA is the major component of the HRV. But it is important to recognise the other non respiratory related components of the HRV, as they contamine a quantification. There exist a couple of methods to calculate, or more accurately, estimate RSA from time-series of discretisised heart rate samples. The methods are naturally divided into two groups. The first group contains methods for sample transformations, and the second group contains statical and geometrical methods. Both groups seem to be equally influenced by non respiratory related components [20]. This implies that the best choise of quantification method for RSA is application dependant rather than method dependant.

Transformation methods

A common approach to quantification is to calculate a frequency spectrum [14, 19] through a discrete Fourier Transformation. A Fast Fourier Transformation (FFT) [21, 22] is often sufficient. The FFT produces a vector of complex numbers. The numbers represent the *power* and *angle* of each frequency from $0Hz$ to $Fs/2\ Hz$. Fs represents the sample frequency of the transformed samples.

The FFT is not always sufficient, because the output vector does not say where the frequencies occur in time, only that they occur somewhere inside the sample sequence. Wavelets [23, 21] solve the time-domain issue by using a dynamic function window instead of a static window. The window expands in the time-domain during lower frequencies while contracting the window on the higher frequencies.

Statical and geometrical methods

A quantification can be achieved without transforming the samples. The quantification is produced directly in the time domain, by either statistical or geometrical methods [14].

Statistical methods are more often used on measurements recorded from longer time periods, normally 24h periods. The sample sequence is created by measuring either the Normal-to-Normal (NN) heartbeat interval or the differences in the NN interval. There exist a couple of statistical methods, but they are basically the same since they are all based on standard derivations of the NN.

Geometric pattern methods use sample density distributions in their quantifications. The geometrical calculations are based on NN histograms; and are often in the form of geometrical shapes. 24h recording periods are recommended for the geometrical methods.

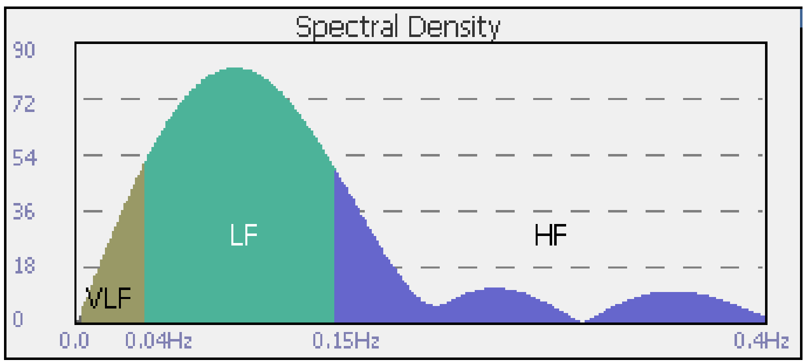

Figure 1.4: The figure shows the VLF, LF and HF frequency bands in the HRV. The ULF is the tiny black dot farthest to the left (between $0Hz$ and VLF). These frequencies are quantified with a FFT based on the measurements in figure 1.3.

Interpretation

Clinicians are only interested in specific frequency bands when they study the quantified HRV. These frequency bands range from $0Hz$ up to $0.4Hz$. Four bands exist within this range, the Ultra Low Frequencies (ULF), the Very Low Frequencies (VLF), the Low Frequencies (LF) and the High Frequencies (HF). The frequency bands is illustrated in figure 1.4.

The ULF frequency band is in the range of $0 - 0.003Hz$,
followed by VLF at $0.003 - 0.04Hz$, LF at $0.04 - 0.15Hz$,
and finally the HF band at $0.15 - 0.4Hz$.

The HF band is only affected by the parasympathetic system. The lower frequencies is affected by both the sympathetic and the parasympathetic systems. Other biological systems is also detectable at the lower frequencies. For instance, changes in the body temperature during the diurnal rythm, is detectable in the VLF band.

Parasympathetic activities related to respiratory activity can be isolated during so called paced breathing, or pacing. Pacing is a method for breath control. The patient is in this case coached to breathe at exactly 6 breaths per minute. A peak appear at 0.1Hz when the patient is breathing steadily at 6 breaths a minute. The peak is a good indicator of the vagal tone.

1.3.4 Treatment methods

Biofeedback training [24, 15, 17, 18] is often the prefered method for treating psychophysiological dysfunctions, such as post traumatic stress and Burn-Out syndrome. The biofeedback method is often focused on breathing techniques. The patient gradually learns how to control their respiration. The controlled breath improves the balance between the parasympathetic and the sympathetic systems. Other methods of treatments include relaxation techniques combined with regular exercises [25].

Chapter 2

Related work

The research in medical CBR is concentrated to Europe and US, as with CBR in general. The medical domain of CBR is generally focused on producing systems for specific tasks, such as diagnosing a specific symptom. These systems are rarely in every-day-use and they are seldom commersially exploited [1]. Most of them are still purely academic.

Some CBR issues are especially interesting for the medical domain. Fault tolerance is one of the more important issues. Bichindaritz *et al* [26, 27] has developed a medical CBR system based on a safety-insurance plan to insure that no local faults are spread beyond its scope. Limitations of CBR in medical systems are addressed by Schmidt and Gierl [28, 3] where they approach the difficult task of case adaptation. Atzmulluer *et al* [29, 30] approach the issue of handling multiple faults, or diagnoses, in a medical system. Their solution is to decompose the problem part of a case into several smaller ones and find solutions to them instead. When the system comes up with solutions to the smaller problems they combine all solutions to solve the original problem.

A description of some of the the closest related medical CBR systems is given in the next section.

2.1 System descriptions

The headlines of the following system descriptions indicate the name of the specific system. The main author is in the headline if the system lack a formal name.

CARE-PARTNER

CARE-PARTNER [27, 26] is a decision support system for the long term follow-up of stem cell transplanted patients at Fred Hutchinson Cancer Research Center (FHCRC) in Seattle. The CARE-PARTNER system gives medical and decision support to the home care providers that follow up the transplant patients, using the Internet to connect the home care providers with the FHCRC transplant specialists. The system uses a multi modal reasoning framework, combining Case Based Reasoning and Rule Based Reasoning. A safety insurance plan at three levels (a procedural, a software engineering and a knowledge level) is adopted to ensure fault tolerance. One main characteristic of the system is that it uses a rich knowledge base of prototypical cases and practice guidelines to interpret medical cases and guide the case based reasoning.

ICONS

ICONS [31, 32] is a CBR system which analyses renal time-series. ICONS forecast kidney functions through an extended CBR cycle. The extension includes a state abstraction step and a temporal abstraction step. These steps are located in front of the regular steps in the CBR cycle. Kidney states are abstracted from renal measurements in the state abstraction step. Creatine clearance is always included as a state, other conditional states are included based on Tverky's measure of dissimilarity of concepts [31]. Only states that are very probable, i.e. have a low dissimilarty, are included. The temporal abstraction step makes prognostic models, i.e. trends, from the abstracted states of the kidney. The revision step has been removed.

Auguste project

The **Auguste** project [33], is an effort to provide decision support for planning the ongoing care of Alzheimer's Disease (AD) patients. The first reported system prototype supports the decision to prescribe neuroleptic drugs for behavioural problems. The prototype is a hybrid system where a CBR part decides if a neuroleptic drug is to be given, and a Rule-Based Reasoning (RBR) part decides which neuroleptic to use. The system uses approximately 100 features, manually extracted from medical charts, in each case for determining the right neuroleptic drug. The patient is initially screened for behavioral problems before a Nearest Neighbour match makes a suggestion on whether or not to give neuroleptics to the patient. If the CBR module finds it appropriate to give neuroleptics and no contradictions are found, e.g., allergies to certain drugs etc., the RBR module determines which neuroleptic (of five available) to use. This prescriptive task, although termed "planning" in the vernacular, may be best characterized as one of design.

FM-Ultranet

FM-Ultranet [34, 35] is a medical CBR project implemented with CBR-Works. FM-Ultranet detects malformations and abnormalities of foetus through ultrasonographical examinations. The detection, or diagnosis, uses attributes derived from scans of the mother's uterus, and identifies abnormal organs and extremities. Cases are arranged in a hierarchical and object oriented structure. The hierarchy is organized in 39 concepts, and every concept has one or more attributes. The attributes consists of anatomical features, medical history and general domain knowledge. Similarity between attributes in the concepts (objects) are mathematically calculated or compared through a look up table, depending on the attribute type. A report of the system's findings are generated when the detection (CBR) process is completed.

Montani *et al.*

Montani *et al.* has developed two separate CBR systems for the medical domain. The first system attempts to integrate different methodologies into a Multi-Modal Reasoning (MMR) system [36, 37]. The system is used in therapy support for diabetic patients. The authors argue that most systems trying to utilize more than one methodology do so only in an exclusive fashion, with methodologies functioning merely as extensions to one another. Montani argues that a MMR system needs much closer integration of technologies to get the full benefits of a multi-modal solution. Integration allows tackling well known problems of single methodologies, i.e. the qualification problem in RBR and the too-small-a-library problem in CBR. The proposed system tries to use a fuller integration and utilize CBR, Rule-Based Reasoning, and Model-Based Reasoning (MBR).

The second system by Montani *et al.* is focusing on CBR in hemodialysis treatments for end stage renal disease [38]. This system is applied to the efficiency assessments of hemodialysis sessions. Each new dialysis session, i.e. assessment, is represented as a case in the system. Patterns of failures over time, from the patients past history, and cross references with other patients, can be found with this solution. Features are both statically and dynamically collected. The static features are patient information of a general nature (age etc.), and the dynamical features originates from online measurements in the form of continuous time series. The online features used for assessment is mainly derived from the extracorporeal circuit during a dialysis session, like measuring the arterial pressure.

Perner *et al.*

Perner *et al* has developed a couple of systems. One of the most interesting is the Alzheimer imaging project. They present a system that uses CBR to optimize image segmentation at the low level unit according to changing image acquisition conditions and image quality [39]. The system has been used to detect degenerative brain disease, in particular Alzheimer disease in CT images of a patient. The cases are comprised of images and image features as well as non-image information about the image acquisition and the patient. The solution of a case is the parameters of the image segmentation unit. The control of the parameter of the image segmentation unit is done by the CBR mechanism. This ensure high image quality of the output image. Similarity is calculated over the image information according to a special image similarity measure and over the non-image information. Finally, both similarity measures are combined to an overall similarity measure. The system was used at the Radiology Department at the University of Halle.

Chapter 3

Contributions

Three papers are included in this thesis. The included papers are ordered to guide the reader in to my specific area of research, instead of ordering them in a plain chronological order.

The first paper, paper A, was presented at FLAIRS '04. FLAIRS is the annual Florida AI Researchers International Conference, and was held in South Beach, Miami. The title of the paper is *Advancements and Trends in Medical Case-Based Reasoning: An Overview of Systems and System Development.* The second paper was presented at ICCBR '03 in Trondheim. ICCBR is the International Conference on Case-Based Reasoning. ICCBR is arranged every other year. The paper was titled *Complex Measurement Classification in Medical Applications Using A Case-Based Approach.* The third and final paper was written for ECCBR '04 in Madrid. ECCBR, i.e. the European Conference on Case-Based Reasoning, is also arranged every other year. The title of this paper is called *A Case-Based Classification of Respiratory Sinus Arrhythmia.*

3.1 Paper A

Paper A contains an analysis of CBR in the medical domain. A number of systems are analysed, explained and summarised. System architectures are also presented, where the information has been available.

The paper begins with a discussion of the advantages and the disadvantages of CBR in the medical domain. The included systems are categorised accoring to their system properties, i.e. purpose, or typ of task to perform. Construction oriented properties are also investigated. Construction oriented trends consist of such things as system architechtures, case library sizes, autonomicity etc. Identifiable trends and other interesting observations conclude the paper.

3.2 Paper B

Paper B is describing a CBR system designed for classification of psychophysiological dysfunctions. The design is divided into three distinct parts. Each part is responsible for a specific task, and the result of a part is depending on the input from the previous part.

The first part is the pre-processing. The pre-processing handles the signals, i.e. the physiological measurements streamed from various hardware sensors. The signals are processed and cleaned from noise and distortions. The aim is to produce a signal that reflects the original physiological event. The second part receives the cleaned signals from the pre-processing. This part is responsible for feature identification. Features are extracted from the signals by feature extraction templates. A template contains information on how to identify a specific feature. All extracted features are collected into a feature vector. Other complex features like trends and cross referenced features are also added to the vector. Classification based on the features vectors is conducted in the third part. The feature vectors are matched against stored cases. The stored cases contain classifications of psychophysiological dysfunctions. Cases are ranked in order of importance once a possible dysfunction is identified. The ranking process uses patient specific information.

3.3 Paper C

The final paper describes an implementation of a classification system for Respiratory Sinus Arrhythmia. The system architechture is based on the design in paper B. The system analyses both time-series of heart rate and carbondioxide measurements. The carbondioxide analysis identifies respiration cycles, i.e. individual breaths. The breaths are piped to the heart rate analysis. Heart rate measurements that are within the same time period as the respiration cycle are extracted. The extracted heart rate measurements are processed to fit a case. Some of the processes are for instance a transformation of the measurements to the frequency domain. The carbondioxide measurements, the heart rate measurements and the processed heart rate measurements are clustered to a feature vector. Rule based reasoning limits the number of cases the CBR cycle has to match by counting notches, i.e. dips in the heart rate measurements. CBR matches the remaining cases and suggests similar classes of Respiratory Sinus Arrhythmia to the user.

