George D. Dounias and Abdul V. Roudsari (Eds.)

INTELLIGENT E-HEALTH APPLICATIONS IN MEDICINE
and
Other Medical Applications of Adaptive / Hybrid Intelligent Systems

Special Session Proceedings of the
EUNITE – 2002 Symposium
Albufeira, Portugal, September 19-21, 2002
Organized by IBA D: HUMAN MEDICAL & HEALTHCARE

Ειδική Έκδοση του Πανεπιστημίου Αιγαίου
Σχολή Επιστημών της Διοίκησης
Τμήμα Μηχανικών Οικονομίας και Διοίκησης
Φωστίνη 31, Χίος GR-82100

Τυπώθηκε σε 100 αντίτυπα με τη συνεργασία
του τυπογραφείου της ABC Διαφημιστικής, στη Χίο
Καλαμπόκα 20, Τηλ.: 02710 42308

Διανέμεται Δωρεάν κατόπιν
γραπτής παραγγελίας απευθυνόμενης
προς τους εκδότες στην ανωτέρω διεύθυνση.

Δείτε σχετικά και τον δικτυακό τόπο:
decision.ba.aegean.gr/book-orders

Special Edition of the University of the Aegean
Business School, Dept. of Financial and Management Engineering
31 Fostini Str., 82100 Chios, Greece

Printed in 100 items by:
ABC Chios Greece
Kalamboka 20, Tel. 02710 42308

Delivered free of charge.
Please mail your order to the editors in the above address.
Visit also the url: decision.ba.aegean.gr/book-orders

ISBN: 960-7475-20-8
A Publication of the University of the Aegean

To peace

Contents

Intelligent e-health Applications in Medicine and other Medical Applications of Adaptive / Hybrid Intelligent Systems: An Introduction to the Present Edition

Georgios Dounias* and Abdul Roudsari**

*University of the Aegean,
Dept. of Business Administration,
8 Michalon Str., 82100 Chios, Greece,
Phone: +30-271-35165, Fax: +30-271-93464
e-mail: g.dounias@aegean.gr

**City University
School of Informatics
Centre for Measurement and Information in Medicine (MIM),
Northampton Square, London, EC1V 0HB, UK
Phone: +44-020-7477-8367
e-mail: A.V.Roudsari@city.ac.uk

SUMMARY

This introductory text contains an initial description and categorization of the 20 papers included within this special edition, and provides a brief reference to the criteria set for the evaluation process applied to each paper. In total 51 scientists from different fields of specialization (medical doctors, computer scientists, engineers, etc.) contributed this year their research reports to this edition, representing 7 different European countries. The 51 contributors also represent approximately 13 different Universities, 6 research institutes and at least 6 different medical agents (i.e. medical companies, organizations or clinics, and medical institutions belonging to Universities). Most of the contributors have a direct and continuous relation and experience to medical decision-making and clinical environments, while the rest of them, represent university research groups and research institutions, mainly working in computer science. In respect to intelligent methodologies, the majority of the papers included in this edition refers primarily to decision support systems, fuzzy rule based systems, and artificial neural networks and secondarily to aspects of learning, data mining and genetic optimization. The most popular domain of application within the collection of papers is with no doubt, that of diabetes.

INTRODUCTORY CONCEPTS AND INTELLIGENT E-HEALTH

As it was stated in [1], [4], computational intelligence tools and techniques have become the method of choice for problems and domains having specific characteristics, such as, high degree of complexity, linguistic representation of concepts or decision variables, high degree of uncertainty, lack of precise or complete data sets, etc. Some areas of computational intelligence became particularly famous over time, due to their effectiveness in facing specific real-world problems. Nowadays, fuzzy rule-based systems and soft computing, neural networks, evolutionary computation and finally, machine learning, are the most recognised sub-topics of computational intelligence. Computer science lately focuses in the discovery of combinations of intelligent tools and techniques that make sense in a theoretical and also a practical basis, such as hybrid (i.e. containing more than one intelligent components) schemes, or adaptive systems, or both the above. The majority of the application domains with a successful use of computational intelligence tools and techniques is closely related to medical and biomedical problem-solving. During the last two years, numerous scientific meetings and events have been taken place with the support and involvement of the Human, Medical and Healthcare

Intelligent e-Health Applications in Medicine
Special Session Proc. EUNITE 2002, Albufeira, Portugal, Sept 19-21, 2002
Published by the University of the Aegean

Committee related to medical applications of intelligent systems, see for example the workshop[1] on fuzzy diagnostic and therapeutic decision support [2], the workshop[2] on intelligent systems in patient care [3], or the organisation of an interdisciplinary summer school teaching soft computing in medicine to medical doctors and engineers[3], as well as last year's special sessions [4]. This year's special session deals with the construction of intelligent e-health schemes, i.e. usually web-based or telemedicine-based decision support systems, used for specific real-time / real-world medical applications.

AIMS AND SCOPE OF EUNITE'S "IBA D: HUMAN MEDICAL & HEALTHCARE"

EUNITE (European Network on Intelligent Technologies) came into existence in January 2001, being the successor to ERUDIT (European Network of Excellence on Fuzzy Technology), which had been funded by the European Union over a period of 5 years. The scope of EUNITE has been extended significantly from that of ERUDIT including two main additional areas. These are Smart Adaptive Systems (SAS) and Hybrid Systems (or Integration and Hybrid Systems (IM)). The structure of EUNITE provides for two Research, Theory and Development (RTD) groupings on SAS and IM. The core of the industrial applications areas are divided into 5 IBA (Industrial and Business Applications) tracks, one of which is that on Human, Medical and Healthcare (HMH) which is responsible for this Proceedings volume. In addition, there are committees responsible for Technology Transfer (into industry) and Training and Education (including Web-based delivery).

The aims of each IBA include participation in the Annual Symposium of EUNITE, the first held in Tenerife last year, on 13, 14 December 2001, see [4]. Smaller Workshops are also being organised in member states, the HMH IBA having held such events in Edinburgh (Sept '01), Vienna (Oct '01) and a training course on Neural Networks for Clinicians in Aachen (Oct '01). A Roadmap is being developed for each application sector, including material based on state-of-the-art surveys in the relevant fields. In addition, case studies are being elicited to illustrate relevancy of the techniques, and Task Forces are being considered for areas where there are perceived gaps in important topics. In particular, we hope to encourage younger researchers to interchange their ideas and meet with established researchers in the specialist topics.

Currently, the HMH track has 10 founding Key Nodes, together with many more other Nodes who are interested and active in medical systems. It is intended that this number should grow steadily in coming years, with the proviso that Node members should show genuine participation in at least one of the activities mentioned above.

RESEARCH WORK CONTAINED IN THE PRESENT COLLECTION OF PAPERS

Papers are divided to three categories, which correspond, not primarily to their area of specialization, but to their presentation sequence, according to the sessions organized during EUNITE-2002 Symposium.

Papers [5]-[10] are closely related to intelligent e-health applications in medicine, announced as the special topic of this year's activities. The work in [5] describes a system that blends the ANNs' advantages for non-linear information processing with the capabilities of DSS to provide help on specific problems that arise in medical practice. The application domain has to do with the prediction of risk of intoxication in patients treated with digoxin, a commonly used drug, for the treatment of cardiac diseases. In [6], a web-based effective management of diabetes mellitus is described. The diabetes management system was developed in order to provide instant access to an electronic diabetes database and real-time information for health care professionals, in order to facilitate their decision-making. The work described in [7], discusses the possibility to utilize a multiagent system as a platform for development of smart adaptive systems, through a knowledge-based model of the agents' mutual awareness (social knowledge). In [8], a remote multimedia -based approach is proposed, involving multicriteria decision-making, classical networking considerations and user preferences, for back pain treatment. In [9], the just in time concept is introduced, as a procurement technique for hospital pharmacy. The study demonstrates the use of a software model for healthcare logistics, based on the intelligent agent technology and the internet, technology. Finally in [10] the development of an intelligent agent is described, that is based on causal probabilistic networks and interprets blood glucose monitoring data through telemedicine technology, for generating alarms when some deviations in the patient state are detected.

Papers [11]-[16] focus on the application of adaptive and/or hybrid computational intelligence schemes in medicine. In [11], classification of damages existing in the central nervous system (CNS) of newborn infants is attempted, by performing cry analysis with the aid of ANNs. The work described in [12], integrates a verification tool, which can

---

[1] See details in http://www.akh-wien.ac.at/mes/home_en.html

[2] Same as before, see http://www.akh-wien.ac.at/mes/home_en.html

[3] See also details about this summer school in : http://fuzzy.iau.dtu.dk/summer.html

classify kinds of solicited dental surfaces into a specific therapy instrument called the ultrasonic scaler. In fact, the problem corresponds to the application of a feature extraction algorithm, to the data of a vibrational signal. Next, [13] presents an attempt to automatically define feedforward ANNs using genetic programming. The proposed system is tested on two real world data sets related to thyroid and Pima Indians diabetes. The work in [14] discusses the fuzzy-controlled glaucoma monitoring based on arden syntax. The approach is assisted by a knowledge base, containing the outcomes of an ANN and a set of fuzzy control rules. In [15], an application of learning vector quantization is presented, for detecting drivers dozing-off, by analysing EEG and EOG recorder signals during an overnight simulation task. In [16], gene expression data based cancer status prediction is attempted for individual patients, with the aid of an optimised fuzzy cluster and rule-based system.

Finally, papers [17]-[24] consist other research attempts related to computational intelligence, applied in medical domains and thus, mainly consist intelligent computer applications to specific medical decision making problems. In [17], it is showed that trapezoid functions with larger core regions are the more appropriate functions for calculating the membership degrees within neuro-fuzzy systems. Preliminary benchmark examples are given for the analysis of septic shock patient data. Next, [18] attempts a classification of eyetracking signals with vector-based ANNs. In [19], a method for enhanced evaluation of classified input vectors is proposed, applied to motoric ability tests for alcohol detection. In [20], an ANNs based multisensor fusion is presented, applied in remote sensing in ecology. The work in [21], deals with the design and implementation of an intelligent system for the estimation of the human supervisory operator stress level, under the normal and safety critical circumstances. Then, [22] presents a universal software tool based on computational intelligence, for biotechnological process monitoring. In [23], ANNs are used for solving different medical problems such as handling patients with potential risk of post-chemotherapy emesis and potentially intoxicated patients, treated with digoxin. Finally, [24] proposes adaptive segmentation and classification of non-stationary and time-varying signals. The work illustrates the method on classification of events on EEG data.

EVALUATION OF PAPERS AND FUTURE PROSPECTS

The 20 papers included within this volume are to be presented in the EUNITE-2002, the European Symposium on Intelligent Technologies, Hybrid Systems and their Implementation on Smart Adaptive Systems[4], 19 - 21 September 2002 Albufeira, Portugal. The papers will be subjected to an evaluation procedure, aiming finally to judge if they could appear, after revision and improvement of their content and meaning, in high quality journals of related scientific fields. First phase evaluation will be performed by the audience of the EUNITE-2002 Symposium following a secret voting process. Generally, evaluation will be based on the relevance of each paper to intelligent e-health aspects, or, to adaptive / hybrid systems, to the originality of the work presented, to likely exploitation in the future and to the quality of oral / poster presentation and /or written text. Best papers will be announced at the end of EUNITE-2002 Symposium. The aim far ahead is to produce high quality journal's special issues on subjects closely related either to intelligent e-health topics, or to smart adaptive systems and hybrid computational intelligence in medicine. All the contributors of this edition are greatly acknowledged.

ACKNOWLEDGEMENT

Thanks are due to the University of the Aegean, Departments of
(1) Business Administration[5] and
(2) Financial and Management Engineering[6],
for providing considerable assistance in bringing this collection of papers at the 2nd EUNITE Annual Symposium at Albufeira, Portugal, to successful publication.

REFERENCES

[1] The European Network on Intelligent Technologies for Smart Adaptive Systems: EUNITE – Symposium, 16-17 March 2001, ELITE Foundation, Aachen, Germany.

---

[4] See for details. http://www.eunite.org/eunite

[5] See www.ba.aegean.gr

[6] See www.fme.aegean.gr

[2] Klaus-Peter Adlassnig (Ed.), (2000), Fuzzy diagnostic and therapeutic decision support, Proceedings of the EUNITE-Workshop, Austria, May 11-12, 2000.

[3] Klaus-Peter Adlassnig (Ed.), (2001), Intelligent Systems in Patient Care, Proceedings of the EUNITE-Workshop, Austria, October 5, 2001, ISBN 3-85403-154-8.

[4] Georgios D. Dounias and Derek A. Linkens (Eds.), (2001), Adaptive Systems and Hybrid Computational Intelligence in Medicine, Special Session Proceedings of the EUNITE 2001 Symposium, Tenerife, Spain, December 13-14, 2001, A Publication of the University of the Aegean, ISBN 960-7475-19-4

[5] Antonio José Serrano, Emilio Soria, Gustavo Camps, Jose D. Martin, Jose R. Sepulveda, Rafael Magdalena, N. Victor Jimenez, Web-based Clinical Decision Support System Using Neural Networks (this edition)

[6] Abdul Roudsari, Nengquan Zhao, Ewart Carson, A Web-Based Diabetes Management System: "DiabNet". (this edition)

[7] Lenka Lhotska, Olga Stepankova, Agent Architecture for Smart Adaptive Systems (this edition)

[8] George Ghinea, George D. Magoulas, Andrew O. Frank, Intelligent Multimedia Transmision for Back Pain Treatment (this edition)

[9] Konstantinos. K. Danas; Abdul V. Roudsari, Panayiotis H. Ketikidis, The VPIS System: A New Approach to Healthcare logistics (this edition)

[10] Elena Hernando, Enrique J. Gómez, Pedro Alvarez, Francisco del Pozo, Intelligent Alarms Integrated in a Multi-Agent Architecture for Diabetes Management (this edition)

[11] Mannes Poel, Taco Ekkel, CNS Damage Classification in Newborn Infants by Neural Network Based Cry Analysis (this edition)

[12] Detlef Mitzschke, Jens Strackeljan, Fiamma Ferrero, Bernd Oehme, User-Dependent Adaptability for Implementation of an Intelligent Dental Device (this edition)

[13] Athanasios D. Tsakonas, Georgios Dounias, A Scheme for the Evolution of Feedforward Neural Networks Using BNF-Grammar Driven Genetic Programming (this edition)

[14] Sven Tiffe, Fuzzy-Controlled Glaucoma Monitoring Based on Arden Syntax (this edition)

[15] David Sommer, Thilo Hink, Martin Golz, Application of Learning Vector Quantization to Detect Drivers Dozing-Off (this edition)

[16] Michael Pfaff, Reinhard Guthke, Dirk Woetzel, Dominik Driesch, Gene Expression Data Based Cancer Status Prediction for Individual Patients by an Optimized Fuzzy Cluster and Rule Based System (this edition)

[17] Jürgen Paetz, A Note on Core Regions of Membership Functions (this edition)

[18] Martin Golz, David Sommer, Classification of Eyetracking Signals with Vector-Based Neural Networks (this edition)