Chapter 4

Conclusions

4.1 Summary

The main contributions of this thesis can be summarised into a couple of statements, they are:

- A novel approach for analysis of physiological time-series.
- A Multi-Modal reasoning system for classification of Respiratory Sinus Arrhythmia.
- A decision support system for treatment personnel in psychophysiological medicine.

A screenshot of the system is pictured in figure 4.1. The system is christened HR3modul, based on its capability to analyse heart and respiration time-series (Heart Rate, Respiratory Rate). The screenshot of the HR3modul system is taken during a learning mode, i.e. when an expert is pushing new cases to the library.

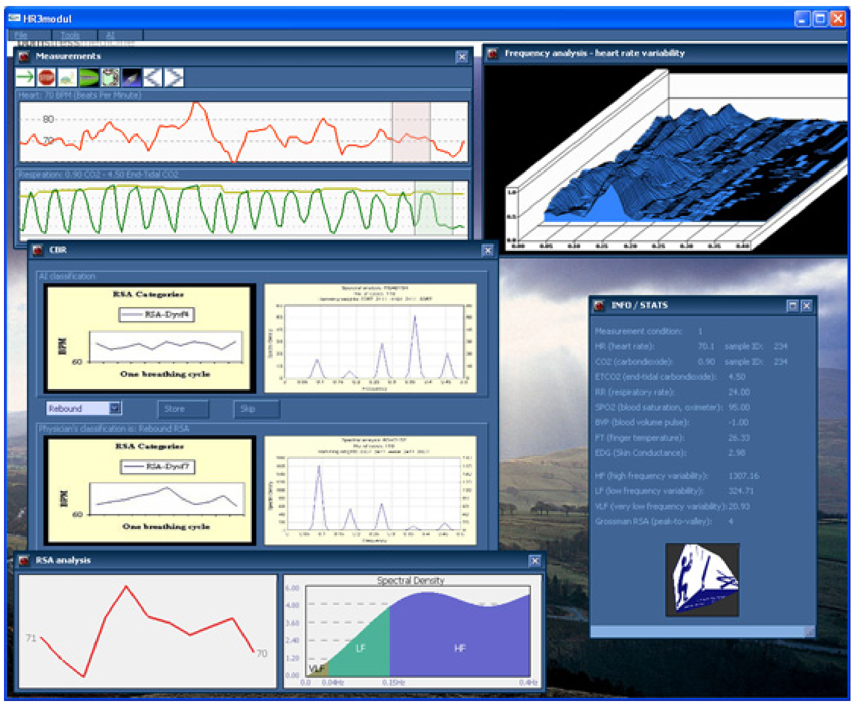

Figure 4.1: HR3modul.

4.2 Future work

An interesting extension to this research would be to incorporate a full breath classification within the respiration analysis. The current respiration analysis identifies the respiration period, i.e. the beginning and the end of individual breaths. A full analysis could identify dysfunctions within the respiration by classify the capnograpy measurements.

Another interesting project is to make an automatic weighting system. The weights for the features is manually adjusted in the current system. The weights could be adjusted offline, that is, when the system is idle. Such a weighting system could use genetic algorithms to try new combinations of weights.

A faster retrieval technique may be necessary as the case library grows. Each case contains, among other things, a FT vector of 1025 points (transformed by a 2048 point FFT). This vector may slow the matching process as all cases in the case library are currently tested against the new case. This matching is performed on every breath, and a normal breath is usually in the range of 4-20 times per minute.

Bibliography

[1] Markus Nilsson and Mikael Sollenborn. Advancements and trends in medical case-based reasoning: An overview of systems and system development. In *Proceedings of the 17th International FLAIRS Conference*. Florida Artificial Intelligence Research Society (FLAIRS'04), May 2004.

[2] Lothar Gierl, Mathias Bull, and Rainer Schmidt. Cbr in medicine. In *Case-Based Reasoning Technology: From Foundations to Applications*, pages 273–297. Lecture Notes in Artificial Intelligence, May 1998.

[3] Rainer Schmidt and Lothar Gierl. Case-based reasoning for medical knowledge-based systems. XVI International Congress of the European Federation for Medical Informatics, 2000.

[4] Markus Nilsson and Peter Funk. Classification of psychophysiological measurements in medical systems. In *Proceedings of the SAIS-SSLS joint Workshop*. Swedish Artificial Intelligence Society and Swedish Society for Learning Systems (SAIS-SSLS'03), April 2003.

[5] Markus Nilsson, Peter Funk, and Mikael Sollenborn. Complex measurement classification in medical applications using a case-based approach. In *Workshop proceedings: Case-Based Reasoning in the Health Sciences*, pages 63–72. International Conference on Case-Based Reasoning (ICCBR'03), June 2003.

[6] Janet Kolodner. *Case-based Reasoning*. Morgan Kaufmann Publishers Inc, 340 Pine St, 6th floor, San Fransisco, CA 94104, USA.

[7] David B. Leake. *Case-based Reasoning: Experiences, Lessons, and Future Directions*. AAAI Press/MIT Press, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, 1996.

[8] Agnar Aamodt and Enric Plaza. Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications*, 7(1):39–59, 1994.

[9] Padraig Cunningham and Barry Smyth. A comparison of model-based and incremental case-based approaches to electronic fault diagnosis. In *Proceedings of the Case-Based Reasoning Workshop*. American Association of Artificial Intelligence, 1994.

[10] M. Frize, H.C.E. Trigg, F.G. Solven, M. Stevenson, and B.G. Nickerson. Decision-support systems designed for critical care. In *Proceedings of the American Medical Informatics Association Conference*. AMIA'97, October 1997.

[11] Laurene Fausett. *Fundamentals of Neural Networks*. Prentice Hall, Upper Saddle River, New Jersey 07458, USA, 1994.

[12] Ian Watson. *Applying Case-Based Reasoning: Techniques for Enterprise Systems*. Morgan Kaufmann Publishers Inc, 340 Pine St, 6th floor, San Fransisco, CA 94104, USA, 1997.

[13] B.H.C. von Schéele and I.A.M. von Schéele. The measurement of respiratory and metabolic parameters before and after incremental exercise on bicycle of patients and control participants: Supporting the effort syndrome hypothesis? *Applied Psychophysiology and Biofeedback*, 24(3):167–177, 1999.

[14] Marek Malik and John Camm et al. Heart rate variability - standards of measurement, physiological interpretation, and clinical use. *European Heart Journal*, 17:354–381, March 1996.

[15] Marie J. Cowan, Helen Kogan, Robert Burr, Sue Hendershot, and Lynne Buchanan. Power spectral analysis of heart rate variability after biofeedback training. *Journal of Electrocardiology*, 23:85–94, 1990.

[16] Gary G. Berntson, J.T. Bigger, D.L. Eckberg, P. Grossman, P.G. Kaufmann, M. Malik, H.N. Nagaraja, S.W. Porges, J.P. Saul, P.H.

Stone, and M.W. van der Molen. Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology*, 34(6):623–648, 1997.

[17] Paul M. Lehrer, Evgeny Vaschillo, and Bronya Vaschillo. Resonant frequency biofeedback training to increase cardiac variability: Rationale and manual for training. *Applied Psychophysiology and Biofeedback*, 25(3):177–191, 2000.

[18] Paul M. Lehrer and Alexander Smetankin et al. Respiratory sinus arrhytmia biofeedback therapy for asthma: A report of 20 unmedicated pediatric cases using the smetanikn method. *Applied Psychophysiology and Biofeedback*, 25(3):193–200, 2000.

[19] Tuan Pham Dinh, Helene Perrault, Pascale Calabrese, Andre Eberhard, and Gila Benchetrit. New statical method for detection and quantification of respiratory sinus arrhytmia. *IEEE Transactions on Biomedical Engineering*, 46(9):1161–1165, September 1999.

[20] P. Grossman and J. Van Beek et al. A comparison of three quantification methods for estimation of respiratory sinus arrhytmia. *Psychophysiology*, 27(6):702–714, November 1990.

[21] Ralph D. Hippenstiel. *Detection Theory - Applications and Digital Signal Processing*. CRC Press, 2000 N.W. Corporate Blvd., Boca Raton, Florida 33431, USA, 2002.

[22] Samuel D. Stern. *Digital Signal Processing*. CRC Press, 2000 N.W. Corporate Blvd., Boca Raton, Florida 33431, USA, 2003.

[23] Karsten Sternickel. Automatic pattern recognition in ecg time series. *Computer Methods and Programs in Biomedicine*, 68:109–115, 2002.

[24] Carl F. Weems. The evaluation of heart rate biofeedback using a multi-element design. *Journal of Behaviour Theraphy and Experimental Psyciatry*, 29:157–162, 1998.

[25] Jan van Dixhoorn. Cardiorespiratory effects of breathing and relaxation instruction in myocardial infarction patients. *Biological Psychology*, 49:123–135, 1998.

[26] Isabelle Bichindaritz, Emin Kansu, and Keith M. Sullivan. Case-based reasoning in care-partner: Gathering evidence for evidence-based medical practice. In *Advances in CBR: 4th European Workshop*, pages 334–345. ECCBR'98, September 1998.

[27] Isabelle Bichindaritz, Carol Moinpour, Emin Kansu, Gary Donaldson, Nigel Bush, and Keith M. Sullivan. Case based reasoning for medical decision-support in a safety critical environment. *Artificial Intelligence in Medicine*, 2780:314–323, November 2003.

[28] Rainer Schmidt, Olga Vorobieva, and Lothar Gierl. Case-based adaptation problems in medicine. 2nd German Workshop on Experience Management, April 2003.

[29] Joachim Baumeister, Martin Atzmueller, and Frank Puppe. Advances in case-based reasoning. In *Proceedings of the 6th European Conference on Case Based Reasoning*, pages 28–42. European Conference on Case-Based Reasoning (ECCBR'02), 2002.

[30] M. Atzmueller, W. Shi, J. Baumeister, F. Puppe, and J.A. Barnden. Case-based approaches for diagnosing multiple disorders. In *Proceedings of the 17th International FLAIRS Conference*. Florida Artificial Intelligence Research Society (FLAIRS'04), May 2004.

[31] Rainer Schmidt and Lothar Gierl. Temporal abstractions and case-based reasoning for medical course data: Two prognostic applications. In *Machine Learning and Data Mining in Pattern Recognition*, volume 2123 of *Lecture Notes in Computer Science*, pages 23–34, 2001.

[32] Rainer Schmidt, Dagmar Steffen, and Lothat Gierl. Evaluation of a case-based antibiotics therapy adviser. In *Artificial Intelligence in Medicine*, pages 462–466. AIME'01, July 2001.

[33] Cindy Marling and Peter Whitehouse. Case-based reasoning in the care of alzheimer's disease patients. In *Case-Based Research and Development*, pages 702–715. International Conference on Case-Based Reasoning (ICCBR'01), July/August 2001.

[34] Ziad El Balaa and Ralph Traphoner. Case-based decision support and experience management for ultrasonography. In *German Workshop on Experience Management.* (GWEM'03), April 2003.

[35] Ziad El Balaa, Anne Strauss, Philippe Uziel, Kerstion Maximini, and Ralph Traphoner. Fm-ultranet: a decision support system using case-based reasoning, applied to ultrasonography. In *Workshop on CBR in the Health Sciences*, pages 37–44. ICCBR'03, June 2003.

[36] Stefania Montani, Paolo Magni, Abdul V. Roudsari, Ewart R. Carson, and Riccardo Bellazzi. Integrating different methodologies for insulin therapy support in type 1 diabetic patients. In *Artificial Intelligence in Medicine*, pages 121–130. AIME'01, July 2001.

[37] Stefania Montani, Paolo Magni, Riccardo Bellazzi, Cristiana Larizza, Abdul V. Roudsari, and Ewart R. Carson. Integrating model-based decision support in a multi-modal reasoning system for managing type 1 diabetic patients. *Artificial Intelligence in Medicine*, 29:131–151, 2003.

[38] Stefani Montani, Luigi Portinale, Giorgio Leonardi, and Riccardo Bellazi. Applying case-based retrieval to hemodialysis treatment. In *Workshop on CBR in the Health Sciences*, pages 53–62. International Conference on Case-Based Reasoning (ICCBR'03), June 2003.

[39] Petra Perner. An architeture for a cbr image segmentation system. *Journal on Engineering Application in Artificial Intelligence*, 12(6):749–759, 1999.

II

Included Papers

Chapter 5

Paper A: Advancements and Trends in Medical Case-Based Reasoning: An Overview of Systems and System Development

Markus Nilsson, Mikael Sollenborn
In Proceedings of the 17th International FLAIRS Conference (FLAIRS'04), Miami, May 2004

Abstract

Case-Based Reasoning (CBR) is a recognised and well established method for building medical systems. In this paper, we identify strengths and weaknesses of CBR in medicine. System properties, divided into construction-oriented and purpose-oriented, are used as the basis for a survey of recent publications and research projects. The survey is used to find current trends in present medical CBR research.

5.1 Introduction

Ever since Shortliffe's seminal work on diagnosis of infection diseases [24], Artificial Intelligence has been applied in numerous applications in the health science domain. In the late 1980's, followed by ground-laying work done by Koton [13], and Bareiss [3], Case-Based Reasoning (CBR) appeared as an interesting alternative for building medical AI applications, and has since been further established in the field. Certainly, one of the intuitively attractive features of CBR in medicine is that the concepts of *patient* and *disease* lends itself naturally to a case representation. Although several advantages of using CBR in medicine has been identified, the medical field certainly is not without its problems, some of them specifically affecting CBR systems.

Gierl and Schmidt [10] identify the following key advantages of medical CBR;

- Cognitive Adequateness. CBR resembles the way physicians are reasoning about patients and the way they use their case expertise.
- Explicit Experience. A CBR system is naturally suited for adjusting itself to the specific requirements of a certain clinic or a surgeon.
- Duality of Objective and Subjective Knowledge. Instead of using the subjective knowledge of one or more experts to build systems (as is done for e.g. rule-based expert systems) CBR systems are built upon existing cases (which may or may not be fully understood).
- Automatic Acquisition of Subjective Knowledge. CBR systems exhibit an incremental knowledge acquisition, and knowledge can be abstracted by generalizing cases.
- System Integration. Patient records are already being collected by hospitals and practitioners and stored on machine readable mediums, which simplifies integration with CBR systems which can utilize them (after varying degrees of modification).

However, a number of disadvantages and problems can also be identified;

- Adaptation. Because of the often extremely large number of features involved in a medical case, adaptation of cases becomes problematic. Generalization and efficient feature identification methods helps to partly remedy this issue, but generally the problem persists. [20]

- Unreliability. Although the reliability of a CBR system increases with the proportion of coverage of the problem domain, reliability cannot be guaranteed. Adding new cases will not necessarily make a system converge towards greater reliability, as cases add only local improvement. Indeed, Bichindaritz argues that the strictly local properties of cases makes convergence an inappropriate notion for CBR systems.[4]

- Concentration on reference. CBR systems are concentrated on reference as opposed to underlying diagnostic factors. Thus, systems cannot function as sources of previous experience unless a suitable case exists in the case-base.

In this paper, we take a look at a number of the most influential medical CBR research projects in late years, with the aim of identifying trends in the development of such systems. Basing our work on the 1998 survey by Gierl and Schmidt [10], we focus primarily on systems created or reported about after 1998. In particular, we are interested in investigating if, and to what degree, the focus has changed on what type of medical CBR systems are constructed, and how they are constructed.

The method of identifying current trends involves examining systems from recent years by focusing on a set of distinctive system properties. We divide system properties into *purpose-oriented* and *construction-oriented*, where the first are characterized by the general type of action the system is supposed to perform (classification, planning, diagnosis, and tutoring), and the second indicate different types of constructions, such as systems supporting adaptation, hybrid systems, varying degrees of autonomicity etc. Additionally, we attempt to find trends of more general importance, looking at medical CBR systems from a broader perspective.

The rest of the paper is organized as follows. Next section contains a description of the different comparison properties used to differentiate a system from another. The section *Recent medical CBR systems* describes a selected number of influential works in the medical CBR domain. In section *Trends in medical CBR*, we present a system property matrix and identify construction-oriented and overall trends.

5.2 System properties

5.2.1 Purpose-oriented properties

With purpose-oriented properties, we refer to the separation of overall system purpose into planning, classification, tutoring, and diagnostic.

Diagnostic systems The majority of medical CBR systems belong in the diagnostic systems category. Diagnostic systems attempt to provide the user of the system with various degrees of assistance in the diagnosing process of a medical condition, possibly up to the point of a completely autonomous diagnose.

Classification systems Classification systems attempt to identify the group or group affiliations of real-world cases. One typical example is image classification systems that do not attempt a complete diagnosis.

Tutoring systems A medical tutoring system based on CBR is typically built closely around the concept of learning by examples, providing students with access to real patient cases.

Planning systems Planning systems are characterized by their intention to help in solving a process involving a number of steps. Therapy support is an often seen example of planning in medical systems.

5.2.2 Construction-oriented properties

Looking at medical CBR systems, we are interested not only in which systems have been recently constructed, but also how they were constructed and the motivation behind their construction. Once again, this is done to ease the identification of current trends in the development of medical CBR. However, in some cases it is not possible to derive the state of all these properties from the papers describing the projects in question.

Hybrid systems A hybrid medical CBR system denotes a multi-faceted solution to a problem space, using CBR as one of a number of AI technologies forming a complete system. Many such systems use CBR as the main organizer of data, and data-intense techniques such as neural networks to handle lower-level case identifications. Others match CBR with the Rule-Based Reasoning used in traditional expert systems to gain the advantages of both Rule-Based and Case-Based Reasoning.

Adaptive systems The problem of doing successful adaptation in the medical domain, because of the often enormous amount of features in a case, has been documented by Schmidt and Gierl [20]. In the system summary in section *Recent medical CBR systems*, we investigate if and to what degree medical CBR systems from recent years has started to utilize adaptation methods.

Case library size The size of the case library does not only involve the actual number of cases in the case library, but also the degree of case generalisation into prototypes, i.e., the degree to which the system tries to merge existing cases into more general ones.

Autonomicity The degree of autonomicity is arguably of the most importance for diagnostic systems, where it denotes the level of interaction needed with a physician or corresponding medical expert before and after the diagnosis is complete. A purely autonomous system would produce diagnoses that would be accepted and used without having a human expert look at them, which is rarely the case in current systems. The degree of autonomicity implies the need for human intervention in the reasoning cycle and for evaluating its results.

Constraints System constraints concerns reliability and safety-criticality. Safety-criticality denotes the need to always provide correct answers, e.g., whether incorrect system behaviour could potentially create dangerous or even life threatening situations. A system is reliable if it is always operational when needed.

5.3 Recent medical CBR systems

As was mentioned in the introduction, the focus of the survey is on systems created or reported about during the last five years. An overview of medical CBR systems before 1998 was done by Gierl et al. in [10]. From this overview, we adopted the division of systems into diagnostic, classification, tutoring, and planning systems.

5.3.1 Diagnostic systems

FM-Ultranet [1, 2] is a medical CBR project implemented with CBR-Works. FM-Ultranet detects malformations and abnormalities of foetus through ultrasonographical examinations. The detection, or diagnosis, uses attributes derived from scans of the mother's uterus, and identifies abnormal organs and extremities. Cases are arranged in a hierarchical and object oriented structure. The hierarchy is organized in 39 concepts, and every concept has one or more attributes. The attributes consists of anatomical features, medical history and general domain knowledge. Similarity between attributes in the concepts (objects) are mathematically calculated or compared through a look up table, depending on the attribute type. A report of the system's findings are generated when the detection (CBR) process is completed.

Perner [18] proposes a system that uses CBR to optimize image segmentation at the low level unit according to changing image acquisition conditions and image quality. The system has been used to detect degenerative brain disease in particular Alzheimer disease in CT images of a patient. The cases are comprised of images and image features as well as non-image information about the image acquisition and the patient. The solution of a case is the parameters of the image segmentation unit. The control of the parameter of the image segmentation unit is done by the CBR mechanism. This ensure high image quality of the output image. Similarity is calculated over the image information according to a special image similarity measure and over the non-image information. Finally, both similarity measures are combined to an overall similarity measure. The system was used at the Radiology Department at the University of Halle.

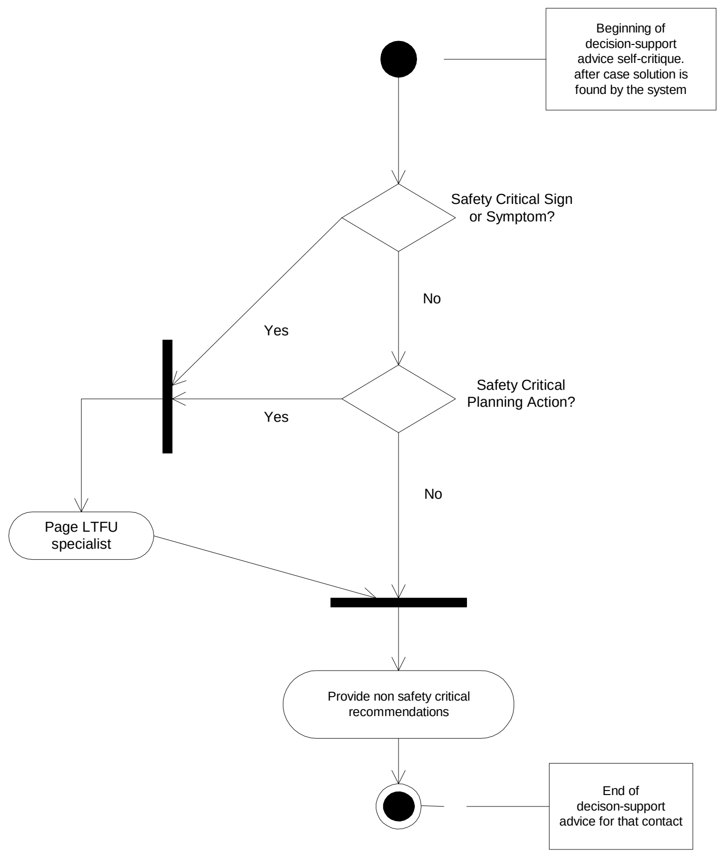

Figure 5.1: caption.

Jaulent *et al.* [12] is diagnosing histopathology in the breast cancer domain. Their system uses cases that are derived from written medical reports. A case has an internal tree structure, and represents a collection of macroscopic area. Every macroscopic area is a collection of histological areas, and each histological area contains a cytological description of subjective features, like a big cell size. The features are also weighted for importance. Cases are compared for structural (structure of the histological tree), surface (semantic resemblance of microscopic areas) and feature similarity. A translation transposes the subjective features into numerical values.

CARE-PARTNER [4, 5] is a decision support system for the long term follow-up of stem cell transplanted patients at Fred Hutchinson Cancer Research Center (FHCRC) in Seattle. The CARE-PARTNER system gives medical and decision support to the home care providers that follow up the transplant patients, using the Internet to connect the home care providers with the FHCRC transplant specialists. The system uses a multi modal reasoning framework, combining Case Based Reasoning and Rule Based Reasoning. A safety insurance plan at three levels (a procedural, a software engineering and a knowledge level) is adopted to ensure fault tolerance. One main characteristic of the system is that it uses a rich knowledge base of prototypical cases and practice guidelines to interpret medical cases and guide the case based reasoning. CARE-PARTNER's safety insurance plan is illustrated in figure 5.1.

Schmidt *et al.* deal specifically with prototypes in [22], where a prototype denotes a generalisation occurring as a result of grouping/clustering single cases into more general ones. The claim is made that generating prototypes is also an adequate technique to learn intrinsic case knowledge, especially if the domain theory is weak. Storing new cases may improve the ability to find solutions for similar cases, but to understand the knowledge included within, generalisation is needed. Schmidt and Gierl have developed several systems focusing on generalising into prototypes, as described in their 1998 medical CBR survey [10], such as **ICONS** [23] for antibiotic therapy advice, **GS.52** for diagnosis of dysmorphic syndromes, **COSYL** for liver patient treatment strategies, and **TeCoMED** for forecasting epidemics of infection diseases. These are all further described in [10] and [22]. In [22], Schmidt argues that the reason for using prototypes varies with the type of application and task. In areas where the domain theory is weak, prototypes help to guide the retrieval. In other systems, prototypes may correspond directly with the physicians view and be absolutely necessary for the project. Prototypes also help to speed up retrieval by decreasing the number of cases. The general drawback of prototypes is however loss of information when generalizing.

MED2000. Goodridge *et al.* [11] presents a theoretical diagnostic model for dealing with medical CBR domain problems. The theoretical model, referred to as the Case-Based Neural Network Model, incorporates CBR within a neural network, and the concept of representing knowledge using frames. The CBR-specific problems addressed are all of those mentioned in the introduction, but unfortunately the description lacks a thorough investigation of how, or even if, most of the problems can be remedied with the proposed method. The heart of the model is the separation of case information into two layers, keeping all information identifiers and case features in layer one, and the actual solutions in layer two. Doing this, the system can eliminate the problem of case representation as the information entities in layer one are independent of the solutions. The paper also introduces the concept of pure cases as a way of dealing with the adaptation problem, but it is unclear whether it will actually present an improvement. MED2000 is a hybrid system, has low autonomicity due to experts accepting/declining every hypothesis, and contains a fairly small number of cases, approximately 40 cases. The neural network architecture provides a level of "natural" prototype usage.

5.3.2 Classification systems

Montani *et al.* has focused on CBR in hemodialysis treatments for end stage renal disease [15]. Their system is applied to the efficiency assessments of hemodialysis sessions. Each new dialysis session, i.e. assessment, is represented as a case in the system. Patterns of failures over time, from the patients past history, and cross references with other patients, can be found with this solution. Features are both statically and dynamically collected. The static features are patient information of a general nature (age etc.), and the dynamical features originates from online measurements in the form of continuous time series. The online features used for assessment is mainly derived from the extracorporeal circuit during a dialysis session, like measuring the arterial pressure.