[19] David Sommer, Martin Golz, Multiple Training of Vector-Based Neural Networks to Detect Density Centers in Input Space (this edition)

[20] Dimitrios Moshou, Cedric Bravo, Stijn Wahlen, Dimitri Lemaire, Henri Maraite, Hermann Ramon, Neural Network Based Multisensor Fusion Applied in Remote Sensing in Ecology (this edition)

[21] Ladislava Janku, Lenka Lhotska, Vaclav Kremen, Additional Method for the Evaluation of the Training Capturing: Intelligent System for the Estimation of the Human Supervisory Operator Stress Level under the Normal Safety Critical Circumstances (this edition)

[22] Klaus-Dietrich Kramer, Steffen Patzwahl, Thomas Nacke, Dieter Frense, Computational Intelligence - Design Tool for Use in Biotechnological Process Monitoring (this edition)

[23] José David Martin, Emilio Soria, Gustavo Camps, Antonio J. Serrano, Jose R. Sepulveda, Victor Jimenez, Solving Clinical Problems With Artificial Neural Networks: Some Case Studies (this edition)

[24] Lenka Lhotska, Milos Zvolanek, Adaptive Segmentation and Classification of Non-Stationary and Time-Varying Signals (this edition)

Web-based clinical decision support system using neural networks

Antonio J Serrano*, Emilio Soria*, Gustavo Camps*, José D. Martín*, José
R. Sepúlveda*, Rafael Magdalena* and N. Víctor Jiménez†
*Digital Signal Processing Group, Electronics Engineering Department,
Faculty of Physics, University of Valencia
50 Dr. Moliner Ave. E46100 Burjassot SPAIN
Phone: +34 96 3160 198, Fax: +34 96 3160 466
email: nn@gpds.uv.es
†Pharmacy Service, University Hospital Dr. Peset, Valencia, SPAIN
email: jimenez_vic@gva.es

ABSTRACT: This paper describes a system that blends the artificial neural networks (ANNs) advantages for non-lineal information processing with the capabilities of decision support systems to provide help on specific problems that arise in medical practice. System has been devised to bring on users capabilities to easily create, develop and query to neural models with no previous knowledge about the algorithms. This performance have been achieved by automating the training process of the ANNs. In order to provide a real-world effectiveness, it has been developed as a network-based application, concept that comprises in a single word a web-based browser runnable (WB^{2}R) application. Finally, a real application example is shown: prediction of risk of intoxication in patients treated with digoxin, a commonly used drug for the treatment of cardiac diseases.

KEYWORDS: decision support systems, artificial neural networks, web-application, e-health

INTRODUCTION

Nowadays, process of economic and cultural globalization is compelling to several European governments to increase the efficiency and quality of the services offered to citizens. This circumstance is forcing to the Health Administrations to invest a considerable amount of efforts and money in resources targeted to increase the Quality of Life (QoL) of patients. The most of the efforts are driven to develop modern decision support systems (DSS). Its aim is to improve the medical task results, either by detecting anomalous situations or suggesting actions, but never replacing the human physicians.

Clinical decision is a choice among several possible options, being last and only goal the patient's health. This task is, generally speaking, very difficult, due to the amount of different criteria that may apply from the fact analysis. Clinicians consider the information coming from a body exploration, clinical history and laboratory reports, before expressing a possible diagnosis. Final decision is often entangled by uncertainty due to lack or imprecision of the available data. In these cases, decision support systems can provide a second opinion to increase the conviction in the decision to take, if matches with the clinician decision.

The decision problem analysis could not be linear, as usually supposed in several clinical cases, and data can hide irregularities, incongruities or mistakes. ANNs are one of the most important tools in analysis, modeling, regression and prediction now available [1]. Its use in clinical environments is a well proved fact with excellent results in plenty of cases ([2], [3]). Differently from other approximations used in development of DSS, ANNs do not rely on any hypothesis about data distribution, and do not need access to the expert knowledge coded in rules: they learn relationships between the available patient information and the disease from the real data. This advantage is taken to develop a new decision support system based in ANNs.

The DSS must provide several capabilities to be a useful tool, apart from being technically powerful to bring an acceptable performance. From the user point of view, DSS must offer ease of use and must be available from nearly everywhere. The implementation of the DSS as a web application is truly advantageous over the classic computer applications. We can remark the easy and uniform access to the application by using a common web browser, the information centralization and management, which will allow to detect and prevent incongruities, duplications or information leakage, and the ease of maintenance, due to the existence of an only *application*, common to every user, which guarantees the user access to the last and most updated version.

Intelligent e-Health Applications in Medicine
Special Session Proc. EUNITE 2002, Albufeira, Portugal, Sept 19-21, 2002
Published by the University of the Aegean

SYSTEM STRUCTURE

Application structure is an inter-connected four elements based architecture (Figure 1): the Web Browser, the Web Server, the Data Manager and the Decision Support Module. User access the application with the Web Browser, connecting to a computer that acts as a server; the Web Server sends the content and the data according to the request or the allowed permissions. In our developed system, Web Server acts as coordinator and interacts with the Data Manager and the Decision Module to solve the requests from the Web Browser. The Data Manager contains a database which structure allows the storage and retrieve of all information needed by the application (problems, variables and data, performed tests, users, etc.) and the Decision Support Module is made up of a bunch of processes with the only purpose of processing the whole information arrived form the different users and problems, and reporting the Web Server every aspect related to the decision or the problem solution state.

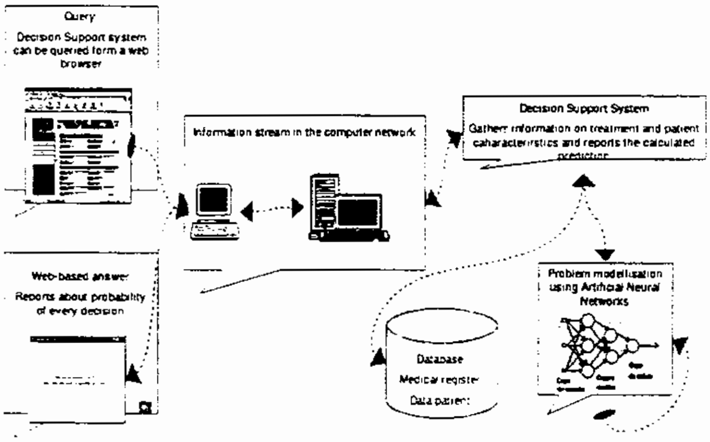


Figure 1. Block diagram of the system, showing the relationship among the several components that comprise the Decision Support System based on Artificial Neural Networks

Hyper-Text Mark-up Language (HTML) [4] has been the chosen programming language to create web pages of the application, and functionality at the client-side has been achieved by using JavaScript [5]. Dynamic pages in the server have been developed with PHP language [6]. Web application is running in an Athlon 800 MHz processor PC, with 256 MB of RAM memory, Linux operating system [7] and Apache web server [8]. The access to the mySQL database [9] is programming-language driven.

THE DECISION MODULE.

The decision module uses neural models to provide a specific response to the clinician's inquiry. Automatic training of neural networks is delegated to a Process Server which works in parallel with the Web Server[1]. The Web Server centralizes the rest of necessary processes for the study and treatment of the available information. They are *programs* that are loaded and unloaded in the server when needed. Four basic *functions*, namely *procedures*, for every problem are built: *migrate*, *analyze*, *modeling*, *inference* (Figure 2)

---

[1] At present it is only available the classical multilayer perceptron trained with the *backpropagation* learning algorithm [3]. Networks were trained with cross-validation as a stop criterion, i.e. part of the data are excluded from training and are used to evaluate the generality of predictions.

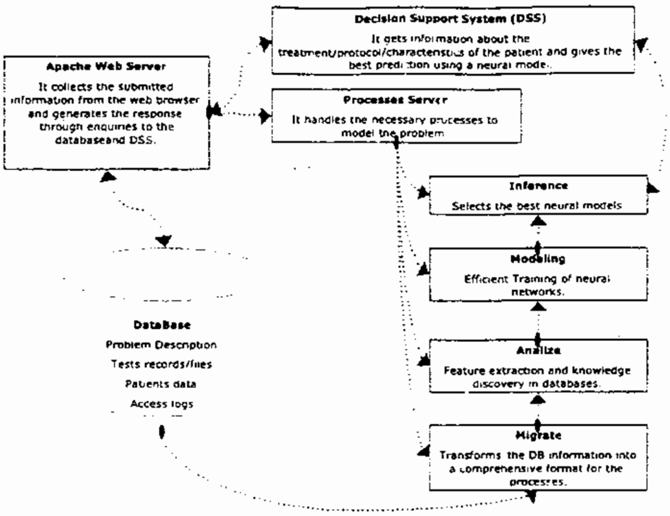


Figure 2. Scheme of the DSS.

USER INTERFACE

Neural-based DSS has been designed to assist clinicians in decision-making. This fundamental idea leads to a simple, user-friendly and intuitive interface since expertising on neural networks or machine learning is not required (Figure 3). In this way, user get the best of both worlds: cutting edge neural computing and the convenience of a web-browser interface. Moreover, the interface can adapt itself in appearance and options to the user, according to his/her knowledge of the specific problem. In addition, data confidentiality is constantly protected while in transit between sites through the "SSL" protocol. The components are (a) site certification and (b) transfer encryption .

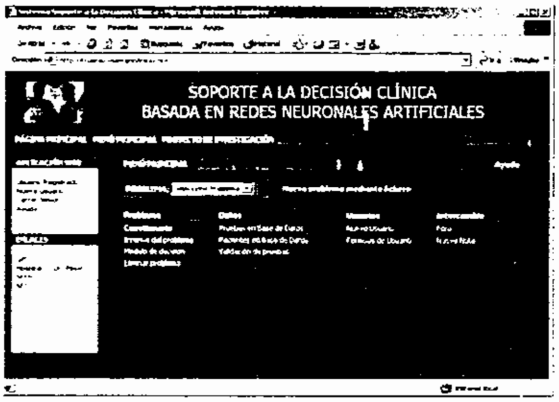


Figure 3. Appearance of the main screen of the neural-based clinical DSS.

MAIN MENU

The application is ready to deal with multiple authentified users with several problems in hand at the same time. Every user can personalize his/her environment depending on the allowed tasks established by the master, that can be defined as follows:

- **Problem Management Menu**: Options have been reduced, basically, to inquiry the decision and to retrieve reports regarding the efficiency, state and specific features of the model.
- **Data Menu**. This menu allows database access and retrieving.

- **Users Menu**: Every problem in the Server allows more than a single user. In this menu, we find options for users handling, licenses, security issues, etc.
- **Exchanging Info Menu.** This menu provides ()with additional characteristics to the application such as specific newsgroups and talk applets. This is of relevant importance in cases when different clinicians are treating the same or similar problems.

Figure 4 illustrates the main menu of the application.

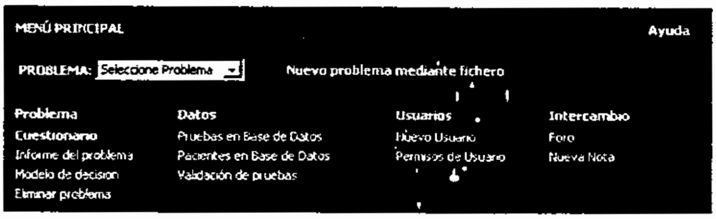
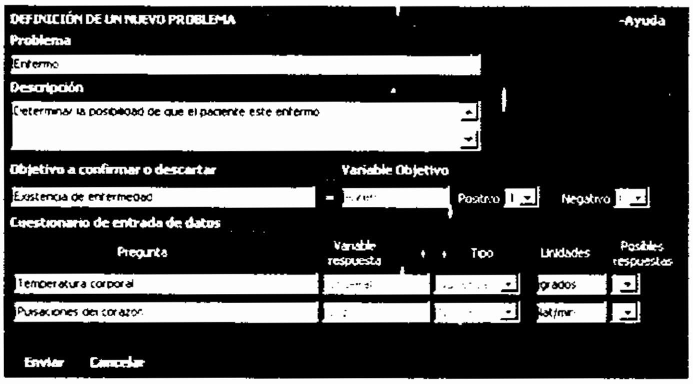


Figure 4. Main menu of the application.

DEFINITION OF A NEW PROBLEM

The final objective of this application is to provide a second opinion in complex clinical problems, i.e. the clinical symptoms or characteristics of a patient must be transformed in a diagnose of disease. Nevertheless, this must not disturb the reader since models do not provide a diagnosis but rather an estimation, leaving the final decision to the clinicians. The problem is defined through a simple form to fill in (Figure 5). The user fills in the goal (problem), a short description of the problem in hand, and which are the dependent and independent variables.



Problema: Enfermo

Descripción: Determinar la posibilidad de que el paciente este enfermo

Objetivo a confirmar o descartar: Existencia de enfermedad = [illegible]

Variable Objetivo: Positivo 1 Negativo [illegible]

Cuestionario de entrada de datos

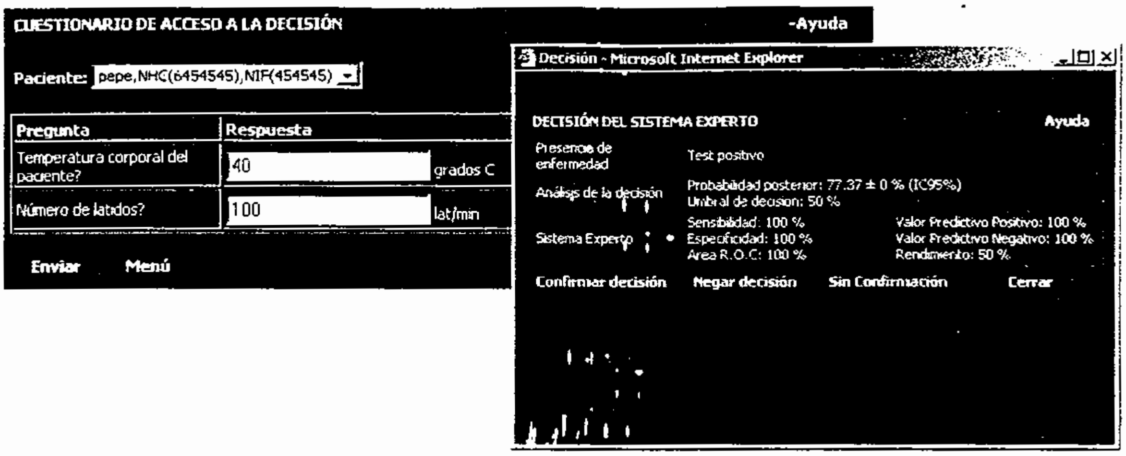
Pregunta	Variable respuesta	Tipo	Unidades	Posibles respuestas
Temperatura corporal	[illegible]	[illegible]	grados	
Pulsaciones del corazon	[illegible]	[illegible]	lat/min	

Figure 5. Training data input form for a especific problem. The actual scale for either input or output variables is not relevant since all data is automatically normalized to have zero mean and unit standard deviation prior processing.