Costello and Wilson [7] is focusing on the classification of mammalian DNA sequences, and are using a case library of nucleotide (A,T,G,C) segments. The stored segments are already classified as exons (carrying information on how to create proteins) and introns (junk segments that do not carry any information). The system is identifying exons in a seemingly random mix of exons and introns in strands of DNA. An edit

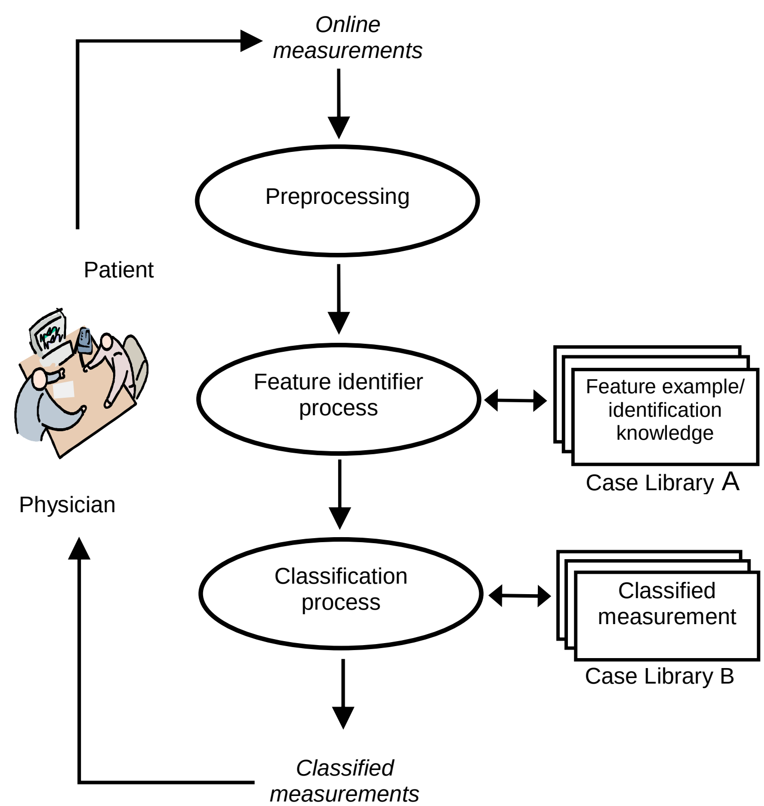

Figure 5.2: caption.

distance calculation of, insertion, substitution and deletion of individual nucleotides in the tested exons is used to evaluate the similarity between the test strand and the store exon cases. Matched exons is then grouped through activation levels (number of similarities) to find new segments of exons in the test strand.

Nilsson *et al.* [17] address the domain of psychophysiological dysfunctions, a form of stress. The system is classifying physiological measurements from sensors. The system is divided into smaller distinct parts. Measurements, like signals from an ECG, are filtered and improved. A case library of models of distortions etc. is applied to the filters. Features are extracted from the filtered signals (measurements). An additional set of features are extracted from the first set, for trend analysis etc. The features from the first and second set, and patient specific data, are used as a case. The cases are classified with a k-nearest neighbour match. The architechture is presented in figure 5.2.

TeCoMED. Further information about the TeCoMED system was given in [21]. Schmidt and Gierl attempt to use a prognostic model to forecast waves of influenza epidemics, based on earlier observations done in previous years. TeCoMED combines CBR with Temporal Abstraction to handle the problem of the cyclic but irregular behaviour of epidemics. Trends are discretized into *enormous decrease*, *sharp decrease*, *decrease*, *steady*, *increase*, *sharp increase*, and *enormous increase*, based on the percentage of change. TeCoMED utilizes former courses and similar cases in a way similar to early kidney problem warnings in the ICONS system. Attempting to commercialize the system, a small software company has incorporated warnings that are generated by the system into web pages of a health insurance scheme and a page of the health authority of the federal state.

Montani *et al.* [16] attempt to integrate different methodologies into a Multi-Modal Reasoning (MMR) system, used in therapy support for diabetic patients. The authors argue that most systems trying to utilize more than one methodology do so only in an exclusive fashion, with methodologies functioning merely as extensions to one another. Montani argues that a MMR system needs much closer integration of technologies to get the full benefits of a multi-modal solution. Integration allows tackling well known problems of single methodologies, i.e. the qualifi-

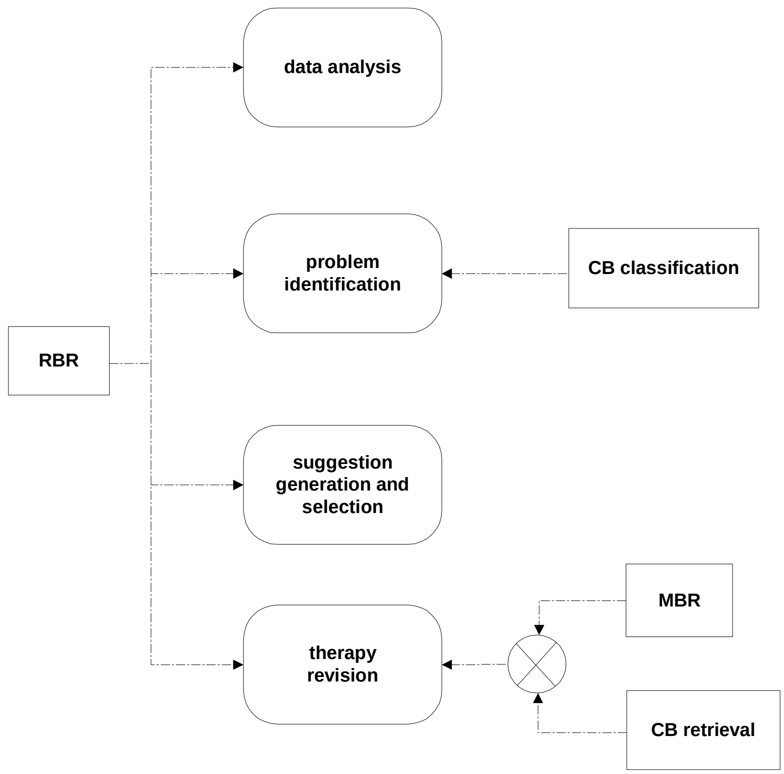

Figure 5.3: caption.

cation problem in RBR and the too-small-a-library problem in CBR. The proposed system tries to use a fuller integration and utilize CBR, Rule-Based Reasoning, and Model-Based Reasoning (MBR). The system architechture is illustrated in figure 5.3.

Perner *et al.* [19] has developed a system for the identification of airborne fungi. The fungal strains have a high biological variability, i.e. dissimilarity between the features of individual fungi is quite extensive. A strain can not be generalised to a few cases because of this variability. The images used originate from microscope enhanced pictures. A case is described by attributes (features) derived from the images. Attributes are in the abstraction level of colour, shape, size etc. New and original cases (descriptions of individual fungi) are retained in the case library, which is constructed by decision tree and prototype learning methods.

5.3.3 Tutoring systems

WHAT [9] is a tutoring medical CBR system for the education of sports medicine students. WHAT is designed to give better matching exercise prescriptions than the conservative rule-based approach taught by most books. The system provides two separate recommendations for exercise prescriptions, one which is based on the rules found in the books, the other uses CBR with a stored case base made by an expert. The prescribed exercises are applied to cardiac and pulmonary disease patients, as well as issues of general health and lifestyle. The prescriptions are based on features from the patients' medical history and on physiological tests.

Bichindaritz *et al.* [6] have evolved CARE-PARTNER into a medical training system on the Internet. The intention is to help medical students improve their knowledge by solving practice cases. Prototypical cases consist of clinical pathways, which can be tailored to generate cases of varying levels of complexity. The system is also able to evaluate the solutions given by the students for the practice cases. Due to the unlikelihood that a student solution matches the stored solution exactly, a correctness score is calculated and the student solution is placed into one of three categories: Fails to meet standards, Adequate, and Meets all standards.

5.3.4 Planning systems

The **Auguste** project [14], is an effort to provide decision support for planning the ongoing care of Alzheimer's Disease (AD) patients. The first reported system prototype supports the decision to prescribe neu-

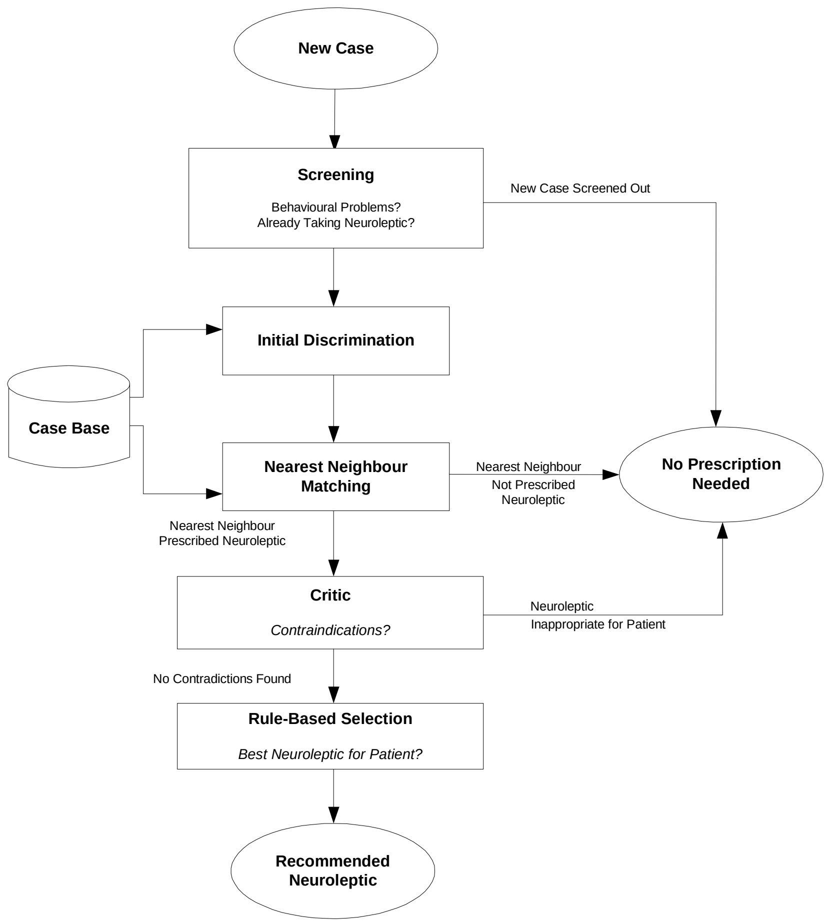

Figure 5.4: caption.

roleptic drugs for behavioural problems. The prototype is a hybrid system where a CBR part decides if a neuroleptic drug is to be given, and a Rule-Based Reasoning (RBR) part decides which neuroleptic to use. The system uses approximately 100 features, manually extracted from medical charts, in each case for determining the right neuroleptic drug. The patient is initially screened for behavioral problems before a Nearest Neighbour match makes a suggestion on whether or not to give neuroleptics to the patient. If the CBR module finds it appropriate to give neuroleptics and no contradictions are found, e.g., allergies to certain drugs etc., the RBR module determines which neuroleptic (of five available) to use. This prescriptive task, although termed "planning" in the vernacular, may be best characterized as one of design. The design of the Auguste system is depicted in figure 5.4

Davis *et al.* [8] are using a planning system based on the ReCall CBR shell. The system decides what kind of SMARTHOUSE devices disabled and elderly people need in their homes for independent living. Features are constructed from manual translations of written reports. The system contains 10 clustered problem space groups and 14 solution groups. Every group is subdivided by a C4.5 decision tree for efficiency and as an easy way to explain the reasoning process.

5.4 Trends in medical CBR

Naturally, the selection of papers in the previous section is highly subjective. None the less, certain trends are distinctive enough to deserve mentioning.

5.4.1 Property matrix

The research papers used as underlying documentation for the system descriptions does not always contain sufficient information about whether or not a system exhibits a certain construction-oriented property. For completion, the system authors were therefore contacted and asked specifically about the missing property information. Additionally, the authors were asked about the practical use of the systems in every-day life and whether there had been any attempts at commercialization. The answers to the questionnaire are presented in Figure 5.5.

Notably, the majority of systems are multi-modal. Only one of the systems utilizes adaptation. Generalisation using prototypes appears to be rare; however, in several projects the intention is to extend the system with prototypes at a later stage. The majority of systems are dependant on some level of user interaction in the reasoning cycle.

A few of the systems has been commercialized to some degree, but typically the projects are kept on a research level. Safety and reliability constraints are not too common. Systems that do have safety-critical constraints usually depend on operational reliability as well.

5.4.2 Construction-oriented trends

Looking at the previously defined construction-oriented properties, a number of trends can be identified.

Hybrid systems, also commonly referred to as Multi-Modal Reasoning Systems, constitute the majority of medical CBR systems. The combination of CBR with assisting technologies seems especially successful when CBR acts as the top level coordinator at the system level. Medical systems based on a straight CBR approach may suffer from unreliability, since all reference information is concentrated to the cases. Hence, systems like CARE-PARTNER have built in safeguards.

The autonomicity of the majority of systems is relatively low. Considering the inherent problem of unreliability in CBR, and the fact that systems typically does not reach a 100% correspondence with reported correct solutions even for controlled sets of cases, not relying on complete autonomicity appears to be sound.

The use of prototypes through case aggregation seems to be a commonly intended future extension, although only partly apparent in the property matrix. Prototypes are already used by many of the systems created by Gierl and Schmidt (as described in Diagnostic Systems), and prototype support is planned for both TeCoMED and WHAT.