Once the problem is defined, the system prepares itself to collect input data, building an internal data structure to handle efficiently relevant information. This implies an additional "intelligent" module to analyze each especific problem. New data, e.g. new patients, do not necessarily need the system to be re-trained since this only depends on the relevance of the new data entered. Data whitening, Principal Component Analysis (PCA), clustering and other feature extraction functions are used to discern this issue. This procedures are transparent to the user.

USING THE DSS WITH NEW DATA.

Once a problem is defined, the data have been entered and the model trained, the DSS can give solutions for new input data through an auto-generated form. Questions are sent and subsequently evaluated by the system which responses with probability scores (success rate, sensitivity, specificity, ROC area, Positive Predictive Value, Negative Predictive Value). Figure 6 shows the process of testing the model with new data.



CUESTIONARIO DE ACCESO A LA DECISIÓN -Ayuda

Paciente: pepe,NHC(6454545),NIF(454545)

Pregunta	**Respuesta**	
Temperatura corporal del paciente?	40	grados C
Número de latidos?	100	lat/min

DECISIÓN DEL SISTEMA EXPERTO Ayuda

Presencia de enfermedad: Test positivo

Análisis de la decisión: Probabilidad posterior: 77.37 ± 0 % (IC95%)
Umbral de decision: 50 %

Sistema Experto:
Sensibilidad: 100 %
Especificidad: 100 %
Area R.O.C: 100 %
Valor Predictivo Positivo: 100 %
Valor Predictivo Negativo: 100 %
Rendimiento: 50 %

Figure 6. Testing the model with new data for a specific problem. The actual scale for either input or output variables is not relevant since all data is automatically normalized to have zero mean and unit standard deviation prior processing. The system gives a diagnose of disease, a posteriori probability and confidence scores on the model.

Results depend, obviously, with the number of real data examples learned by the system. The higher the number of patterns is, the more representative is the solution provided and, consequently, more confidence can be kept in new cases (generalization capability). In this sense, the system offers the possibility to include each new prediction example in the database in order to re-train the model.

RESULTS

The application followed a preliminary testing period, including different lines: testing the learning *core* of the application with clinical and pharmaceutical problems, environment refinement, computational burden reduction through MEX-files compilation of m-files and security testing. After this initial period with clinician's supervision, the system was ready to be tested. The WB^{2}R DSS has been tested in a great amount of synthetic and real problems. Since the final use of the application involves clinical problems, testing has been carried out in collaboration with two hospitals in the city of Valencia (SPAIN).

A case study is described in this paper. We tested the application in a common problem in cardiac patients; risk stratification of intoxication in patients treated with digoxin (DGX). Digoxin is a drug widely used for treating congestive cardiac failure and symptomatic alterations of the heart rate such as auricular fibrillation and paroxysmal supraventricular tachycardia. The main drawback derived from using this drug refers to the possibility of producing intoxication in the patient due to its narrow therapeutical range of application, whose value is usually accepted between 0.8 and 2 ng/ml [13]. Patients monitored in the *Servicio de Farmacia of the Hospital Universitario Dr. Peset* in the city of Valencia (Spain) were included in this study. Training set was formed by 172 patients and 85 more patients constituted the validation set. Patients are then classified into two categories: patients with plasmatic concentrations (PC) greater than 2 ng/ml (high risk) and patients with PC < 2 ng/ml (low risk). Results were excellent (SR = 90.6%, SE = 80%, SP = 93% in the validation set) and revealed the automatic approach proposed as very useful in the clinical management of patients.

CONCLUSIONS.

In this paper we have presented a novel DSS to help in clinical decision making. The system is web-based, user friendly and intuitive and incorporates neural models as decision models. Neural networks have proven to be well suited in clinical problems not only because of the accuracy of their estimations but also because of their robustness. Our future work is tied to the use of more sophisticated neural networks an kernel-based methods. We are at present working on rules and knowledge extraction from trained networks and expect that an understanding of accurate models will provide us with information about which variables are important and which meaningless, and about treatments and suitable protocols, etc.

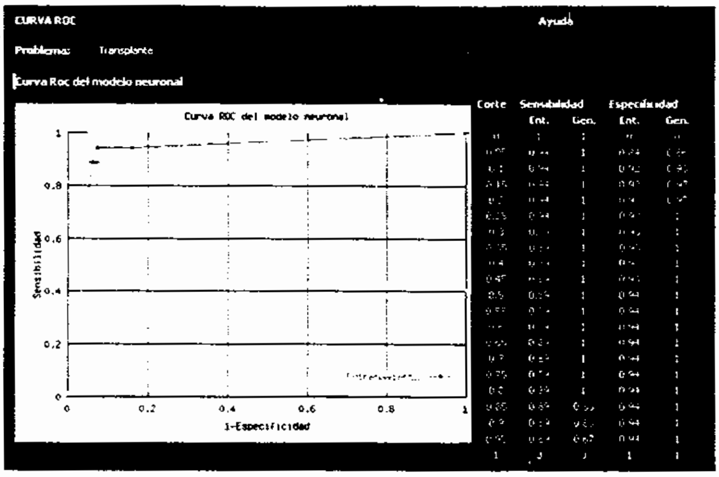


Figure 7. Receiving Operative Curve (ROC) of an specific model. Results are shown for training and validation. The decision limit γ is varied along the output range [-1, +1] to obtain the curve. In the table aside on can select the optimal decision limit, inspecting the sensitivity and specificity scores achieved.

REFERENCES

[1] Turban, E. Decision Support Systems and Expert Systems. Prentice Hall. 1995.
[2] Weinstein, M.C., Finenberg, H.V. Clinical Decision Analysis. W.B. Saunders Company. 1980.
[3] Haykin, S., Neural Networks: A Comprehensive Foundation. NY: Macmillan. 1994.
[4] Dybowski, R., Gant,V. Clinical applications of artificial neural networks. Ed. Cambridge University Press. UK. 2001.
[5] HyperText Markup Language. http://www.w3.org/MarkUp
[6] JavaScript Developer Central. http://developer.netscape.com/tecn/javascript/index.html
[7] PHP. Scripting language for Web development. http://www.php.net
[8] Elboth, D. The Linux Book. Prentice Hall PTR, 2001.
[9] Apache HTTP Server Project. http://httpd.apache.org
[10] MySQL Open Source Database. http://www.mysql.com
[11] Hunt, B.R., Lipsman, R.L., Rosenberg, J. A Guide to Matlab : For Beginners and Experienced Users Ed. Cambridge University Press. 200.
[12] Schurman, S.J., McEnery, P.T. Factors influencing short-term and long-term pediatric renal transplant survival. Journal of Pediatrics. 1997;130:455-62.
[13] Evans et. Al. Applied Pharmacokinetics. Principles of Therapeutic Drug Monitoring. Applied Therapeutics, 1993.

A Web-based Diabetes Management System: Diabnet

Abdul Roudsari, Nengquan Zhao and Ewart Carson
Centre for Measurement and Information in Medicine
City University
Northampton Square
London EC1V 0HB
UK

a.v.roudsari@city.ac.uk
e.r.carson@city.ac.uk

Abstract

Effective management of diabetes mellitus requires the processing and interpretation of blood glucose, insulin and other lifestyle data. This paper presents a web-based diabetes management system (DiabNet) which was developed to provide instant access to the electronic diabetes database and real-time information for health care professionals in order to facilitate their decision-making. DiabNet offers innovative online diabetes management involving online appointment and consultation. It integrates portable glucometer, handheld computer, mobile phone and Internet access as a combined telecommunication and mobile computing solution for diabetes management. Active Server Pages (ASP) embedded with advanced ActiveX controls and VBScript were programmed to allow remote data upload, retrieval and interpretation. Internet-based learning features, together with a video teleconferencing component make the DiabNet web site an informative platform for tele-consultation. Security steps have also been taken to protect sensitive patient information and the authentication of registered user.

Introduction

Consultations which outpatient diabetic patients have with specialised health care services account for a tremendous amount of time and money. Patients sometimes find they have to wait several months for their appointment, although their problem often turns out to be a minor clinical problem that could have been dealt with in general practice. Many seriously ill patients have to endure long waiting periods for their appointment, during which time their illness will worsen and the chances of recovery diminish.

Telemedicine offers efficient means for overcoming some limitations in the healthcare system. By using telemedicine, health care institutions can increase their patient population throughput, thereby achieving shorter stays and increased outpatient care, increased referrals, and improved competitive differentiation. In the UK, the National
Health Service (NHS), for example, is currently using Racal's Healthlink nation-wide data communications network and messaging service to connect GP practices to their FHSAs (Family Health Service Authorities) for its high-profile "Links" project. GPs can already send patient registration data and Items Of Service to and from their FHSAs this way.

This paper describes a web-based system called DiabNet, that integrates handheld computer, mobile phone and Internet access into a portable tele-monitoring and diabetes management device. The system is completely hardware/software-independent and runs independently from the patient-location. For the end-user it offers a system which is convenient for storing and retrieving information in and from the recommended portable devices. They are able to obtain

Intelligent e-Health Applications in Medicine
Special Session Proc. EUNITE 2002 Albufeira, Portugal Sept 19-21, 2002
Published by the University of the Aegean

substantial information generated by the web service. In addition, a teleconferencing session can be established on demand. All these components make the system a virtual clinic in close proximity to the patient and a portable tool to facilitate decision making for health care professionals.

In the sections which follow, the architecture and functions of the system are described and its use in routine diabetes care is illustrated.

Telecare in Diabetes Management
Diabetes affects over 100 million people world-wide and is likely to increase to affect over 230 million people by the year 2010 and 260 million by 2025, according to the World Health Organisation (WHO) [1]. Without proper management of the disease, diabetes can lead to severe long-term health complications. These include blindness, kidney disease, heart disease, and peripheral nerve damage, potentially leading to amputation. Diabetes is a major cause of death by disease in the UK, principally as a result of its secondary complications, and affects up to 2.4% of the adult population. The disease and its complications often cause severe problems for affected individuals and their carers and impose heavy burdens on health services, consuming at least 5% of national health care resources [2,3].

Self-management strategies involve different components of the health care and lifestyle regimen: blood glucose self-monitoring, dietary behaviour, exercise, medication taking, safety/preventive actions and appropriate integration and timing of all these activities. Frequent measurements of blood glucose are important to determine the effectiveness of therapeutic interventions.

The following is a list of widely used glucometers currently available on the market: Refloflux S (Boehringer Mannheim, Germany), One Touch II (LifeScan, USA), Glucocard Memory (Menarini, Italy), Precision QID (Medisence, USA), HaemoCue (HaemoCue, Sweden), Accutrend DM (Boehringer Mannheim, Germany) and
Samsung Fine Chemicals Co., Ltd. (SFC). SFC announced that it had developed the world's first non-invasive blood glucose meter, a hand-held measuring device named TouchTrakO, which measures blood glucose levels without a blood sample. By using a non-invasive electromagnetic radiant ray to determine glucose levels, Samsung's new device eliminates the need for diabetic patients to prick their fingers several times every day.

Another important event in 1998 occurred in December, with the announcement of preliminary results for the GlucoWatch™ automatic and continuous glucose-monitoring device from Cygnus company [4]. Several testing objectives were satisfied, including the establishment of internal accuracy target values to be used later in protocols for clinical studies. The GlucoWatch™ performs frequent measurements that can be used by diabetic patients to control blood glucose levels more effectively. Reverse iontophoresis provides non-invasive extraction, and data can be collected and stored at the rate of three measurements an hour for up to 12 hours. Other features will include an alert system, longer data storage times, and the capability to download recorded information to a personal computer.

With the use of these non-invasive blood glucose meters such as the GlucoWatch™ and its spin-off products, diabetic patients should be willing to measure their blood glucose more times each day, providing more data for accurate aggregation and interpretation. Moreover, these light and handy devices will become essential components of the telemonitoring system for diabetes care.

These measurements are usually delivered to the diabetologist during a conventional consultation session, and are used as displayed in Figure 1. This consultation model requires the patient to visit the Diabetic Day Centre (clinic) every several months based on the status of the disease. However, such a consultation is subject to postponement or cancellation due to the patient's uncertain location, i.e. when travelling or due to illness.

Figure 1: A conventional consultation session

A number of major clinical trials undertaken during the 1990s showed that patients who maintained lower levels of blood glucose through intensive management (frequent self-monitoring) had significantly lower rates of these complications [5,6].

A variety of diabetes management programmes (i.e. Camit, Diabass, Mellitus Manager, etc.) have been developed to improve compliance among patients with diabetic conditions. Some services are solely educational, while others include an element of intervention to encourage medication adherence and behavioural change. In France using Minitel, some diabetic patients have had access to email and an expert system customised to give individualised dietary advice based upon information supplied by the patient regarding diet and exercise [7]. Patients also had access to libraries of information, for example recipes. In a randomised trial, patients with access to information through the system demonstrated improved diabetic knowledge, improved dietary habits and improved metabolic markers of disease control (fructosamine and HbA1c).

Many of these programmes have demonstrated improvements in patient quality of life and reductions in total treatment costs. The greatest success has been achieved with programmes that deliver a personalised service and provide long-term follow-up. However, these programmes are the most expensive to implement.

With the fast growing utilisation of information and communications technologies (ICT) in diabetes care, self-monitoring can be even more effective in terms of using telemedicine by avoiding unnecessary visits to the clinic and increasing the possibility of immediate consultation initiated by the patient. There are also some successful cases that demonstrate that the use of a handheld computer can enhance patient care. Palm-based mobile devices provide the flexibility and performance physicians need at any time, in any patient care location to help ensure accurate and timely provision of the care which is needed [8,9].

Although self-measuring of blood glucose is very important to diabetic patients, they are still waiting for an approach which can mitigate the painstaking measurement process and facilitate direct delivery of the test results.

Using the Internet, an online approach to diabetes management may be proposed to deliver a highly personalised service, over an extended period, in a cost-effective manner. Blood glucose measurements can be immediately transmitted to the remote server for analysis by physicians. Moreover, it is now increasingly unlikely that any one institution has a representative from every feasible medical sub-specialty within their institution, so the sharing highly specialised expertise across different hospitals is also required.

Shared care is a comparatively new model that is often adopted by many different health care professionals involved in the management of an individual patient. For example, the diabetic patient may interact with a primary care physician (or a diabetologist), an ophthalmologist, a diabetic specialist nurse/district nurse, a chiropodist, a dietician and a diabetes educator for instruction in day-to-day care. All this implies not only communications between the patient and these health care professionals, but also the communication between these team members which are needed.

The idea of using ICT in diabetes care as an integrated patient-doctor approach was first addressed by the DIACRONO system, developed by Gomez, DelPozo and colleagues in the late 1980s [10]. The conceptual model of the "telemedicine based diabetic patient management system" was first introduced in 1991 by the same Madrid-based group, who tried to adapt the DIACRONO system to the specific telemedicine approach to diabetic care [11]. The system brought together the data collection and processing capabilities of hand-held devices, namely a Patient Unit (PU) and a diabetes workstation (WS) with therapy advising capacity. Due to the limitations of portable computers and wide area network technologies at that time, the hand-held devices described in the system were merely used as a data "store-and-forward" tool.