Author / System	Cases	Prototypes	Adaptivity	Hybridity
Schmidt/TeCoMED	3000	No (intended)	No (intended)	CBR most imp.
Nilsson/Stress diagnosis	20	No	No	CBR most imp.
Montani/Hemodialysis	1000	No	No	Pure CBR
Montani/Diabetes (MMR)	150	No	No	Hybrid
Costello/Gene finding	948	No	No	Pure CBR
Evans-Romaine/WHAT	25	No (not yet)	No (not yet)	CBR most imp.
Marling/Auguste	28	No	No	Hybrid
Perner/Fungi identification	100	Some extent	No	Pure CBR
Perner/Image segm.	1000	Some extent	No	
El Balaa/FM-Ultranet	130	No		Pure CBR
Bichindaritz/CARE-PARTNER	4000	Some extent	Largely	CBR most imp.

Author / System	Interaction (autonomicity)	Commer-cialisation	Every-day use	Safety criticality	Reliability
Schmidt/TeCoMED	None	Some extent	Largely	No	Soft const.
Nilsson/Stress diagn.	Some extent	No	No	No	No
Montani/Hemodialysis	Some extent	No	No	No	No
Montani/Diabetes	Some extent	No	Some extent	No	No
Costello/Gene f.	None (intended)	No	No	No	No
Evans-Romaine/WHAT	None	No	No	Soft const.	Soft const.
Marling/Auguste	Largely	No	No	Hard const.	No
Perner/Fungi identif.	Some extent	Some extent	Largely		No
Perner/Image segm.	Some extent	Planned	Planned		No
El Balaa/FM-Ultranet	Some extent	No	No	Soft const.	
Bichindaritz/C.-P.	Some extent	No	Some extent	Hard const.	Must work

An empty cell in the matrix denotes that the property could not be determined.

Figure 5.5: System property matrix.

5.4.3 Overall trends

The majority of systems in the purpose-oriented category belong to classifying and diagnostic systems. True to the nature of the domain, the emphasis in the medical AI domain has and probably will continue to be on clinical use, i.e., systems involved in some sort of treatment.

Features and feature extraction is an important part of most CBR systems. One identifiable trend in medical CBR is the continuation of separate pre-processing methods on the input data, whether it is a human or an automated process. The datasets are often too large for a direct CBR analysis, and therefore needs to be pre-processed. Examples of systems focusing on separate feature extraction are the stress diagnosis system by Nilsson *et al.* and the airborne fungi detection system by Perner *et al.*

As was the case in the 1998 medical CBR survey by Gierl and Schmidt [10], medical tutoring systems utilising CBR are rare. The inherent case- and example-based nature and the cognitively plausible model of CBR should be ideal for teaching medical knowledge; still the number of tutoring systems is remarkably low. There is however an increasing number of systems that could partly be seen as tutorial, i.e. the system covers more than one of the purpose-oriented properties, including the Auguste project, WHAT, and FM-Ultranet.

5.5 Conclusions

Although the recent five years has not seen any dramatic changes in the construction and use of medical CBR systems, the field is evolving steadily but slowly. The potential for automated systems in clinics is high, but has yet to reach its full potential. Most systems tend to concentrate on diagnostic tasks, but the use of CBR for therapeutic planning appears to be on the increase. Medical tutoring systems based on CBR are still rare.

The clear majority of systems built around a combination of CBR and other AI methods indicates that most medical domain problems looked into by researchers in recent years have been too complex and multi-faceted to handle using CBR alone. Arguably, hybrid systems have been utilized in the CBR health science domain from the very beginning, with early projects such as CASEY [13] utilizing a mixture of CBR and RBR. There is, however, an increasing majority of hybrid systems being developed, which appears to reflect the increasing complexity and scope of the problem domains.

5.6 Acknowledgments

The authors would like to thank the participants of the questionnaire for the feedback, on which parts of the findings were based.

Bibliography

[1] Ziad El Balaa, Anne Strauss, Philippe Uziel, Kerstion Maximini, and Ralph Traphner. Fm-ultranet: a decision support system using case-based reasoning, applied to ultrasonography. In *Workshop on CBR in the Health Sciences*, pages 37–44. ICCBR'03, June 2003.

[2] Ziad El Balaa and Ralph Traphoner. Case-based decision support and experience management for ultrasonography. In *German Workshop on Experience Management*. GWEM'03), April 2003.

[3] Ellis Raymond Bareiss. *Exemplar-Based Knowledge Acquisition: A unified Approach to Concept Representation, Classification and Learning*. UMI, 300 North Zeeb road, Ann Arbor, MI 48106-1346, 1989.

[4] Isabelle Bichindaritz. Solving safety implications in a case based decision-support system in medicine. In *Workshop on CBR in the Health Sciences*, pages 9–18. ICCBR'03, June 2003.

[5] Isabelle Bichindaritz, Emin Kansu, and Keith M. Sullivan. Case-based reasoning in care-partner: Gathering evidence for evidence-based medical practice. In *Advances in CBR: 4th European Workshop*, pages 334–345. ECCBR'98, September 1998.

[6] Isabelle Bichindaritz and Keith Sullivan. Generating practice cases for medical training from a knowledge-based decision-support system. In *Workshop Proceedings*, pages 3–14. ECCBR'02, September 2002.

[7] Edwin Costello and David C. Wilson. A case-based approach to gene finding. In *Workshop on CBR in the Health Sciences*, pages 19–28. ICCBR'03, June 2003.

[8] Genevieve Davis, Nirmalie Wiratunga, Bruce Taylor, and Susan Craw. Matching smarthouse technology to needs of the elderly and disabled. In *Workshop on CBR in the Health Sciences*, pages 29–36. ICCBR'03, June 2003.

[9] Kathleen Evans-Romaine and Cindy Marling. Prescribing exercise regimens for cardiac and pulmonary disease patients with cbr. In *Workshop on CBR in the Health Sciences*, pages 45–52. ICCBR'03, June 2003.

[10] Lothar Gierl and Rainer Schmidt. Cbr in medicine. *Case-Based Reasoning Technology, from Foundations to Applications*, pages 273–297, 1998.

[11] Wayne Goodridge, Hadrian Peter, and Akin Abayomi. The case-based neural network model and its use in medical expert systems. In *Artificial Intelligence in Medicine*, pages 232–236. Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making (AIMDM1999), June 1999.

[12] Marie-Christine Jaulent, Christel Le Bozec, Eric Zapletal, and Patrice Degoulet. A case-based reasoning method for computer-assisted diagnosis in hisopathology. In *Artificial Intelligence in Medicine*, pages 239–242. AIME'97, March 1997.

[13] Phyllis A. Koton. *Using Experience in Learning and Problem Solving.* MIT Press, 1988.

[14] Cindy Marling and Peter Whitehouse. Case-based reasoning in the care of alzheimer's disease patients. In *Case-Based Research and Development*, pages 702–715. ICCBR'01, July/August 2001.

[15] Stefani Montani, Luigi Portinale, Giorgio Leonardi, and Riccardo Bellazi. Applying case-based retrieval to hemodialysis treatment. In *Workshop on CBR in the Health Sciences*, pages 53–62. ICCBR'03, June 2003.

[16] Stefania Montani, Paolo Magni, Abdul V. Roudsari, Ewart R. Carson, and Riccardo Bellazzi. Integrating different methodologies for insulin therapy support in type 1 diabetic patients. In *Artificial Intelligence in Medicine*, pages 121–130. AIME'01, July 2001.

[17] Markus Nilsson, Peter Funk, and Mikael Sollenborn. Complex measurement classification in medical applications using a case-based approach. In *Workshop on CBR in the Health Sciences*, pages 63–72. ICCBR'03, June 2003.

[18] Petra Perner. An architeture for a cbr image segmentation system. *Journal on Engineering Application in Artificial Intelligence*, 12(6):749–759, 1999.

[19] Petra Perner, Thomas Gunther, and Horst Perner. Airborne fungi identification by case-based reasoning. In *Workshop on CBR in the Health Sciences*, pages 73–79. ICCBR'03, June 2003.

[20] Rainer Schmidt and Lothar Gierl. Case-based reasoning for medical knowledge-based systems. In *German Workshop on Experience Management*, pages 720–725. Proceedings of MIE'00 and GMDS'00, 2000.

[21] Rainer Schmidt and Lothar Gierl. Prognostic model for early warning of threatening influenza waves. In *German Workshop on Experience Management*, pages 39–46. GWEM'02, March 2002.

[22] Rainer Schmidt, Bernhard Pollwein, and Lothar Gierl. Experiences with case-based reasoning methods and prototypes for medical knowledge-based systems. In *Artificial Intelligence in Medicine*, pages 124–132. AIMDM'99, June 1999.

[23] Rainer Schmidt, Dagmar Steffen, and Lothat Gierl. Evaluation of a case-based antibiotics therapy adviser. In *Artificial Intelligence in Medicine*, pages 462–466. AIME'01, July 2001.

[24] Edward H. Shortliffe. *Computer-Based Medical Consultations: MYCIN*. Elsevier, North Holland, New York, NY, 1976.

Chapter 6

Paper B: Complex Measurement Classification in Medical Applications Using A Case-Based Approach

Markus Nilsson, Peter Funk, Mikael Sollenborn
International Conference on Case-Based Reasoning (ICCBR'03), workshop on the health sciences, workshop proceedings, pages 63-72, June 2003

Abstract

In many domains it is sufficient to classify measurements obtained by sensors by filtering noisy data and performing a classification of the measurements by means of some mathematical function. Classification of measurements is difficult in some medical domains, e.g. determination of stress levels. Measurements and their classification may be too complex for an algorithmic approach. Matching may result in a number of classification candidates, historic data (previously classified measurements) may be sparse, and same measurements may indicate several diagnoses. We propose a Case-Based Reasoning approach in which feature vectors are matched against cases in a case library, and in which indeterministic or weak classifications are validated by an experienced physician, pointing out the features relevant in the classification of psychophysiological dysfunctions. The physician's classifications are stored in the case library, continuously improving the system's performance as it is being used. If only a few examples of classified measurements are available, the experienced physician initially points out which feature combinations are used in the classification/diagnosis.

6.1 Introduction

In many application domains, classification of complex measurements is essential in a diagnosis process. Correct classification of measurements may in fact be the most critical part of the diagnostic process. Diagnosing stress, a psychophysiologic dysfunction, is a medical application in this category and measurement classification is difficult and complex even for a physician. Case-based reasoning (CBR) is a concept that is recognized in medical domains and has been applied in a number of medical diagnosis projects [12], but not for direct classification of complex sensor readings from patients. CBR has good potential for many medical applications since reasoning from cases is commonly applied in medicine [5] and also reduces the expertise bottle-neck [2].

Physicians measure physiological activities in the patient with a number of sensors under different conditions. When measurements made under different conditions are obtained a physician can begin to analyse the measurements. One such analysis is the balance and recovery of sympathetic and parasympathetic levels in stressed situations through observing changes of the heartbeat interval [10]. This enables the detection of dysfunctions and abnormalities, ideally before the patient notices any severe physiological symptoms. These dysfunctions can in some cases be corrected with biofeedback training [3].

Figure 6.1 shows a standard diagnosis and treatment session. Only experts in the field are able to perform a reliable classification and incorrect classification of measurements may lead to serious risks for the patient. Classifications of measurements may also have complex relations to other measurements, complicating classification further. The skill of classification is difficult to transfer to less experienced physicians.

In this paper we propose an interactive case-based classification technique (ICBC) for classification of medical measurements from patients with stress symptoms. The application has similarities with Case-Based image interpretation [11] but with a stronger focus on features. ICBC elicits classifiable features from patient measurements, and at the same time creates a language of features in which physicians can discuss measurements. ICBC helps physicians determine the features on which they are to base their classification[1] and record these for reuse by other prac-

[1]It was observed that experienced physicians are able to perform complex classi-

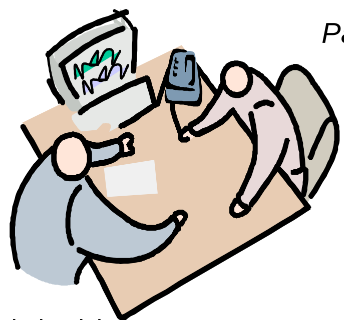

Figure 6.1: Patient and physician in a diagnosis and treatment session.

titioners (experience sharing among experts). Feedback from physicians is used together with previously stored feedback to refine matching and to adjust the weighting of features, an interactive process applied in some CBR systems. The value of Case-Based Reasoning for diagnostics in medical applications has been recognized earlier, e.g. in diagnosing myocardial infarctions [7].

6.2 Classifying Measurements

This section gives a short background to measurement classification and outlines some of the methods and techniques used to prepare, analyze and classify measurements, especially in the medical domain.

fications without being able to point out all the features on which they base their classifications and diagnosis. The ICBC approach transforms implicit knowledge into an explicit, sharable and transferable skill.

6.2.1 Preparation and Filtering

Filtering is used to restore the measurement to its "original" shape, e.g. by removing noise and distortions. A number of different techniques, such as adaptive filters, spectral subtraction and linear prediction models have been developed for filtering and noise reduction (see [1, 13]). The filtering process may be complicated in some domains by distortion and certain features important for classification having similarities. Filtering is a sensitive process and if not well performed, may remove essential information necessary for a correct classification. Domain knowledge, e.g. the relation between different measurements and heuristics may be used to improve filtering. The filtering process should be visible to and adjustable by the experienced physician.

6.2.2 Features and Feature Vector

Physicians classifying measurements often express themselves in terms of the presence or absence of features. The approach using feature vectors (example in figure 6.2), as commonly used in CBR systems, is well received amongst physicians. In many CBR applications feature selection is obvious, but in medical domains selecting features and creating a feature vector may require more effort.

[[peak-to-valley 45],[RR frequency 0.2 Hz],[small notch 2.46-2.49],[strong incline 0-2.87]]

Figure 6.2: Example of a simplified feature vector (this is an unusually short feature vector).

Physicians may look for features on a more intuitive basis and a cardiologist may need only inspect an ECG (Electrocardiogram) curve to be able to propose a diagnosis on the basis of his/her experience. Surprisingly a physician may not always be able to point out all the features which he/she uses to classify the measurement. The ICBC system will aid in the identification of these intuitive features and make them explicit. Features may also be derived from the sketches of features of importance made by experts. Feature identification functions may be extracted automatically from the drawings or constructed semi-automatically (with or without interaction with an expert). Fourier

Transformation (FT) [6] is a common technique transforming an analog signal to a frequency diagram/phase diagram. FT analysis is a powerful technique used when searching for specific frequency-dependent features in signals in medical applications [14] and a physician may need to consider the absence or presence of certain frequencies.

6.2.3 Classification Process

A number of different methods are available for the classification of measurements. The selection of classification method is based on the complexity of the task. A simple classification may only require a single test (e.g. above or below $37C$) for a complete classification. A set of complex measurements is classified by physicians comparing their features with the known features of previous measurements.

A different approach to the classification of similar feature vectors is to use Artificial Neural Nets (ANN). The performance of reliable classification using the ANN approach requires training with large numbers of classified measurements. If an interesting new case with an important classification becomes available, the net must be retrained, and only after this retraining does the new classification reflect the additional experience. In the domain of diagnosing stress, the number of classified measurements required for training is not always available.

6.3 Case-Based Categorization of Measurements

This section gives an overview and introduction to the interactive case-based classification approach, ICBC. The different steps, pre-processing, feature identification and curve classification are described in sections 6.4, 6.5 and 6.6, respectively. Online measurements are received from a patient to whom sensors are attached as shown at the left in figure 6.3. In medical applications, sensors mostly provide a continuous analog output. An analog/digital converter is used to produce a bit stream arriving at the pre-processing. The pre-processing process in figure 6.3 is responsible for filtering and restoring the measurement to a state as close as possible to its original state. The pre-processing also uses a library with models of known distortions to simplify the restoration process as explained in section 6.4.

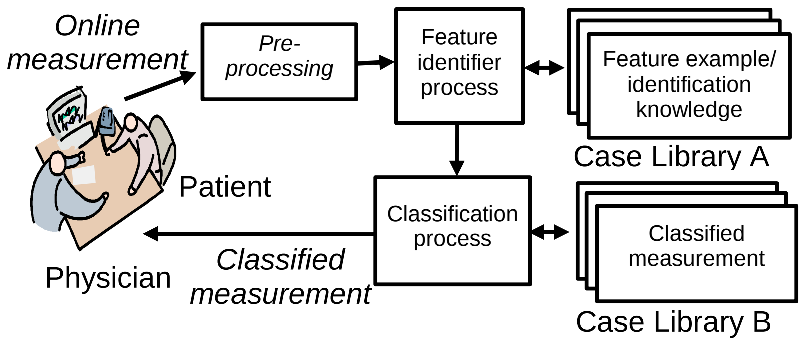

Figure 6.3: Schematic picture of the ICBC system .

In the feature identification process, ICBC uses a two pass model, first identifying features and then creating a vector with features. For feature identification a library (case library A in figure 6.3) is used to identify those relevant. In the second phase of the feature identification, more complex features based on other features and relations to other measurements are identified, see section 6.5 for more details.

Once the features are identified, the system classifies the feature vector. The classification is based on previously classified measurements (case library B) in figure 6.3. When a new measurement has been classified, the new case is added to the case library. In medical applications, the same measurement may be classified differently based on patient characteristics. The classification process is described in section 6.6. The measurements are shown to the physician together with their classification. In the research prototype (not all parts have been fully implemented) normal measurements are shown in green and measurements indicating dysfunction are shown in red.

6.4 Pre-processing

Sensor data typically contain noise and distortion. Before a classification attempt is made in ICBC, the sensor data received is pre-processed to remove as much noise, distortion and unsound data as possible. In figure 6.4 the two pre-processing steps are shown; identification and removal of noise and unsound data (first left box), and restoration of distorted data (right box). A library with models of known noise, distortions and how to identify and restore unsound data is used (bottom part in figure 6.4).

Noise may be caused by internal/external interference such as electrical interference. Noise is handled in the same way as distortions (next section). If the preconditions relating to certain noise which must be filtered have not been set in the model, all the data must be filtered.

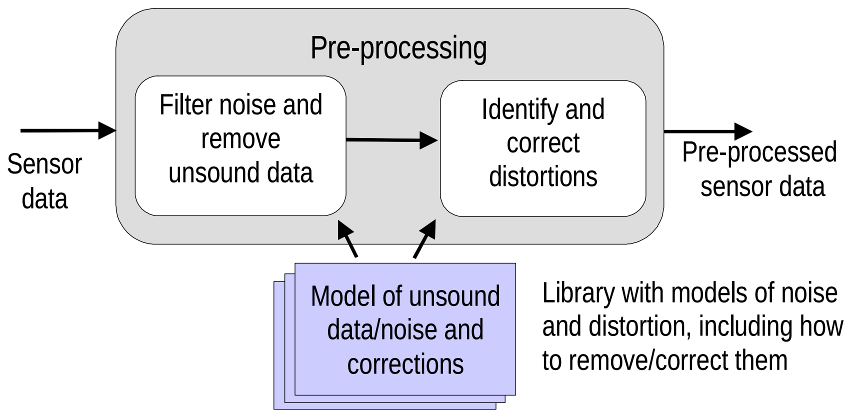

Figure 6.4: Pre-processing of sensor data in the ICBC process.

Distortion may be caused by the measurement procedure (e.g. by the patient moving while being measured) or an imperfect sensor. In the restoration step, we remove known noise from the data received and correct any distortion using the models and corrections recorded in the model library. The models are of different types; simple limit triggering rules, neural nets, a case-based reasoning system in itself etc. A model has a specific task, to restore data to its original state, to remove noise and to correct distortions.

6.4.1 Restoring Data Using Models

To restore the data, knowledge of what to expect of the data is needed, e.g. characteristics of sensors, including the hardware and the channel conveying the signal. This information permits the construction of models which can be used to identify and remove distortions. The models may also include correction functions such as compensation for data delays, amplification of signals etc. A model should restore the signal to its original state with respect to the particular distortion. When sensor data enters the pre-processing stage, the models stored in the model library are applied to determine if they should trigger and perform relevant correcting actions. In a default situation the correction actions should always be applied. Unsound data is data that could not possibly have originated as sensor output, given knowledge about the sensor. Models may be constructed to identify and correct unsound data.

6.5 Feature Identification Process

It is necessary to find a suitable form in which to represent and compress the information from the sensor data while storing enough information to be able to classify data correctly. The pre-processed data is divided into periods, e.g. individual heart beats. Features are identified within the period and stored in a feature vector (left box in figure 6.5). All feature vectors are temporarily stored in a database (lower middle in figure 6.5). A complex feature is a feature that is based on other features in more than one feature vector. Complex features are identified and appended to the vector.

The first task in the classification is to identify a number of features sufficient to perform identification of similar measurements. Figure 6.2 gives an example of a measurement containing some features which the expert believes should be used in classification. The features are automatically recognised and used to find similar cases in the case library. If the expert classifies this measurement in the same way, these features are sufficiently accurate to classify the measurement examples correctly. The system may also select a number of near misses and ask the expert to classify these, and if they are classified correctly, the classification is validated with respect to the selected cases.

The system can thereafter be tested with new measurements, previously classified by the expert but not used in the training. If the system is able to classify these correctly we can calculate the accuracy of the system. If the accuracy is sufficient, the CBR system can be used for matching and reuse of experience.

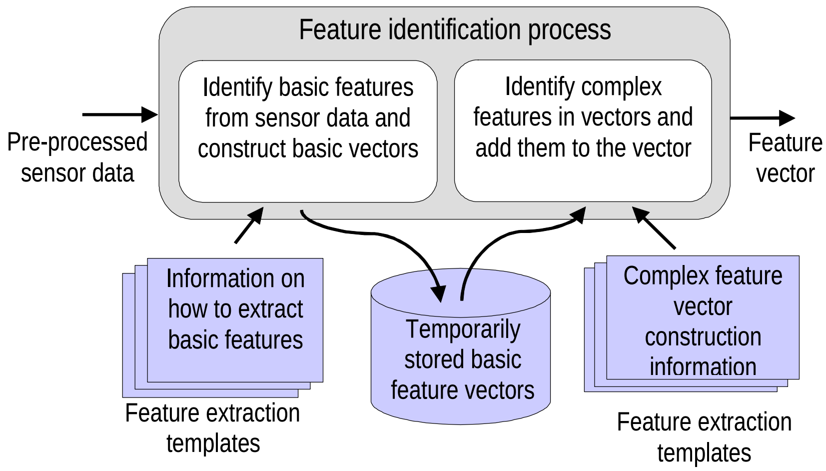

Figure 6.5: Feature identification in the ICBC process.

6.5.1 Feature Identification and Extraction

Extracting the appropriate features is dependent on the kind of data to be analysed. For heart rate and carbon dioxide levels we use angle of slopes, notches and peak values, period length etc. Figure 6.2 illustrates some of these features. If we obtain a derivative of the period data and combine this information with the measurement before derivation, we can detect notches and slopes when and where they occur. As within the pre-processing step, we can use different models to extract additional information from the data. One such model studies the frequency spectrum in a FT (see section 6.2.2). The FT analysis enables us to find any heart rate variability in the frequency spectra [9] and the relationship between sympathetic and parasympathetic systems [8]. We can also read the Respiratory Sinus Arrhythmia (RSA) [4] in the FT. RSA is the natural rise and fall of the heartbeat rate controlled by the autonomic nervous system, the heart rate increasing with an inhalation. Different techniques may be used for the feature extraction models, stored in the lower left in figure 6.5.

6.5.2 The Feature Vector

A feature vector, illustrated in figure 6.2, is a fairly simple construction. The feature vector contains some of the features mentioned in the previous section. A feature vector is basically a list of features found in the same sample period, i.e. a description of a period with features instead of with measured values. The amount of data in the ICBC process is reduced in this representation form.

6.5.3 The Complex Vector

A feature vector is incomplete if all its dependencies on other measurements are not included. To classify a feature vector, we may need to include additional information which is not present in the period itself. This information denotes complex dependencies between feature vectors. All recently created feature vectors are stored in a temporary database. In this database we collect features from recent vectors with the same sensor source and also fetch and compare measurements with parallel or other sensors. By combining information from several feature vectors we can detect trends over time within the same measurement session. A complex dependency can also be a relation between feature vectors with different sensor sources. Finally we create a new feature vector by combining the complex features with the existing feature vector.

6.6 Measurements Categorisation

The measurements are finally classified as shown in figure 6.6. in which a feature vector in ICBC is used as a case. The feature vector is matched with a regular nearest neighbour algorithm against the case library with previously classified feature vectors. A scored list with the best matching cases is created, and sent to the patient customisation.

Patient information and physiological domain knowledge which may affect the final ranking is also taken into account, for example, a patient may have pre-requisitions that must be observed. A ranked classification list is created, based on a reordering of the scored case list and with the additional information from the patient customisation part. The

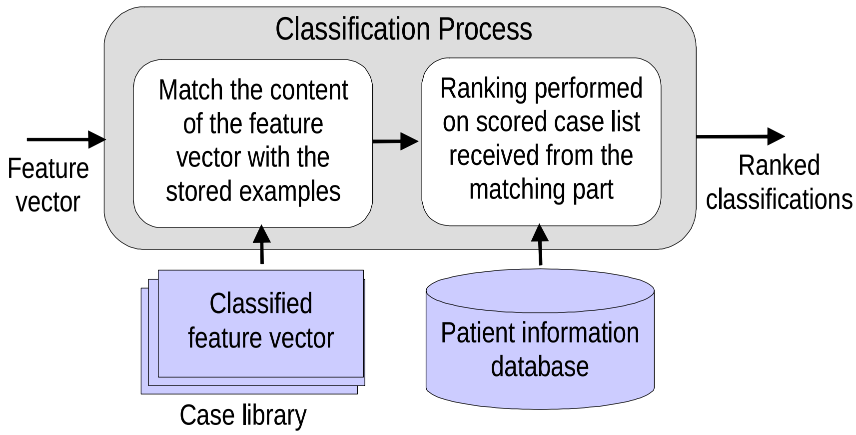

Figure 6.6: Classification process in detail.

ranked classification list with the classified complex feature vectors is now targeted at a specific patient or a group of patients.

6.7 Conclusions

Physicians diagnosing stress-related dysfunctions use different sensor readings to diagnose the patient. The measurement classification process is difficult and long experience is needed to learn the skill of making accurate classifications.