Also, the design of the PU-WS model, via the public telephone network, not only limited the location of health care staff (he or she must sit down in front of the WS), but limited the number of patients involved as well (he or she must have a handheld device running specific software).

The subsequent DIABTel project intended to overcome the inherent difficulties and problems of telemedicine systems. The project protocol was still to be based on its preceding concept and design. However, it represented a further practical implementation of the first telemedicine model for diabetes care. With the rapid development of the portable computer, this project employed a compact palmtop computer with decision-aid software to collect and analyse data. It was limited, though, by the resolution and colour of the screen. Separate software design for the Medical Workstation (MW) and the Patient Unit (PU) was emphasised to suit the specific needs of health care staff and the patient individually [11,12]. In the mean time, the absence of the integrated patient administration module and hardware-specific programming in this project would deter the wide use of this service.

The need to manage diabetes care information and develop quality control European-wide was identified in the DIABCARE Q-Net project [13]. This European project aimed at improving the quality of health care for diabetic patients. Patient data have been collected in all the countries of Europe, by using the so-called Basic Information Sheet (BIS), Data Entry and Feedback Software (stand alone PC software), Fax System and Internet component. Data exchange is supported by paper, fax (OCR), modem and Internet. However, transfer via the Internet by evaluation of HTML-forms and linkage to the DIABCARE database has not been fully developed.

This project also faced many challenges such as (i) the need for multi-language support, (ii) security problems due to the current use of paper and fax-based data sheets and (iii) management and control issues among various data centres located in different countries. Besides, the standalone software which has been introduced highlighted the need for additional training and technical support for end users (i.e. diabetes specialists, GPs). Also, the use of PCs with modem and Internet connectivity, based on the local conditions in each country, does not yet allow the international benchmarking of the quality of care across all centres in Europe.

In 1997, another European project was launched to develop a telematic service to assist IDDM patients. The T-IDDM project (Telematic Management of Insulin Dependent Diabetes Mellitus) inherited from the technical architecture of the DIABTel project the MW and PU concept, but used a Web based architecture for patient-doctor communication [14]. In the implementation of the T-IDDM architecture, the PU is connected to the hospital through standard Public Switched Telephone Lines, while the distributed services of the MW are accessible over the Local Area Network (LAN, and specifically to the Intranet) of the hospital. In particular, the project introduced a powerful Oracle database, hierarchical security measurement, and various organisational aspects in its development. In those PUs with slow CPU and low memory, unpredictable delay in producing the graphics is inevitable, which will distract the end-user from continuous use.

In terms of hardware, the glucose testing device is undergoing rapid improvement from finger pricker plus glucometer to pain-free non-invasive blood glucose meter, which may give impetus to the telemedicine approach in diabetes care. Also, the portable personal computer has decreased much its size and weight in the last 20 years, though at the same time it has become more and more powerful and faster. Although some projects are currently undergoing feasibility testing on a small scale, there is no evidence that this new telemedicine service has yet taken the place of the traditional consultation between the patient and the doctor. Diabetes management applications on the Internet have not been fully developed and the advantages of using the web could be utilised to a greater extent.

DiabNet

The new DiabNet concept of mobile diabetes management has been introduced as a result of the rapid development of personal telecommunication technology and involved the following three steps. First, terminal mobility means that patients always carry a portable telephone which can be used anytime and anywhere. This kind of mobility is already possible with conventional cellular phones and the GSM network. The second step, called personal mobility, is to guarantee the mobility of the telephone number. Personal mobility will release the patient from the cellular phone: smart card storage of their personal number as well as their relevant health care information will be enough to establish a consultation session when required. The third step is to introduce personalised services. In this way one can receive personalised communication services by storing one's own communication mode in the network, that is the network can assign personalised space to store and forward specific multimedia health care information for a specific patient no matter whether they are online or not.

The scheme of the DiabNet system and telemedicine consultation is depicted in Figure 2. The blood glucose measurements are produced by a conventional glucometer or non-invasive GlucoWatch™. The patient can then easily input those measurements into their handheld computer on the DiabNet web site while their mobile phone is connected via an infrared port. A wireless modem PC card such as Aircard™ is another good choice to connect to the Internet. A portable PC card-based camera and ISDN modem could also be an option when teleconferencing is needed.

The combination of the glucometer and the handheld computer is a new idea in the transmission and processing of patient blood glucose data. By combining portable glucometers such as the GlucoWatch™ accompanied with the handheld computer patients can measure their blood glucose level as many times as they think necessary and input those data into the handheld computer for transmission and analysis.

At the GP's office, the desktop computer is connected to the Internet. An ISDN connection is preferred, while a fast analog modem or Ethernet card can also be installed to establish a high performance Local Area Network (LAN). A PCI-based videoconferencing system is necessary to transmit high quality moving images and sound.

Figure 2: Prototype of a telemedicine diabetes care system

A Hospital Web Server, which can be located at the local FHSA or hospital, uses Active Serving Pages (ASP) to generate the dynamic HTML page the user requires. These web pages also possess Microsoft ActiveX controls to produce a run-time diagram of data set analysis. The patient administration database and patient record database are both stored in an NT-based Diabetes Data Server at the same location to provide extra security and data backup technology. Microsoft® Remote Data Service technology is employed to communicate between the user and the remote databases.

During the telemedicine consultation, the diabetologist can receive the patient's measurements immediately from the DiabNet web site while the patient may be a thousand miles away. There is also the possibility for the diabetologist, when remote from their office, to keep close watch on the patient's blood glucose by utilising this system. The diabetologist is able not only to use their own handheld computer, but can use any computer connected to the Internet in a public Internet room or hotel room as well.

Client-Agent-Server Architecture for Remote Diabetes Management

Remote diabetes management uses many mobile technologies and devices. The mobile device will not be connected to the network all the time. For example, a patient may turn off his or her mobile phone for some reason, however, just at the time that their doctor may have requested a download of their personal diabetes database. In a normal client-server application, if the server cannot locate the database via a TCP/IP protocol, the SQL query session will be cancelled. Moreover, the patient will not subsequently know about that request which the doctor had made earlier. To avoid this situation there is the need for a Client-Agent-Server solution such as that which has been adopted in DiabNet.

DiabNet uses a client-agent-server architecture that enables a client (the patient's or doctor's handheld computer) to work offline and connect periodically to the network to submit requests and receive results from the server. The agent component of this architecture functions on behalf of the client in its absence from the network as shown in Figure 3.

Figure 3: The Client-Agent-Server paradigm

By using this technology, if either part of the connection fails to connect the network, DiabNet Agent can store and forward the patient or doctor's request at an appropriate time via a Agent Event Manager, Message Gateway and client side Message Manager.

The conceptual web-based diabetes management model and related products described relate to the development of the integrated diabetes care system - DiabNet. The DiabNet user can manipulate the data even when they are offline. The only calls requiring a trip to the server are calls to update the data server or a request for new data. Multiple languages are supported.

Four different levels of software security are provided before offering the service in a multi-user environment. These are: database security, application security, browser security and system security. Application security is quite different from database security. Application security focuses on securing not only data but also processes. Application security can limit users' ability to utilise some of the functionality provided by ActiveX control, like selecting data entry forms, producing certain graphs or reports, or running critical procedures (such as updating a patient's main measurement of blood glucose (BG) level).

Forcing users to log on to and out of the application also gives the opportunity to create audit logs of all user activity. These audit logs can help to keep track of who is using the application and to locate and fix problems reported by users. Browser security refers to the web browser's (such as Internet Explorer) own security screens. All of the above security strategies need to be combined together in developing the DiabNet project to ensure the maximum protection of patient data.

The sensitive web pages on the DiabNet web site refer to all the pages containing any type of personal information, for example the patient's treatment history, measurements, personal diabetes database (PDD), etc. DiabNet takes a "3As" approach to protect these sensitive web pages - namely, Authentication, Accessibility and Audit.

Accessibility can be defined as the right user seeing the right thing. The appropriate rights should be granted to different users hierarchically, from system administrator, to doctors, to nurses and to registered patients. The system administrator can monitor all the activities on the server and keep the server stable and secure. Doctors can see their list of patients, nurse list and relative patient record. However, nurses can only view the data of certain patients who have been assigned by the doctor. Registered patients can view their own patient records. Those who have not registered at the DiabNet web service can only browse the general information on the web site and, as such, they are called restricted users or visitors.

Audit is an import measurement to secure a web site. It answers: Who was recorded? (User name). When was the event? (Event time). What did he or she do? (Event and files accessed), and Where was he or she? (User's IP address).

All this information is stored in a log file on the server. The system administrator of DiabNet is able to check the log file frequently (using other site management software) to assure the healthy operation of the DiabNet web service.

DiabNet Functionality
The DiabNet web site, as a core component in the telematic diabetes management service, is a dedicated web site aiming to provide a substantial information service and fast connection for both diabetes patients and health care professionals. The main functionality includes:

Data view and interpreter of the Personal Diabetes Database;
Data query service on the basis of an arbitrary period or the blood glucose range;
Data maintenance service, i.e. appending measurements and insulin units, modification of personal information, etc;
Data upload service to transmit the Personal Diabetes Database to the Central Diabetes Database;
Tele-consultation service using telephone or video conferencing software between the doctor and the patient; and
Patient education service providing updated information of telemedicine in diabetes management.

As a web site suitable for both skilled and unskilled Internet user, the navigation amongst these functional pages is designed to be as easy as possible. Moreover, the security measurements are intended to protect sensitive patient information. The web site structure diagram (Figure 4) explains the whole web service and how the user accesses these web pages.

Figure 4: DiabNet web page structure

Once the user is connected to the DiabNet web server, the service requires them to identify whether they are a new user or a registered user. The home page is shown in Figure 5.

Figure 5: The home page of the DiabNet web site

The "Registered User's Page" provides a main menu of the aforementioned services. From each individual service, the user can still click on the "Return to Main Menu" icon to go back to the main menu. Whenever the user has any question about a specific web page, they can click the "help" icon to jump to the relative section of the user guide, so that he or she can get online help immediately.

New users should submit a "New User Registration Form" to use the full service.
The New users register page aims to admit the new user (either patient or doctor) to the DiabNet web service.After the user has sent the form, the information will be stored on the server ready for appending to the central user database. An email letter regarding the ID/Password will be delivered to the user to confirm the registration. Registered users can log on to the services of DiabNet from the "User Login" page (see Figure 6).

Figure 6: Registered user's page

The "Data Viewer and Interpreter" module presents and interprets data sets from a remote database graphically in the common web browser. There are several types of chart to present the same data sets. They are Line Chart, Bar Chart, Scatter or XY Plot, Pie Chart, Area Chart, Hi-Low-Close Chart, Surface Plot, Bubble Chart, CandleStick Chart, Curve Area Chart, etc. All Chart types are available in both 2D and 3D modes. These charts generated using ChartFX software strengthen the statistical functionality and provide impressive visual effects when presenting the data.

A special feature of this module is the integrated toolbar on the left side of the chart window. Users can implement (i) file manipulation (save the current chart and open a saved chart in three different formats), (ii) clipboard manipulation, (iii) print, (iv) data presentation in different type of chart and 3D preview.

Users are allowed to input the particular period, e.g. the interval between two follow-up visits or any other periods he or she is interested in reviewing. A combo box is provided to select among English/French/German/etc., to meet multinational requirements and to select the preferred measure and time units, and define the bounds of the target range.

In the BG by Days of Period Tab, X-Y scatter charts and 3D bar charts visualise the blood glucose and insulin data in a concise but vivid way. The X-axis represents the selected dates from the first to the last day. The X-axis is scrollable which enables the user to view the whole data series by just clicking the scroll bar or dragging the sliding block from side to side; the Y-axis shows blood glucose. The circular dots in four different colours stand for four blood glucose values of Before breakfast, Before lunch, Before supper and Before bedtime. Insulin values are presented by a 3D bar symbol in the bottom of the plot so clinical professionals can compare the insulin intake with the relevant blood glucose level. Glucose levels that are out of target will be displayed in different colours, e.g. hyperglycaemia in red and hypoglycaemia in grey (Figure 7).

Figure 7: BG by Days of Period Tab

When the user double-clicks a particular data point in the chart, DiabNet will pop up a balloon box to tell the user the corresponding date, time and value.

The BG Scatter chart is a Period Overview of Blood glucose. Users can have a general impression as to how the blood glucose fluctuates over the chosen period. This is especially useful when the user wants to know where the peak or valley measurements are (see Figure 8).

Figure 8: BG Scatter Tab

The BG by Days of Period (fitting curve) Tab connects all data points in the same data series, e.g. the Before Breakfast measurements stream. Sufficient information is provided to enable comparison to be made with the values in the same measurement episode.

In the BG by Time of Day chart, the x-axis indicates the four specified sampling times: Before Breakfast, Before Lunch, Before Supper and Before Bedtime; whilst the y-axis is the BG level. The chart gives the range of measurements at certain landmark times (so-called meal-stamped times) (Figure 9).

Figure 9: BG by Time of Day Tab

The BG by Days of Week chart gives range information by comparing measurements on different days of week, so users may pay special attention to their diet or medication on the specific day, usually at the weekend.

The Frequency Distribution chart is a 3D stacked bar chart where two modes are available for the user to choose from. The X-axis stands for BG level, while each colour section on a certain bar indicated one of the four landmark times (BBR, BLU, BSU, BBT). The length of each section is the percentage of the measurements of one landmark time in the total set of measurements (see Figure 10).

Figure 10: Frequency Distribution Tab (in regular stack bar)

Data analysis involves decomposition of observed blood glucose data into trend, cycle and daily components [15,16]. After decomposing a time-series model into its multiple components, the picture can be further clarified with the help of smoothing techniques. These techniques help to give a clearer picture of the behaviour of the time-series data, not distorted by noise or seasonal fluctuations, so that the data can be manipulated to reveal secrets that would otherwise remain hidden.

Centred Moving Averages (CMA) are used to approximate trends (smooth out large variations) in the data, by spanning the period of low frequency variation. When the observations are quarterly (corresponding in our situation to BG level at BBR, BLU, BSU, BBT), the CMA will be formed from overlapping sets of five observations as shown in Figure 11.

Figure 11: Centred Moving Averages

The Logbook is a chronological data viewer in the format of "spreadsheet", which contains the information of Blood glucose measurements, Date, Time, Insulin and other input data in the relevant table of the remote database.

The Statistics tab provides some statistical information relating to the selected period. This can include the number of blood glucose measurements, highest and lowest glucose value and date and time, average blood glucose value, standard deviation (of BG-values) of their population and number and percentage of points below or above target range (see Figure 12).

Figure 12: Statistics Tab

In the Query your Diabetes Logbook page, users are allowed to query the diabetes logbook by (i) an arbitrary period of time or (ii) by a certain blood glucose range. In the case that no records meet the requirements, the user can enter different values and try again. The Maintain your Diabetes Logbook page module is supplied to enter the user's daily measurements, insulin dose and other life events. Alternatively, the user can upload their Personal Diabetes Database (PDD), rather than input each record, into the DiabNet central diabetes database server. The user can upload the Personal Diabetes Database (PDD) to the central database in the Upload your Diabetes Logbook Page, as another way to input their measurements, insulin dose and other life events.

The Teleconsultation Request page can co-ordinate the tele-consultation session between the patient and health care professionals. Ideally, the workflow of this module is as follows. The user selects the proposed quest is a telephone request or a NetMeeting (video-conferencing) request. If this is a telephone request, the user enters the telephone on which he or she can be contacted and when he or she wants to be called. Then the user selects one of the preferred doctors. DiabNet software will check the availability of that doctor using a specified paging system. If the doctor is online, the telephone session will be established. For a NetMeeting request, a NetMeeting® window will pop up to begin a video conferencing session, provided that both parties have video conferencing equipment such as video camera & microphone, etc. (see Figure 13).

Figure 13. Tele-consultation Request Page (NetMeeting session)

The About DiabNet page lists the background and ongoing information about the DiabNet project, to give the user an overview of this project. The Telemedicine in Diabetes page links the user to some other relevant web sites about Telemedicine and Diabetes which have been collected, for instance essential knowledge about diabetes mellitus with animation and other organisation links. It will also be possible to develop an electronic chat board in a future version to allow users to discuss disease information and informative web links in public.

Together these pages comprise the full service of DiabNet. Unlike standalone applications, the web pages can be revised and appended without compiling the whole application. Therefore, DiabNet is fully modularised and easily expandable to include newer services in order to meet the increasing needs of the user. In addition, DiabNet provides an interface for web-based learning and video-conferencing consultation for patients.

Software and Hardware Used

The following software packages were selected to design, code and implement the DiabNet project:

Windows NT 4.0 + Service Pack 3
Visual Basic 6.0
Access 97
Internet Explorer 5.0
FrontPage 98 & Internet Information Service 3.0
NetMeeting 3.0
Chart FX Internet Edition

The promising ActiveX technology has been used in the development ofDiabNet in order to send health care information and data sets between independent web browsers running on a local network or the Internet. This kind of scenario is the focus of the ActiveX for Health Care initiative to date.

The DiabNet server was developed and tested based on the following hardware specification:

Windows NT 4.0 workstation
- Pentium III 500MHz microprocessor
    - 256Mb of SDRAM
    - 12GB hard drive
    - 17 inch ADI Microscan 5GT monitor
    - 8Mb Hercules TNT AGP card
    - 3Com® Etherlink XL PCI 10 MBPS Network Interface card

A notebook computer was tested as a client computer, with the following specification:

AST® Asentia M5160T mobile computer
- 166 MHz Intel Pentium MMX microprocessor
    - 32MB RAM
    - 1.6GB HD
    - 12.1" SVGA TFT colour display
    - Windows 98
    - Data/Fax Modem 33.6 kbps
    - 10Mbps Network Interface card

The video conferencing hardware kit included:

Videum Conference Pro (PCI) - Desktop Video Conferencing Solution
- Videum AV PCI audio and video capture board
    - Colour video camera
    - 640x480 pixels @ 30fps video capture
    - Video compression (up to 48:1) for saving to disk drive
    - Works with all videoconferencing protocols (H.320, H.323, and H.324)

VideumCam Traveler--Digital Video Camera for Notebook PCs

- VideumCam digital video camera
    - VideumCam PC Card (PCMCIA) interface
    - 320x240 pixels @ 30fps capture
    - low power consumption for notebook use

Mobile phone serving as a link between notebook computer and the Internet:

- Ericsson SH1888 mobile phone with:
    - Dual band GSM system
    - Infrared communication
    - Built-in PC card

This hardware setting can satisfy the basic development needs and reduce the network traffic to an acceptable level during small-scale evaluation.

Evaluation of the DiabNet project

The word evaluation has several definitions and covers a large number of applications. The methodology used to evaluate the DiabNet project here is based on the first stage of the evaluation framework (Evaluability Assessment) developed by Bashshur [17], which has been adopted by Cramp and Carson, in evaluating the novel, complex telemedicine system HOMER-D [18,19]. This approach has the potential to be applied usefully to other telemonitoring systems for the management of chronic disease such as diabetes.

The evaluation criteria comprise (i) technical features that affect system use, (ii) feasibility of the service, (iii) user acceptance, and (iv) impact on diabetes management. To implement the evaluation of DiabNet project using the above approach, a questionnaire survey was designed to identify, collect, analyse, and interpret data so as to form a coherent picture of the process that resulted in the effects and impact observed.

A need has been identified to collect quantified feedback data and subjective comments on the technical improvement, efficiency, security and human friendliness issues from end users. Ideally, the survey will result in a report that clarifies user attitudes to a web-based approach in diabetes management and the achievements of this project as developed so far. According to the potential users separate patient-oriented questionnaire (PQ) and a doctor/nurse/technical staff-oriented questionnaire (DQ) have been designed.

From August 1999, requests for evaluation of the DiabNet project were sent to several diabetes discussion groups and mailing lists. The reason for choosing those potential respondents was that they were not only diabetic patients or health care professionals, but also Internet users with a certain level of computer literacy. These users were a representative sample of those potentially requiring the DiabNet service. The discussion groups and mailing lists were Diabetes-UK, Diabetes-talk, Children with diabetes, GP-UK and Teens with diabetes.

Also, DiabNet joined several Diabetes Web Rings, including British Diabetic patients Web Ring, Children with Diabetes Web Ring and The Diabetes Web Ring. The figures and analysis discussed below relate to the period between 15th and 31st August, 1999, but they do not differ much from the general trend observed subsequently. 450 unique users (from distinct IP addresses) who visited the DiabNet web site have been recorded by the IIS web service log. A total of 24 users, out of 450, completed the web-based questionnaires (21 PQ and 3 DQ).

The majority of the respondents are from the United States. According to the server log, actual user sessions were triggered from the United States (492), UK (88), Japan (45), Australia (11), Canada (9), Brazil (3), Denmark (3), South Africa (3), New Zealand (2) and Netherlands (2). 67% patients are type 1 (insulin-dependent) diabetic patients, 14% are type 2 (non insulin-dependent) and 19% are other types.

Table 1 lists the questions, frequency of each answered option, histogram of these options and comments of the author based on the analysis. Questions were asked of both PQ and DQ respondents, unless specified. The choices (C1-C5) correspond to the scores (1-5).

Table 1: Feedback analysis of disease management

From the feedback provided by the DiabNet questionnaire. the youngest user is 13 years old while the oldest is 62. In those diabetes discussion groups dedicated to teenage diabetic patients, juvenile diabetic patients are very active in learning to use the computer as a way of communicating with others. However, older patients seem less active due to age and health reasons. Actually, the Internet, to some extent, is a medium which is ideally suited to the elderly: there is no need to leave the house; they can choose their own large font size and colour to read if the eyesight is not good, etc.There is evidence that more and more elderly people are getting on-line, but the uptake just takes time.

Commentary and Conclusions

DIABNet is a successful example of the combination of various modern Internet technologies and telecommunication technologies into an integrated telemedicine solution for diabetes management.

The Data Viewer and Interpreter module made DiabNet not only a simple electronic logbook, but a data analysis tool as well. At present patients often perform self monitoring of their blood glucose month after month, producing large volumes of data without any appreciable or appropriate feedback from health care professionals. When assisted by DiabNet, they can review and analyse their historical data to facilitate the adjustment of their current treatment and regimen. DiabNet is also a potential tool for analysing existing high-volume data sets via the Internet. such as the UK diabetes data set from 102,927 patients (UKDIBS project, Diabetes UK [20]).

Since the majority of diabetic subjects have non-insulin dependent diabetes mellitus (NIDDM), not all the advanced features of DiabNet mentioned will necessarily apply. Nevertheless, improvements in diabetes care through the use of web-based services, such as the centralised database and customised search facility, provide an affordable method of improving care for patients with both IDDM and NIDDM.

It goes without saying that the audience of the DiabNet web service must be computer literate, or at least web browser literate. A question is then raised: how many diabetic patients are sufficiently computer literate to enjoy the web-based service? Although the exact number of Internet users among diabetic subjects is unknown, there is no evidence to distinguish them from ordinary people who are potential Internet users. An Internet research firm (eMarketer) recently predicted that the number of Internet users, worldwide, will nearly quadruple over the next five years, from 95.43 million in 1998 to 350.0 million by the year 2003. This represents an average annual growth rate of 29.8%.

This work is beneficial to diabetic patients, to health care professionals and to the health care system overall. For the diabetic patient, access to clinical data and treatment regimen information is only a "click away", so is their doctor and medical knowledge more generally. For the health care professional, it will be easy to trace the records of different patients in an efficient way. The data analysis functionality provided by DiabNet will facilitate their making of decisions regarding treatment. Used in the health care system, this approach can enhance the communication in a shared care group, combining primary care, hospital care and tertiary care. The expected savings in cost and time are huge, and may affect the whole health care delivery system. As a user of DiabNet said, "This is a powerful first step in the right direction." Internet and other IT technologies are going to play an increasingly important role in the changing health care system in the next century.

In order to improve the performance and strengthen the functionality of DiabNet, further work is needed to (i) enhance functionality, (ii) reinforce security and (iii) speedy up operations. Further versions of DiabNet will provide services for Type 2 [NIDDM] patients as well as for those on insulin pump therapy. In these cases the DiabNet database should provide data fields for diabetes medication, diet and exercise control information, in addition to insulin units and blood glucose measurements which can be accessed and queried. It is natural for DiabNet to connect to the HIS at a service level or even at a database level so as to reduce costs and improve efficiency. The establishment of a DiabNet paging system will enable the consultant's availability to be traced and thus give a quick response to a tele-consultation request.

The Data Viewer and Interpreter module can be strengthened by combining time series and temporal reasoning techniques with knowledge discovery functionality such as data mining and automated reasoning. Just as with any other Internet-based system, DIABNet also needs more security technology to protect its content. Future DiabNet versions will be complemented with state of art encrypting/decrypting techniques for protecting patient data, access authorisation security and secure network protocols.

The DiabNet server with its current hardware configuration can meet its small-scale test requirement (maximum bandwidth 24,994 k/day). However, some users still reported that the network speed was not very satisfactory on August 24, 1999 when the maximum number of concurrent users exceeded ten. The speed of a web site is crucial as users are going to be frustrated by slow downloads.

Research on a wide variety of hypertext systems has shown that users need response times of less than one second when moving from one page to another if they are to navigate freely through an information space. Based on the current Internet framework, it is impossible to achieve sub-second response times. The minimum goal for response times should therefore be to get pages to users in no more than ten seconds, since that is the limit of people's ability to keep their attention focused while waiting.

Advanced products such as fast analogue modems (56 kbps), ISDN modem (112 kbps) and leased lines are recommended to the individual user to achieve best web performance. Heading today's Internet technology, future networks like NGI (Next Generation Internet), I2 (Internet 2), vBNS (very high-performance Backbone Network Service) are aiming to bring Internet speed, over OC (Optical Carrier level) infrastructure, 100 to 1000 times faster than on today's Internet. Future development of DiabNet is also expected to be implemented over these fast networks.

References

[1] A.F. Amos, D.J. McCarty, and P. Zimmet, "The rising global burden of diabetes and

its complications: estimates and projections to 2010," *Diabet. Med.*, vol. 14 (suppl. 5), pp. S1-S85, 1997.

[2] B. Leese, "The costs of diabetes and its complications," *Soc. Sci. Med.*, vol. 35, pp. 1303-1310, 1992.

[3] A. Gray, P. Fenn, and A. McGuire, "The cost of insulin-dependent diabetes mellitus in England and Wales," *Diabet. Med.*, vol. 12, pp. 1068-1076, 1995.

[4] V. Jablonski, "Computer simulation saves many years of testing in development of first continuous glucose monitor," *Measurement and Control*, vol. 31, pp. 232-233, 1998.

[5] The Diabetes Control and Complication Trial research Group, " The effect of intensive treatment of diabetes on the development and progression of log-term complications in insulin-dependent diabetes mellitus," *New Engl. J. Med.*, vol. 329, pp. 977-986, 1993.

[6] UKPDS 33, "Intensive blood glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes," *Lancet*, vol. 352, pp. 837-853, 1998.

[7] M. Levy, P. Ferrard and V. Chirat, "SESAM-DIABETE, an expert system for insulin requiring diabetic patient education", *Comput. Biomed. Res.*, vol. 22, pp. 442-453, 1989.

[8] Business and Technology Editors, " Symbol announces Duke University Medical Center's use of Symbol Palm Computer for enhanced patient care: shirt-pocket device provides instant patient information." *Business Wire*, 23 February 1999.

[9] S. Osaki. "Handheld computer and calculator collection", http://member.nifty.ne.jp/handheld/handheld/handheld.html (accessed October 2000).

[10] F. Del Pozo. E.J. Gomez, and M.T. Arredondo, "A telemedicine approach to diabetes management." *Diab. Nutr. Metab.*, vol. 4 (suppl. 1), pp. 149-153, 1991.

[11] E.J. Gomez. F. Del Pozo. and M.E. Hernando. "Telemedicine for diabetes care: the DIABTel approach towards diabetes telecare." *Med. Inform.* Vol. 21, pp. 283-296, 1996.

[12] E.J. Gomez. P.J. Caballero. N. Malpica, and F. Del Pozo, "Optimisation and evaluation of an asynchronous transfer mode teleradiology cooperative system: the experience of the EMERALD and the BONAPARTE projects," *Comput. Meth. Prog. Biomed.*, vol. 64. pp. 201-214. 2001.

[13] K. Gerlach. A. Kaeding. S. Kottmair. D. Westphal. G. Henning, and K. Piwernetz, "The implementation of a Quality-Net as a part of the European Project DiabCare Q-Net." *IEEE Trans. IT Biomed.*, vol. 2, pp. 95-104. 1998.

[14] R. Bellazzi. S. Montani. A. Riva. and M. Stefanelli. "Web-based telemedicine systems for home-care: technical issues and experiences." *Comput. Meth. Prog. Biomed.*, vol. 64, pp. 175-187. 2001.

[15] E.R. Carson. T. Deutsch. H.J. Leicester. A.V. Roudsari. and P.H. Sönksen, "Challenges for measurement science and measurement practice: the interpretation of home-monitored blood glucose data." *Measurement*, vol. 24, pp. 281-293, 1998.

[16] R. Bellazzi et al., "Intelligent analysis of clinical time series by comining structural filtering and temporal abstractions." *AIMDM '99, LNAI 1620*, pp. 261-270, 1999.

[17] R.L. Bashshur. "On the definition and evaluation of telemedicine," *Telemed. J.*, vol. 1 (1). pp. 19-30, 1995.