In this paper we outline how complex measurements are classified using an interactive case-based classification approach, ICBC. The classification task is divided into three steps, pre-processing, feature identification and classification. The pre-processing removes distortion in sensor readings caused by limitations in sensors, patient's movements and other outside interference. To remove the distortions, the pre-processing step uses a case library containing models of possible distortions and knowledge of how to remove them. In the feature identification step we identify and extract relevant features used by physicians to classify measurements and create a feature vector. Some features are based on relations between different measurements and their features; hence the feature identification process is divided in two parts as described in section 5. In the third step feature vectors are classified using a nearest neighbour search and a library with previously classified feature vectors. The measurements together with the classification of these are presented to the physician.

Bibliography

[1] M. Golay A. Savitzky. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8):1627–1638, 1964.

[2] Agnar Aamodt and Enric Plaza. Case-based reasoning: Foundational issues, methodological variations, and system approaches. *Artificial Intelligence Communications*, 7(1):39–59, 1994.

[3] Marie J. Cowan, Helen Kogan, Robert Burr, Sue Hendershot, and Lynne Buchanan. Power spectral analysis of heart rate variability after biofeedback training. *Journal of Electrocardiology*, 23:85–94, 1990.

[4] Tuan Pham Dinh, Helene Perrault, Pascale Calabrese, Andre Eberhard, and Gila Benchetrit. New statical method for detection and quantification of respiratory sinus arrhytmia. *IEEE Transactions on Biomedical Engineering*, 46(9):1161–1165, September 1999.

[5] Lothar Gierl, Mathias Bull, and Rainer Schmidt. Cbr in medicine. In *Case-Based Reasoning Technology: From Foundations to Applications*, pages 273–297. Lecture Notes in Artificial Intelligence, May 1998.

[6] Ralph D. Hippenstiel. *Detection Theory.* CRC Press, 2000 N.W. Corporate Blvd., Boca Raton, Florida 33431, USA, 2002.

[7] Marie-Christine Jaulent, Christel Le Bozec, Eric Zapletal, and Patrice Degoulet. A case-based reasoning method for computer-assisted diagnosis in hisopathology. In *Artificial Intelligence in Medicine*, pages 239–242. Proceedings of the 6th Conference on Artificial Intelligence in Medicine Europe, March 1997.

[8] M. Rostrup K. Sevre. Undersokelser av hjertefrekvensvariabilitet og baroreflekssensivitet. *Tidsskr Nor Laegeforen*, 26(121):3059–3064, 2001.

[9] Marek Malik and John Camm et al. Heart rate variability - standards of measurement, physiological interpretation, and clinical use. *European Heart Journal*, 17:354–381, March 1996.

[10] R. McCraty and M. Atkinson et al. The effects of emotions on short-term power spectrum analysis of heart rate variability. *American Journal of Cardiology*, 76(14):1089–1093, 1995.

[11] Petra Perner. Why case-based reasoning is attractive for image interpretation. In *Case-Based Reasoning Reasearch and Development*, pages 27–43. Proceedings of the 4th International Conference on Case-Based Reasoning, July/August 2001.

[12] Rainer Schmidt and Lothar Gierl. Case-based reasoning for medical knowledge-based systems. XVI International Congress of the European Federation for Medical Informatics, 2000.

[13] S. V. Vaseghi. *Advanced signal processing and digital noise reduction*. John Wiley, New York, 1996.

[14] Bo von Schele. *Classification Systems for RSA, ETCO2 and other physiological parameters*. PBM Stressmedicine, Heden 110, 821 31 Bollns, Sweden, 1999.

Chapter 7

Paper C: A Case-Based Classification of Respiratory Sinus Arrhythmia

Markus Nilsson, Peter Funk
European Conference on Case-Based Reasoning (ECCBR'04), Madrid, August 2004

Abstract

Respiratory Sinus Arrhythmia has until now been analysed manually by reviewing long time series of heart rate measurements. Patterns are identified in the analysis of the measurements. We propose a design for a classification system of Respiratory Sinus Arrhythmia through time series analysis of heart and respiration measurements. The classification use Case-Based Reasoning and Rule-Based Reasoning in a Multi-Modal architecture. The system is in use as a research tool in psychophysiological medicine, and will be available as a decision support system for treatment personnel.

7.1 Introduction

This paper describe a system for pattern classification of Respiratory Sinus Arrhythmia (RSA). The patterns are classified with Case-Based Reasoning (CBR) and Rule-Based Reasoning (RBR) through physiological time series measurements. The system is developed to be a decision support system for treatment personnel, as well as a research tool in psychophysiological medicine. The next paragraph defines RSA and put it into clinical context. The reader may skip the next paragraph as it is not required to understand RSA to comprehend the contents of this paper.

Respiratory Sinus Arrhythmia is described as centrally modulated cardiac vagal and sympathetic efferent activities associated with respiration [3]. RSA occurs because the heart rate, i.e. heart beats per minute, is variable. This Heart Rate Variability (HRV) is an effect of inhibitions on the sympathetic and parasympathetic systems while breathing. The sympathetic and the parasympathetic systems, which are a part of the autonomous nervous system, have different activity levels during different stages of the respiration cycle [7].

Physicians detect irregular heart rate patterns by analysing the RSA. Some of the irregularities are dysfunctions caused by physiological and/or psychological stress. A common diagnostic method for detecting dysfunctions in RSA is to manually analyse sampled heart rate measurements together with an analysis of the measurements' frequency spectrum [4, 3, 7]. The dysfunctions are treated with cognitive behavioural sessions with psychologists, and with biofeedback training [6].

Time-series analysis in medical Case-Based Reasoning has previously been studied by Montani *et al.* [9, 10], where they integrate CBR, RBR and Model-Based Reasoning (MBR) in a Multi-Modal Reasoning (MMR) platform for managing, i.e. suggesting insulin therapy, for type 1 diabetic patients. Another CBR system which analyses time series is ICONS [14, 15]. ICONS forecast kidney functions through an extended CBR cycle which abstracts states from measurements and trends from the states. Other related medical CBR systems are CARE-PARTNER [1, 2], Auguste [8] and Perner *et al.*'s airborne fungi detection system [13]. Further information of these systems can be found in Nilsson and Sollenborn's survey on medical CBR [12].

We propose a MMR system design for the classification of RSA, where CBR matches physiological parameters and RBR reduces the domain of cases. A system design for the classification of RSA is introduced in the next section. We evaluate the proposed system in section 7.3, and conclude the paper in section 7.4.

7.2 System Architecture

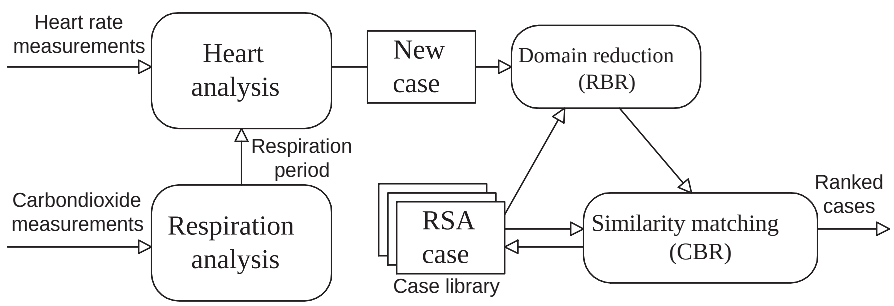

Figure 7.1: A design for a classification system of Respiratory Sinus Arrhytmia.

A classification system for RSA is naturally divided into two initial analytical stages. Each stage analyses time series measurements. The first stage analyses the respiration and the second stage analyses the heart measurements. Cases are there after created based on the findings in the analysis processes. Rules limits the number of cases for the matching procedure to compare to, and the cases that pass the filter are matched and ranked. The design is illustrated in figure 7.1. The system is a revised version of the two later parts of the design described in [11], the first part is processed in the hardware. Each part of the figure is detailedly described in the remainder of this section. The respiration analysis is described in subsection 7.2.1, followed by the heart analysis, a domain reduction, case matching and finally the user interface.

As RSA is quantified during a breath (a respiration cycle), a respiration analysis precedes the heart analysis. The respiration analysis locate where in time the respiration occur and passes that information to the heart analysis.

7.2.1 Respiration analysis

A breath begins, by definition, on an inhalation. Hence, the respiration cycle starts when an exhalation stops and inhalation begins. Capnograph [5] measurements are used to pinpoint the beginning and end of the respiration cycle. The capnograph is a non invasive method, and measures the contents of carbon dioxide (CO_2) in exhaled air. Capnograph measurements are depicted in figure 7.2.

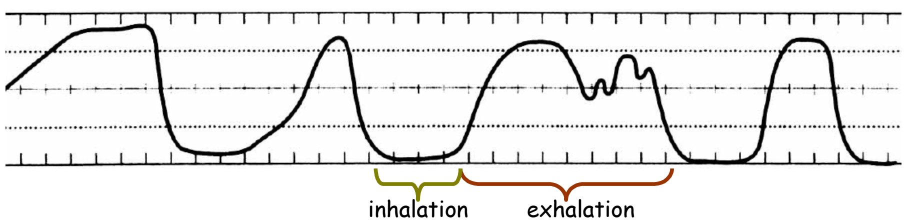

Figure 7.2: Capnography measurements illustrating the respiration cycle, divided into inhalation and exhalation. The picture is adapted from [5].

Finding either the beginning or the end of the respiration cycle is actually sufficient to determine the entire respiration cycle, since the end of a respiration cycle marks the beginning of the next. A new breath start, in the ideal case, when the levels of CO_2 dramatically drops from circa 5% to just above 0%, followed by a steadily low level. This low level of CO_2 occurs during the entire inhalation. The levels of CO_2 never reaches 0% because the surrounding air naturally contain CO_2, and it is also difficult to vacate the measuring sensor from all gases, even with a pump driven device.

A rough estimate of the respiration period is calculated by searching for a local maxima followed by a local minima. The maxima represents the exhalation and the minima the inhalation. A simulated annealing algorithm is then used on the first order derivates of the CO_2 measurements to find an approximate position between the maxima and minima. The position is where the exhalation stop and inhalation start, i.e. where the respiration cycle begins.

Two points are identified, the first as the beginning and a second as the end of the respiration cycle. The samples in the respiration period are shifted in time due to lag in the sensor and additional delays associated with capnography measurements. A major delay is the transportation of CO_2 from the measuring point to the sensor. The CO_2 is sucked through a tube with a pump. The corrected measurements are then sent to the heart analysis as seen in figure 7.1.

7.2.2 Heart analysis

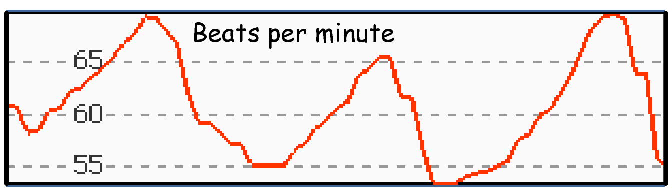

Figure 7.3: The heart rate variability, i.e. the oscillating effect of the heart can easily be seen in these heart rate measurements.

Physicians observe both the HRV and the frequency spectrum of the HRV when they classify RSA. The beginning and the end of a HRV period is based on the respiration analysis. The HRV period span over the same time period as the respiration period, and is calculated from heart rate measurements. The heart rate measurements are mean-valued electrocardiogram (ECG) measurements. The conversion from ECG to heart rate measurements are automatically computed in the hardware [1]. HRV measurements are depicted in figure 7.3.

[1]The AirPas and cStress hardware environments from PBM StressMedicine are used to measure physiological parameters.

$$\sum_{i=1}^{n}\left(HR(i)-\frac{\sum_{j=1}^{n}HR(j)}{n}\right)=0 \tag{7.1}$$

The frequency spectrum is calculated when the samples for the HRV have been determined. However, some pre-processing is required before a frequency spectrum can be calculated. The physicians are only interested in the oscillation of the sequence of samples, HR, that make up the HRV, when they observe the frequency spectrum. The sample sequence have to be shifted to oscillate around it's own mean value, as seen in equation 7.1. If not, a large portion of the lower end of the frequency spectrum is mixed with non relevant oscillations due to the nature of the heart rate samples. The heart rate samples are always positive numbers with a range of about 50-90 beats per minute, which unintentionally create large sine waves, or low frequencies within the measurement sequence.

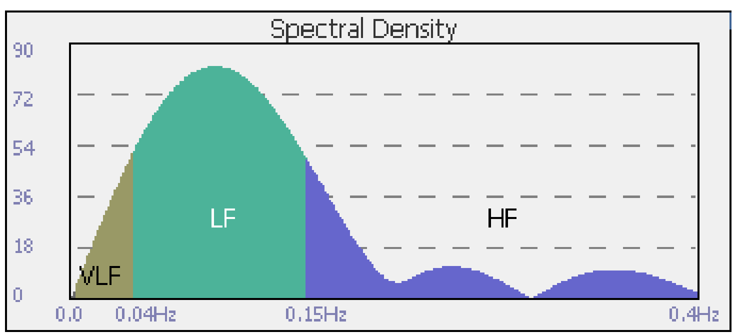

Figure 7.4: A frequency spectrum of a typical RSA. Physicians are only interested in the range from 0 to $0.4Hz$. The spectrum is divided in to three major frequency bands. Very Low Frequencies (VLF), Low Frequencies (LF) and High Frequencies (HF), as various physiological variables appear within these individual bands.

The output sample rate from the hardware sensors is 2 Hz; and a normal breath are in the range of 6-12 seconds. Hence, there are usually too few samples in the HRV to make any useful frequency transformation.

The solution is to pad, or to add, the sample sequence with zeroes. Padding with zeroes does not affect the frequency distribution in the spectrum. The sample sequence is padded to 2048 samples. The samples are then transformed to the frequency spectrum using a Fast Fourier Transformation (FFT). The length, or power value, of each frequency is calculated from the FFT's output of complex numbers, see equation 7.2, and figure 7.4.

$$Power(f) = \sqrt[2]{FFTreal(f)^2 + FFTimg(f)^2} \tag{7.2}$$

Physicians study additional parameters in their classification of RSA. The additional parameters are notch patterns and peak-to-valley differences in the heart rate measurements. The peak-to-valley value is the ΔY difference of the maximum and minimum heart rate sample values. Notches are irregular dips in the otherwise smooth heart rate oscillation. The notches have different significance depending on where they occur. Both peak-to-valley and notches are calculated.

7.2.3 Cases and domain reduction

Cases contain all above described parameters and measurements, with one addition, first order derivates of the heart rate measurements are also included. A case belong to one of the stereotypical classes of RSA identified in [16]. A class can contain an arbitrary number of example cases. The classes are clustered into larger groups, the clustering criteria is based on the number of notches the heart rate measurement contain[2]. A class is not limited to one group. A RSA class may end up in several clusters. Rules trigger new cases for notches. This determines which cluster of classes the matching procedure is to use.

7.2.4 Case similarities

A new case is matched with stored cases by calculating the similarity of the heart rate measurements and the heart rate frequency spectrum. The new case is matched with all cases in all the classes of the local cluster.

[2]Stereotypical classes and their clusters may change whenever new knowledge from psychophysiological research is available.

The frequency match calculates the distance between two frequency vectors by comparing the spectral density of the lowest common frequency window, for that specific frequency region. This is calculated throughout the entire length of the vectors. The phase, i.e. the angle, of the frequency is also taken in to account when the distance is calculated. A similarity of the entire frequency spectrum is compiled and normalised to a floating point number in the range of $0 - 1$. The heart rate measurements are matched through the first order derivates. The derivates of the new case are interpolated to match the number of derivates in the stored case. The distance is calculated for every pair of derivates and compiled to a normalised similarity number for the entire heart rate sequence.

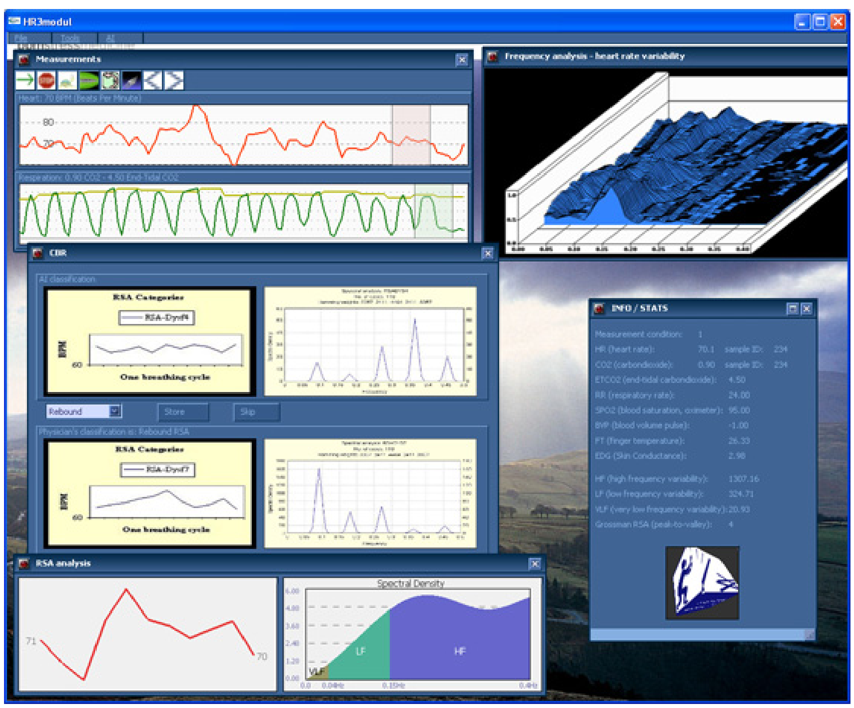

Figure 7.5: A screenshot of the application christened HR3modul. HR3modul is a tool for classification of Respiratory Sinus Arrhythmia.

The similarities of the measurements and the spectrum are merged to one similarity value for the entire case. The cases are ranked based on the similarity value. The cases with the closest similarity are presented to the user. The similarities of the measurements and frequencies are also available for the user.

7.2.5 User interface

As mentioned in the introduction, one of the systems task is to serve as a research tool for researchers in psychophysiology. Hence, a windowed environment was chosen to display the measurements. The user can freely choose what measurements or parameters he/she wants to work with, as well as enabling the RSA classification. A screenshot of the system is displayed in figure 7.5. The screenshot illustrate the complexity of classifying RSA.

The system is currently implemented in C++ as an application for the Windows platform. The application uses OpenGL to display graphics, as it is easier to port the application to other platforms in the future, due to OpenGL's OS independent interface.

7.3 Evaluation

This section contains a first evaluation of the RSA classification. The first evaluation was also the first time leading experts in the field of psychophysiology came in contact with the system.

The case-base was initialized with stereotypical cases produced by domain experts. The cases are described in [16]. The cases are supposed to cover all known classifications of RSA, i.e. cover the entire domain. Additional cases were also added to the case library. The additional cases belong to one of the stereotypical classes, and were added to facilitate an easier matching process. An example of an additional case is were the heart rate is constant during the entire respiration cycle, i.e., $\forall_i(s_i \in S : s_i = 0)$ after the conversion in equation 7.1. There exist no frequencies in a straight line.

7.3.1 Evaluation data set

A data set of approximate 100 pre-recorded measurements was used in the evaluation. The measurements are recorded in a cStress system, and are measured from a normal population of 17 year olds. Pre-recorded measurements are parsed and simulated in the HR3modul system as if they are real-time measurements streamed direct from hardware.

7.3.2 Results

The evaluation was conducted with the help of the domain experts. Cases of special interest for an accurate classification were pushed to the case library. An evaluation mode was enabled when the case library contained enough example cases of RSA. The evaluation mode collects statistics of the accuracy of the classification system. The case library used in the evaluation consists of approximately 50 cases. The cases represent the existing stereotypical classes of RSA.

However, as this was the first time the physicians had an opportunity to view every individual RSA, i.e. the HRV per respiration cycle; new patterns of RSA were discovered. This invalidates the notion of total domain coverage by the cases in the case library, since the new RSA patterns does not fit into any of the stereotypical classes described in section 7.3.

Nevertheless, statistics were collected from the evaluation. A summary of the statistics are presented in figure 7.6. The figure represents the accuracy, i.e. similarity, of the classification system in a comparison with a domain expert. The leftmost column represents the probability that the first RSA class suggested by the system is the same as the expert would choose. The second column from the left represents the probability that the expert's choice of class is the same as either the first or the second RSA suggested by the system. The rest of the columns proceed in the same manner, from left to right. All statistics of the similarities beyond the 5th suggested class have been summarised in to the rightmost column.

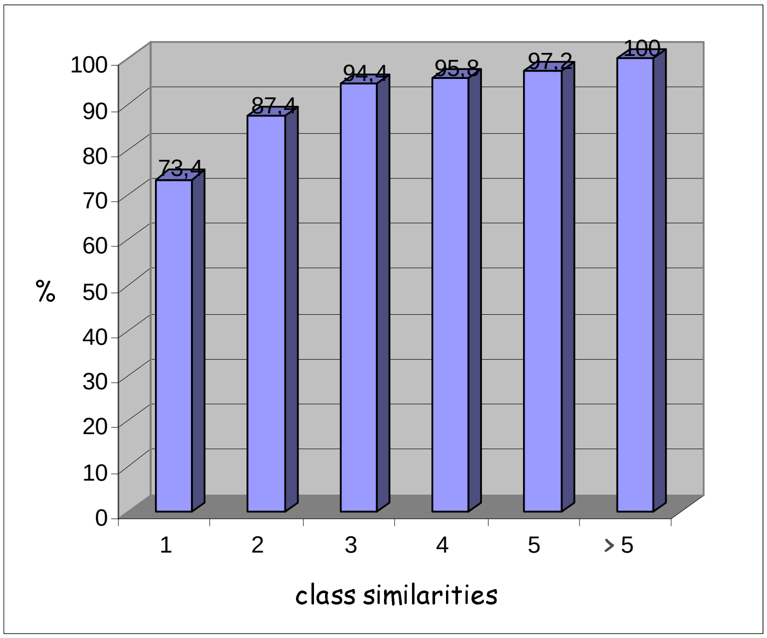

Figure 7.6: Evaluation of the classification system. The columns represent the probability of an accurate classification, ranging from similarity in the first suggested class to the 5th. The remaining classes, i.e. beyond the 5th, are summarised in the rightmost column.

7.4 Conclussions

We have presented a MMR design for the classification of RSA. The design uses two analytical stages of time series measurements from the heart and from exhaled air. The analytical stages process the time series measurements so they will conform in to cases, for later similarity comparisons. A RBR stage limits the number of RSA classes a new case have to be compared to, and a CBR stage make a similarity match with the cases from the remaining RSA classes.

The MMR design for the classification of RSA seems to be reliable, as 19 out of 20 cases in the evaluation data set was among the three top most suggested classes. The evaluation also showed that even the experts benefit from the system, as they discovered new patterns of RSA while using the system.

Bibliography

[1] Isabelle Bichindaritz, Emin Kansu, and Keith M. Sullivan. Case-based reasoning in care-partner: Gathering evidence for evidence-based medical practice. In *Advances in CBR: 4th European Workshop*, pages 334–345. ECCBR'98, September 1998.

[2] Isabelle Bichindaritz, Carol Moinpour, Emin Kansu, Gary Donaldson, Nigel Bush, and Keith M. Sullivan. Case based reasoning for medical decision-support in a safety critical environment. *Artificial Intelligence in Medicine*, 2780:314–323, November 2003.

[3] Tuan Pham Dinh, Helene Perrault, Pascale Calabrese, Andre Eberhard, and Gila Benchetrit. New statical method for detection and quantification of respiratory sinus arrhytmia. *IEEE Transactions on Biomedical Engineering*, 46(9):1161–1165, September 1999.

[4] P. Grossman and J. Van Beek et al. A comparison of three quantification methods for estimation of respiratory sinus arrhytmia. *Psychophysiology*, 27(6):702–714, November 1990.

[5] Bernard Landis and Patricia M.Romano. A scoring system for capnogram biofeedback: Preliminary findings. *Applied Psychophysiology and Biofeedback*, 23(2):75–91, 1998.

[6] Paul M. Lehrer, Evgeny Vaschillo, and Bronya Vaschillo. Resonant frequency biofeedback training to increase cardiac variability: Rationale and manual for training. *Applied Psychophysiology and Biofeedback*, 25(3):177–191, 2000.

[7] Marek Malik and John Camm et al. Heart rate variability - standards of measurement, physiological interpretation, and clinical use. *European Heart Journal*, 17:354–381, March 1996.

[8] Cindy Marling and Peter Whitehouse. Case-based reasoning in the care of alzheimer's disease patients. In *Case-Based Research and Development*, pages 702–715. ICCBR'01, July/August 2001.

[9] Stefania Montani, Paolo Magni, Riccardo Bellazzi, Cristiana Larizza, Abdul V. Roudsari, and Ewart R. Carson. Integrating model-based decision support in a multi-modal reasoning system for managing type 1 diabetic patients. *Artificial Intelligence in Medicine*, 29:131–151, 2003.

[10] Stefania Montani, Paolo Magni, Abdul V. Roudsari, Ewart R. Carson, and Riccardo Bellazzi. Integrating different methodologies for insulin therapy support in type 1 diabetic patients. In *Artificial Intelligence in Medicine*, pages 121–130. AIME'01, July 2001.

[11] Markus Nilsson, Peter Funk, and Mikael Sollenborn. Complex measurement classification in medical applications using a case-based approach. In *Workshop proceedings: Case-Based Reasoning in the Health Sciences*, pages 63–72. International Conference on Case-Based Reasoning, June 2003.

[12] Markus Nilsson and Mikael Sollenborn. Advancements and trends in medical case-based reasoning: An overview of systems and system development. In *Proceedings of the 17th International FLAIRS Conference*, May 2004.

[13] P. Perner, T. Gnther, H. Perner, G. Fiss, and R. Ernst. Health monitoring by an image interpretation system - a system for airborne fungi identification. In *Medical Data Analysis*, volume 2868 of *Lecture Notes in Computer Science*, pages 62–74, 2003.

[14] Rainer Schmidt and Lothar Gierl. Temporal abstractions and case-based reasoning for medical course data: Two prognostic applications. In *Machine Learning and Data Mining in Pattern Recognition*, volume 2123 of *Lecture Notes in Computer Science*, pages 23–34, 2001.

[15] Rainer Schmidt, Dagmar Steffen, and Lothat Gierl. Evaluation of a case-based antibiotics therapy adviser. In *Artificial Intelligence in Medicine*, pages 462–466. AIME'01, July 2001.

[16] Bo von Schele. *Classification Systems for RSA, ETCO2 and other physiological parameters*. PBM Stressmedicine, Heden 110, 821 31 Bollns, Sweden, 1999.