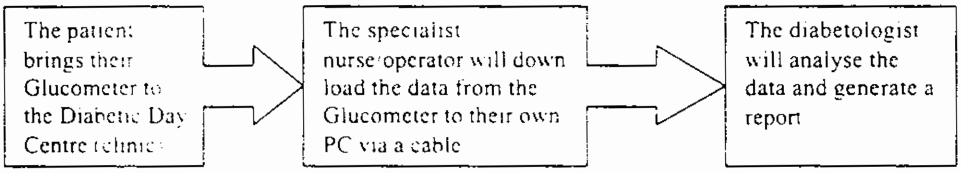
[18] E.R. Carson. D.G. Cramp. A. Morgan. and A.V. Roudsari, "Clinical decision

support, systems methodology, and telemedicine: their role in the management of chronic disease," *IEEE Trans. IT Biomed.*, vol. 2, pp. 80-88, 1998.
[19] D.G. Cramp, and E.R. Carson, "A model-based framework for assessing the value of ICT driven healthcare delivery," *Health Informatics, vol. 7, 89-94, 2001.*
[20] Diabetes UK http://www.diabetic.org.uk (accessed July 2001)

Legends

Figure 1. A conventional consultation session.
Figure 2. Prototype of a telemedicine diabetes care system.
Figure 3. Client-Agent-Server paradigm.
Figure 4. DiabNet web page structure.
Figure 5. The home page of the DiabNet web site.
Figure 6. Registered user's page.
Figure 7. BG by Days of Period Tab.
Figure 8. BG Scatter Tab.
Figure 9. BG by Time of Day Tab.
Figure 10. Frequency Distribution Tab (in regular stack bar).
Figure 11. Centred Moving Averages.
Figure 12. Statistics Tab.
Figure 13. Tele-consultation Request Page (NetMeeting session)

Figure 1: A conventional consultation session



The patient brings their Glucometer to the Diabetic Day Centre (clinic)

The specialist nurse/operator will down load the data from the Glucometer to their own PC via a cable

The diabetologist will analyse the data and generate a report

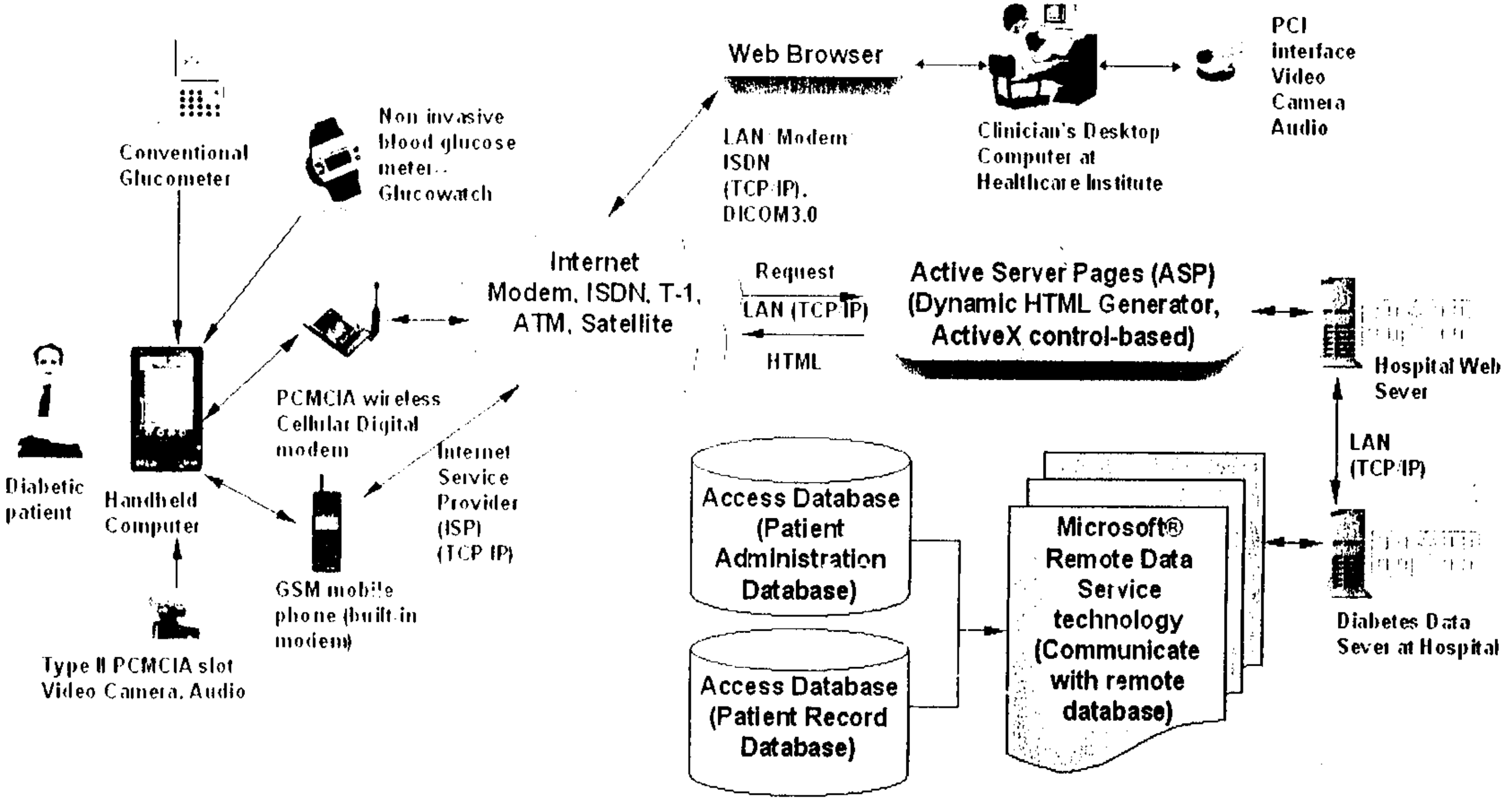


Figure 2: Prototype of a telemedicine diabetes care system

Figure 3 : The Client-Agent-Server paradigm

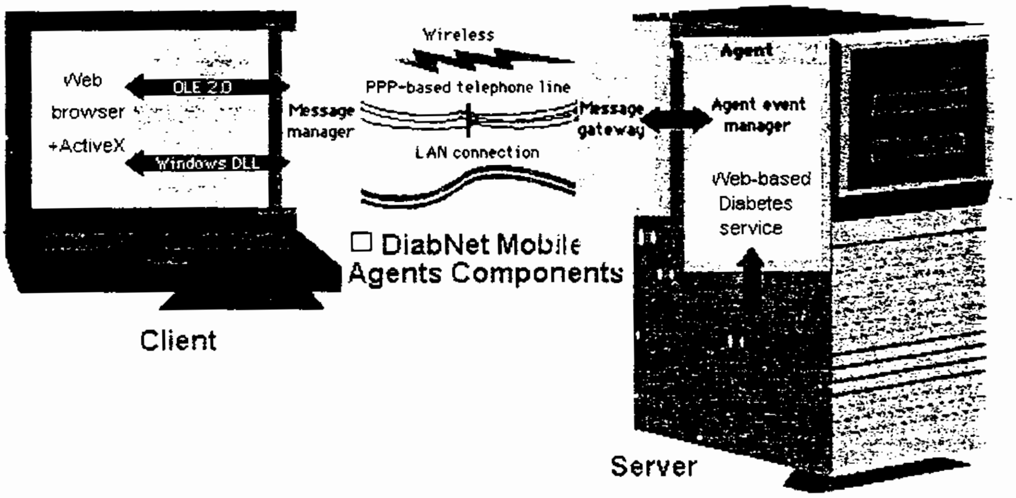


Figure 4: DiabNet web page structure

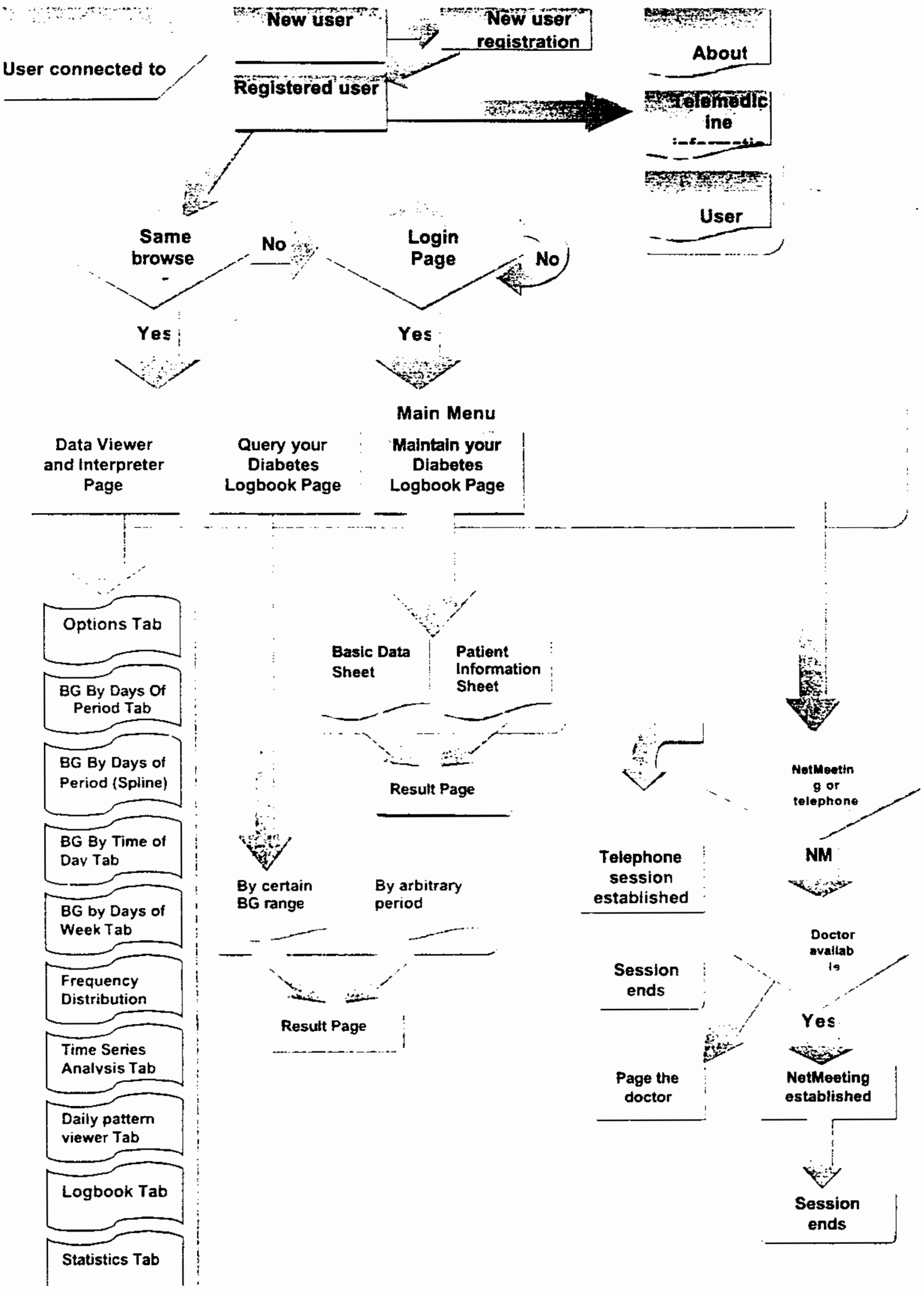
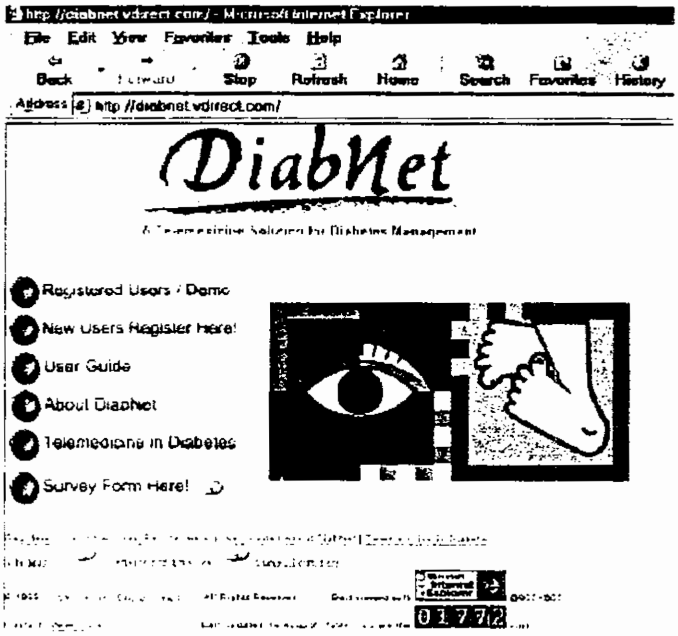


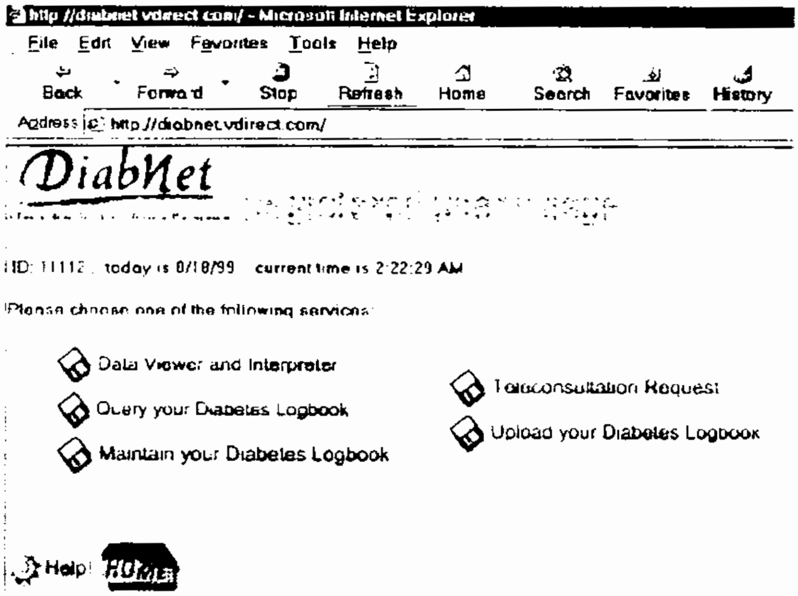
Figure 5 : The home page of the DiabNet web site



A Telemedicine Solution for Diabetes Management

- Registered Users / Demo
- New Users Register Here!
- User Guide
- About DiabNet
- Telemedicine in Diabetes
- Survey Form Here!

Figure 6: Registered user's page



ID: 11112 today is 0/18/99 current time is 2:22:29 AM

Please choose one of the following services:

- Data Viewer and Interpreter
- Query your Diabetes Logbook
- Maintain your Diabetes Logbook
- Teleconsultation Request
- Upload your Diabetes Logbook

Figure7: BG by Days Of Period Tab

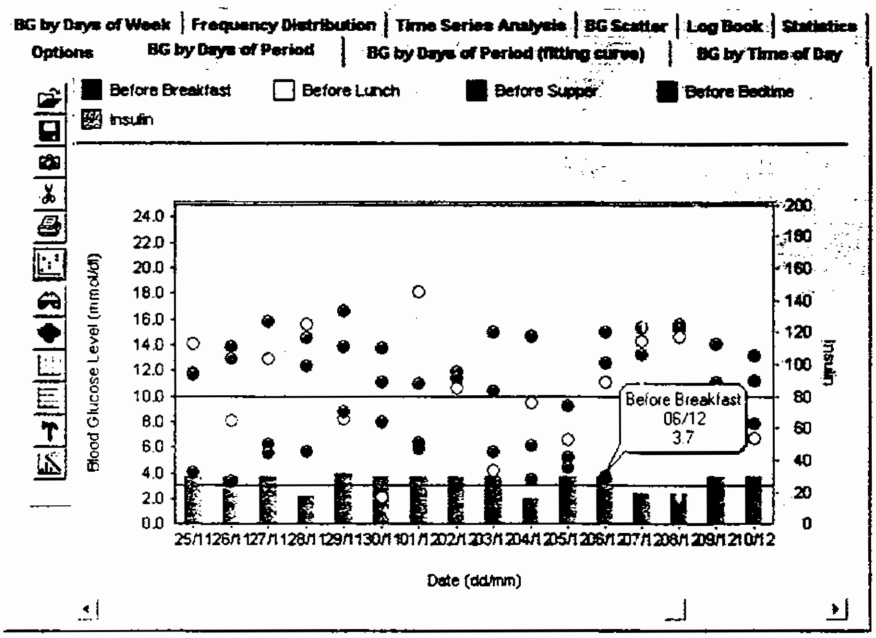


Figure 8: BG Scatter Tab

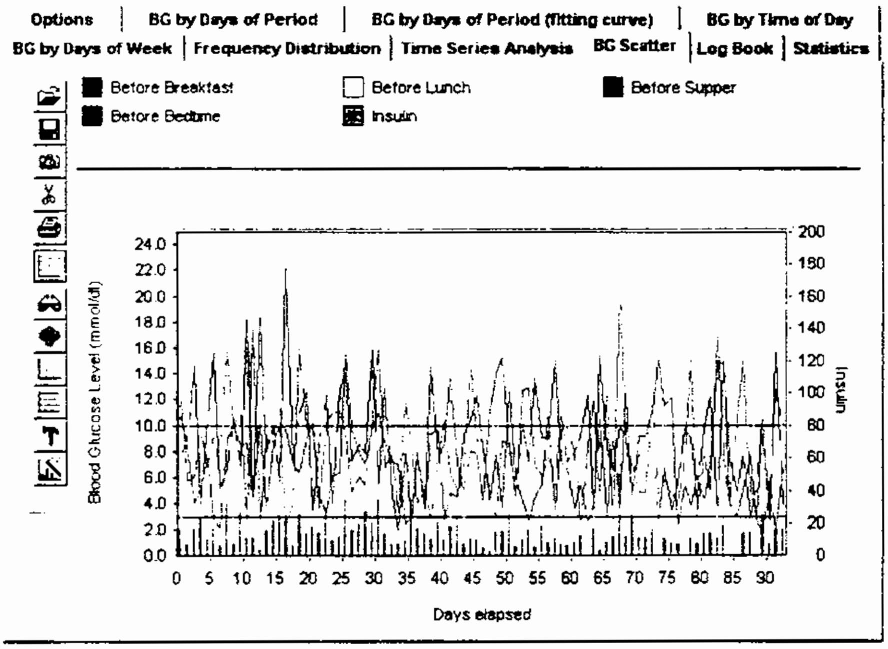


Figure 9: BG by Time of Day Tab

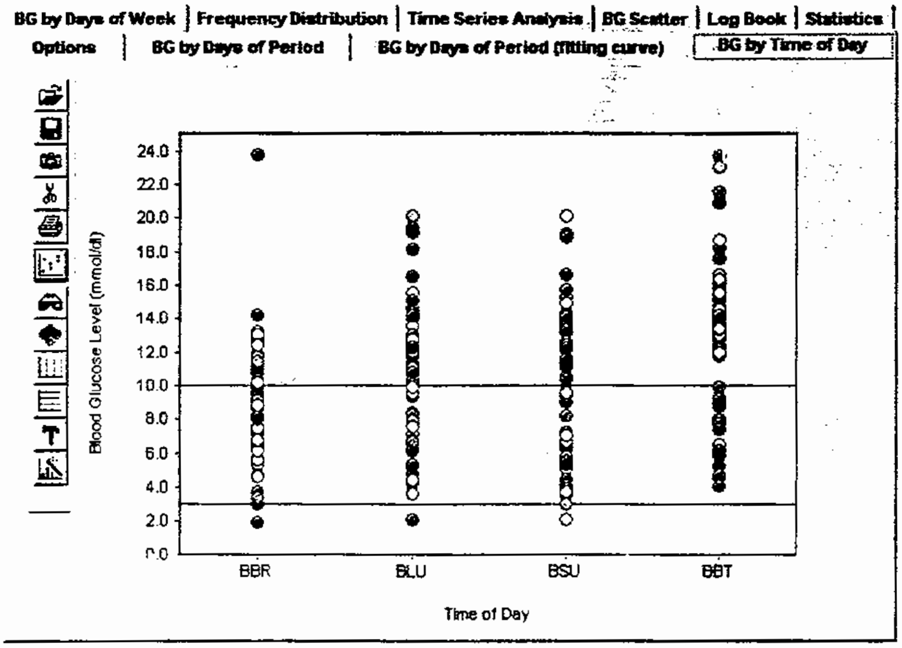


Figure10:Frequency Distribution Tab (in regular stack bar)

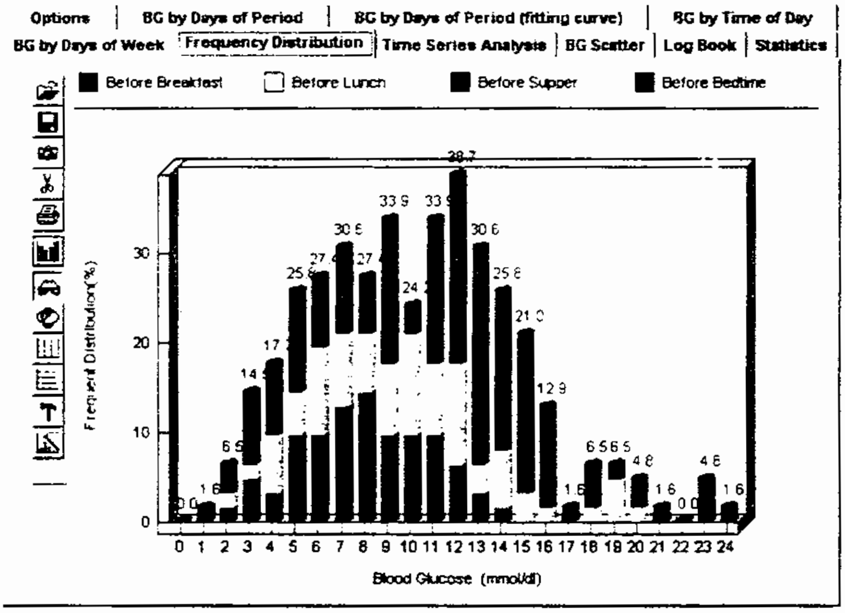


Figure11: Centred Moving Averages

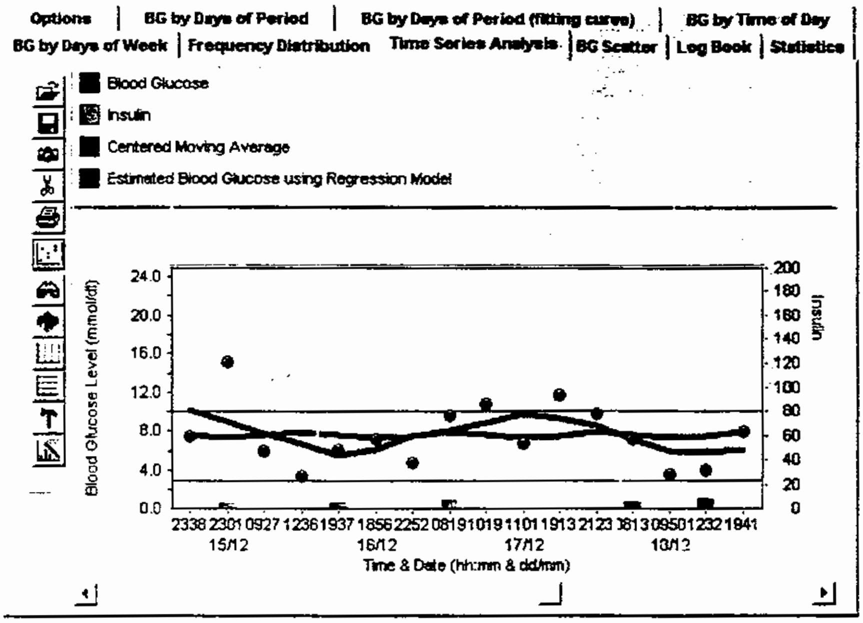
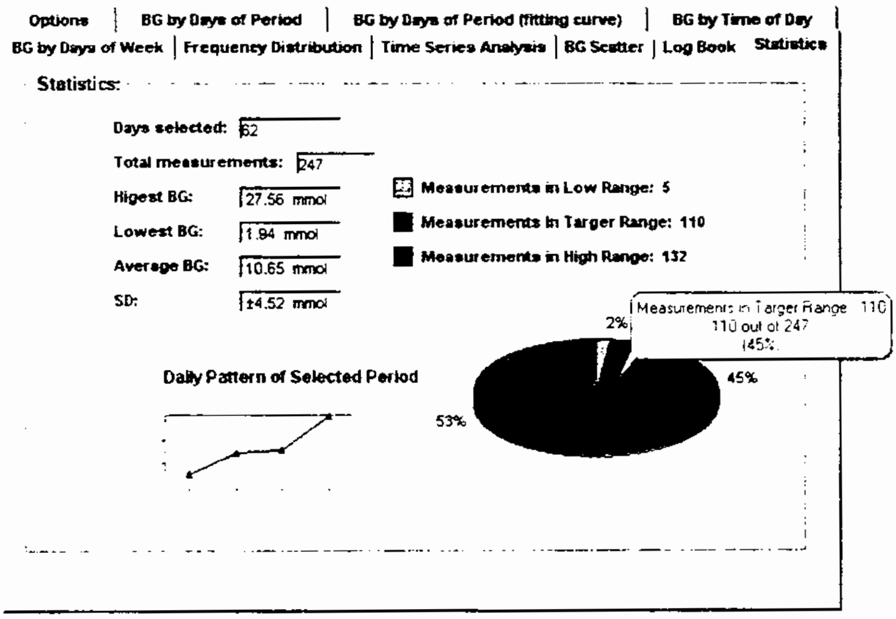


Figure12: Statistics Tab



Statistics:

Days selected: 62

Total measurements: 247

Higest BG: 27.56 mmol

Lowest BG: 1.94 mmol

Average BG: 10.65 mmol

SD: ±4.52 mmol

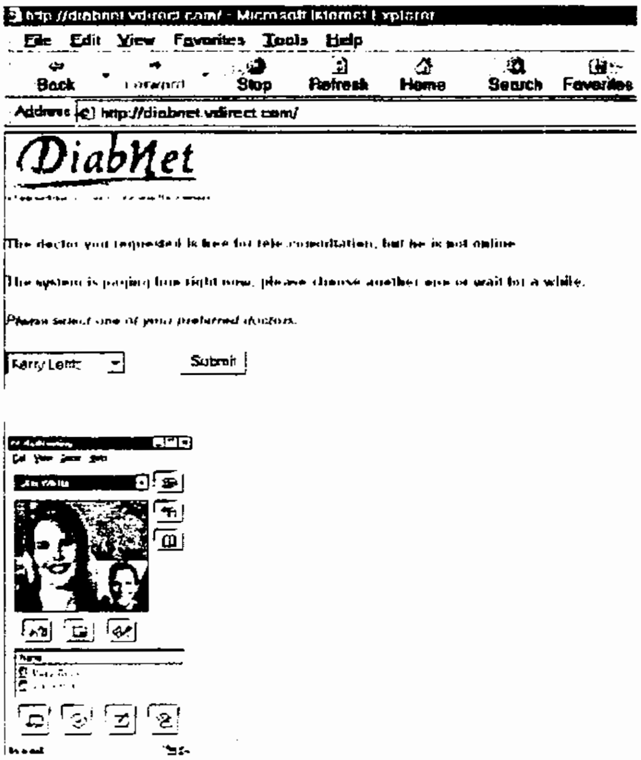
▨ Measurements in Low Range: 5

■ Measurements in Targer Range: 110

■ Measurements in High Range: 132

Measurements in Targer Range 110
110 out of 247
(45%)

Figure13: Tele-consultation Request Page (NetMeeting session)



The doctor you requested is free for tele-consultation, but he is not online.

The system is paging him right now, please choose another one or wait for a while.

Please select one of your preferred doctors.

Kerry Lentz

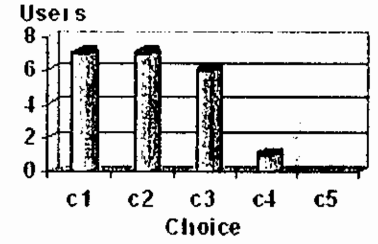
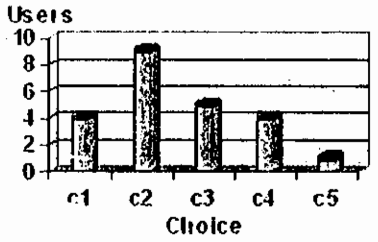
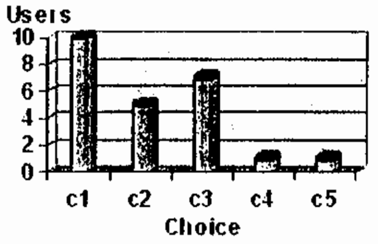
Statement of Questionnaire	From:	C1	C2	C3	C4	C5	To:	Chart
I am satisfied with my current diabetes disease management. (PQ)	very satisfied	7	7	6	1	0	not satisfied	
I am satisfied with the timing of hospital appointments. (PQ) I hear complaints from patients about the timing of hospital appointments (DQ)	very satisfied	4	9	5	4	1	not satisfied	
I think the web-based diabetes management approach will improve the control of blood glucose if it is well designed.	agree	10	5	7	1	1	disagree	

Table 1: Feedback analysis of disease management

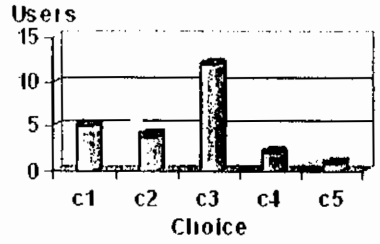
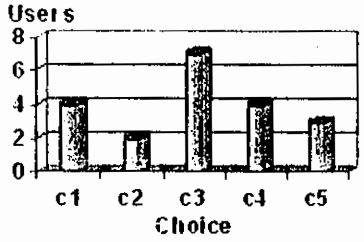
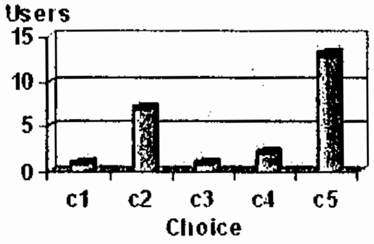
Statement of Questionnaire	From:	C1	C2	C3	C4	C5	To:	Chart
I am satisfied with the ease of getting to the diabetes clinic centre (PQ) I hear complaints from patients about the difficulty of getting to the diabetes clinic/centre. (DQ)	very satisfied	5	4	12	2	1	not satisfied	
		The average score is 2.6. More than half of the respondents are not very happy about the ease of getting to the clinic, which shows the need for an Internet-based consultation.						
I often visit my diabetes consultant/specialist nurses. (PQ)	once a month	4	2	7	4	3	once every 5 months or less	
		The average score is 3, which confirms that once every 3 months is the usual frequency of visits that patients currently follow.						
I am using/will use a palmtop/handheld computer in my diabetes management	most likely	1	7	1	2	13	unlikely	
		The average score is 3.8. More than half said they would not use a handheld computer in diabetes management. Probably the complexity, size and weight of current handheld computers are still unacceptable.						

Table 1: Feedback analysis of disease management (cont'd.)

Agent Architecture for Smart Adaptive Systems

Lenka Lhotska, Olga Stepankova
Gerstner Laboratory, Department of Cybernetics,
Czech Technical University in Prague, Faculty of Electrical Engineering,
Technická 2, CZ-166 27 Prague 6, Czech Republic
Phone: +420-2-24353933, Fax: +420-2-24311081
E-mail:lhotska@fel.cvut.cz, step@labe.felk.cvut.cz

ABSTRACT: This paper discusses the possibility to utilize a multiagent system as a platform for development of smart adaptive systems. A knowledge-based model of the agents' mutual awareness (social knowledge) is presented. A tri-base acquaintance (3bA) model is used for formalization of agent's social behaviour knowledge and agent's cooperation neighbourhood knowledge. The paper focuses particularly on analysis of content of 3bA knowledge bases for applications in medical decision support.

KEYWORDS: agent, multiagent system, tri-base acquaintance model, cooperation, meta-agent, learning

INTRODUCTION

Starting from the definition of adaptivity as the property of a system to adequately perform even in non-stationary environments where significant smooth changes of the main characteristics can be manifested we can claim that the multiagent systems represent a suitable platform for development of adaptive systems. Since many real world problems such as production planning, supply chain management, engineering design, intelligent search, medical diagnostics, robotics, etc. are naturally distributed multiagent systems [1], [2], [7] offer efficient problem solving platform. They eliminate limitations on the processing power of a single monolithic system. Distribution also brings inherent advantages of distributed systems, such as scalability, fault-tolerance, parallelism, robustness, etc.

However, there are questions connected with mutual behaviour of the agents, of ownership of global knowledge, of structure and content of this knowledge.

The applicability of agent architecture developed in the Gerstner Laboratory has been proven by applications in several areas, e.g. in the field of production planning and scheduling (system ProPlanT), supply chain management (project ExPlanTech), coalition formation (system CPlanT) [3], [8]. The intended application we focus on is decision support in medical diagnostics.

This paper presents a short description of the developed 3bA model [3] and analysis of content of individual knowledge bases, in particular for applications in medical decision support. There are discussed questions of learning in multiagent system as well

DIAGNOSTICS

Diagnostics is one of the most frequently encountered decision making tasks of human activity. Diagnostic tasks are usually understood as classification, i.e. an object needs to be assigned into a predefined group or class based on a number of observed attributes related to that object. Many problems in business, science, industry, and medicine can be treated as classification problems. Let us mention several examples: medical diagnosis, fault detection, bankruptcy prediction, quality control, handwritten character recognition, and speech recognition.

Intelligent e-Health Applications in Medicine
Special Session Proc. EUNITE 2002, Albufeira, Portugal, Sept. 19-21, 2002
Published by the University of the Aegean

When we look at data and information from the above-mentioned areas more closely we find out that there is no generally applicable best method or technique for evaluation of particular data. Each one has its relative strengths and weaknesses. Some can only produce an approximate solution, but do so comparatively quickly; others are more accurate, but relatively slow. Furthermore, a given technique's performance is often dependent on the nature of the data set (some work well with noisy data, others do not; some work well with data that has a high signal strength, others work comparatively better with a low signal strength, some can cope with missing data, others do not). However, the things are even more complicated. It is often impossible to determine a priori which technique is the most appropriate for a given data set or its part. There are several reasons for that, namely if the data set is too large the user is usually not able to evaluate the quality of data manually, the user may not be very experienced, the user may skip important part of data, etc. From that basic requirement on the system being developed follows: the system needs to be responsive to its problem-solving context. In medicine, this situation occurs whenever it is necessary to evaluate long-term recordings of ECG, EEG, etc. and to combine the results with other types of data, such as anamnestic data, results of biochemical analysis, etc.

To overcome the problems associated with selecting a single technique, there can be developed a system that allows multiple methods to coexist. However, as examples from other domains show (e.g. image processing), such systems or tools typically place a significant burden on the user. For each technique, the user is expected to know its problem solving characteristics, be able to judge when, where, and how to apply it, and to determine how best to integrate and fuse the results it produces. It would be impossible to solve this problem using a single monolithic system (e.g. expert system) because it does not allow integration of different techniques and evaluation of partial results reached by these techniques. Therefore we have decided to use developed agent model as a platform for design and development an open system that will provide a wide range of uncoupled base techniques (represented by separate agents). The 3bA model allows determining at run-time which techniques are appropriate in which circumstances. The interchange of partial and final results between individual agents is directly supported as well.

TRI-BASE ACQUAINTANCE MODEL AND ITS APPLICATION TO DIAGNOSTIC TASKS

The basic architecture of an agent in our system consists of a functional body (usually a stand-alone program with a well-defined functionality) and a wrapper (which is responsible for involvement of the agent into the community of agents) – see figure 1. The tri-base acquaintance (3bA) models are encoded in agents' wrappers (see figure 2).

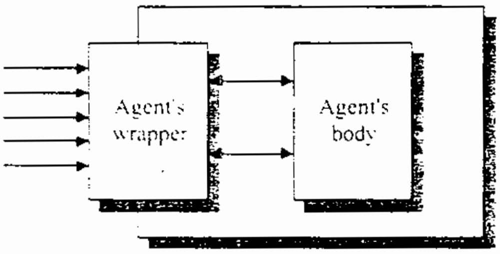


Figure 1. Structure of an agent

The 3bA models have several important purposes

- to limit explosive communication in multiagent system,
- to ensure immediate reply in time-critical situation,
- to generate and maintain databases of information sources

Instead of communicating with the collaborating agents in order to find out certain information about the community, an agent equipped with the acquaintance model consults this social knowledge stored in its wrapper instead. This feature is very closely linked with the second one. If we require immediate reply to an input or stimulation there is usually not much time for communication with collaborating agents. The agent must react without any delay and therefore it must have relevant information at hand, e.g. which agent should execute the task. Using negotiation in such cases is not acceptable.

The agents can browse the Internet and search for relevant information. If such information is later used the source is included on the list of potential sources of information for further use.

Within the 3bA model each agent maintains three knowledge bases where all the relevant information about the rest of the community is stored, namely the Co-operator Base, the Task Base, and the State Base.

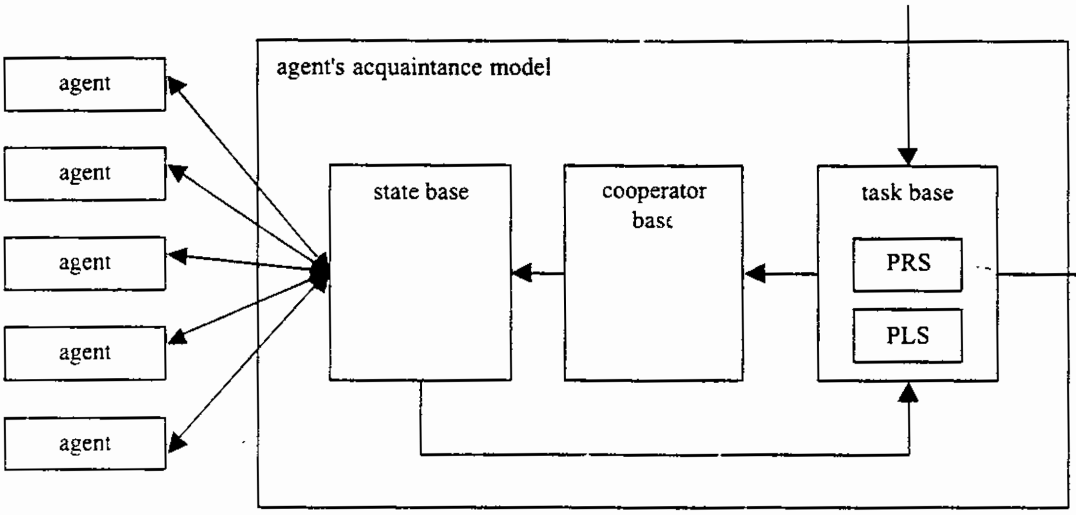


Figure 2: Tri-base acquaintance model

The Co-operator Base (CB) maintains permanent information on cooperating agents, i.e. their address, communication language, and their predefined responsibility. It contains information about required form of input data for agent's body. This may speed up process of selection of a proper agent that is able to perform required task since in diagnostics there is usually used large volume of input data of different nature and form. This type of knowledge is expected not to be changed very often.

The State Base (SB) contains the information about collaborating agents, i.e. about their current state. For example, if the agent is to process data that must be pre-processed by another agent, the former agent must know whether the latter has already started/finished the pre-processing. The SB stores in its agent section (AS) all information on current load of cooperating agents. This part of the state base is updated frequently and informs the agent which of the collaborating agents are busy and which of them are available for collaboration. In the task section (TS) there is stored information on statuses of tasks the agent is currently solving.

The Task Base (TB) has two sections: problem section and plan section. In the plan section (PLS) it maintains the actual and most up-to-date plans on how to carry out those tasks, which are the most frequently delegated to the agent. In addition, it maintains information about the most suitable chains of agents that can perform certain tasks from collecting, pre-processing to evaluation. This information is updated by the metaagent that learns from successful and failed cases in the past (see below). The Task Base stores in its problem section (PRS) general problem solving knowledge on possible decision making with respect to input data type and expected output. In case of an agent responsible for certain data pre-processing the task base will contain knowledge about data types that can be pre-processed by the agent's body. It may contain knowledge about possible outputs of the agent's body, namely whether the output represents intermediate results that should be sent for further processing or whether it represents final results that should be sent to the user. In case of an agent responsible for evaluation of pre-processed data its task base will contain information about data types that may arrive at its input and about procedure for checking data consistency. Considering time critical applications, response time of agents is important as well. Therefore information about average, maximum and minimum response time of the agent is attached to description of tasks the agent's body is able to execute. If the response is required in shorter time than the agent is able to deliver then another agent must be found that is able to deliver response in required time. Depending on type of the task, suboptimal solution delivered in shorter time can be preferred to late optimal solution. For example, classification using neural network or decision tree is usually faster than using case-based reasoning.

How is the knowledge maintained and updated in individual bases? As we have already mentioned, the co-operator base collects knowledge of rather permanent nature and we do not expect to update it very often besides the register phase. Once a new agent registers with a community (by means of contacting a central agent – facilitator that administers all the data about the community members), the facilitator replies the newcomer by providing information about the community members. In addition, it informs other agents about the newcomer.

The state base, which models the actual state of the collaborating agents, is maintained by a simple subscribe/advertise mechanism. After parsing the problem solving knowledge (in PRS), each agent identifies possible collaborators and subscribes them for reporting on their statuses. The subscribe/advertise mechanism facilitates the subscriber to make the best decision with no further communication.

The task base is kept up-to-date by periodic revisions of the pre-prepared plans in the PLS. Such a revision represents verification/modification of the plan by exploring the information kept updated in both the co-operator and state bases. The knowledge contained in the PRS can be maintained e.g. by the meta-agents.

Content of all three bases is dependent on the tasks the agents are supposed to solve. In case of planning agents the dominant role is played by knowledge of task decomposition and responsibility delegation. In case of configuration agents the knowledge stored in the task base can be used to lead the communication scenarios. The diagnostic agents contain the social knowledge about data sources, about the process of finding appropriate data and about the current progress in required data processing by the other agents.

CASE STUDY: MEDICAL DIAGNOSTICS

Each data and signal pre-processing and evaluation technique can be regarded to be an autonomous software agent that cooperates, communicates and coordinates, if necessary, with other agents to try to satisfy the global goal. Since there are several modules at hand that perform pre-processing (e.g. wavelet transform module for ECG signal pre-processing extracts significant attribute values from the raw signal [4], EEG segmentation module extracts significant parts from the signal and calculates parameters of these parts) and evaluation [5] as well we have decided to use wrapping of these modules into agents' bodies.

Components of the system ADIA (Agents for DIAgnostics) can be divided into three main layers, namely agents for data collection, agents for pre-processing and processing (determination of diagnosis), and agent for final evaluation. In addition to these agents, we propose a metaagent that is an independent agent observing the community. It has two roles: passive role (visualisation of community structure, distributed solutions, user interface, etc.) and active role (it affects community operation - invokes operation sequences, learns from observations and tries to improve behavior of the whole community). The proposed architecture is illustrated in figure 3. It is assumed that the agents are running on different machines and connected via Internet. That means that security issues must be taken into account as well. However, they will not be discussed in this paper.

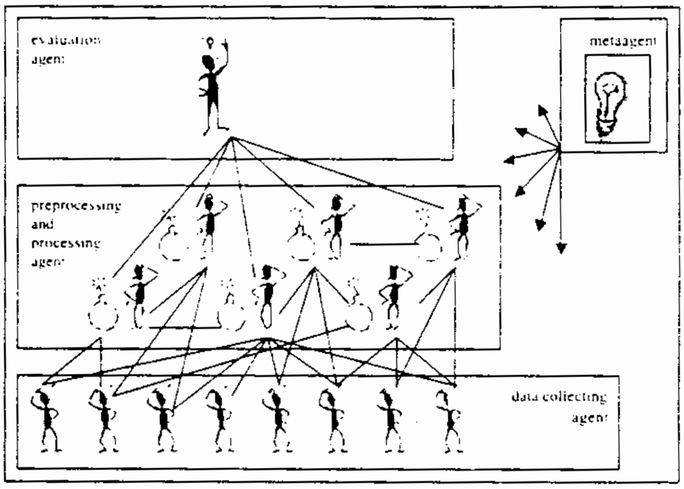


Figure 3 - Proposed architecture

At present, there are several modules ready for encapsulation into agent's structure. On the pre-processing level following modules have been developed, implemented and tested: wavelet transform module for ECG pre-processing (extraction of significant attribute values) [4], GUHA (for identification of relations between certain features), EEG module based on FFT (signal segmentation and calculation of significant values of these segments). There are other modules currently being developed: ICA for EEG analysis, wavelet transform module for EEG pre-processing.

On the processing and evaluation level there have been several modules developed and implemented as well. The rule-based diagnostic expert system FEL-EXPERT [6] has been verified in many practical applications. Other modules that have been already used in practical applications are decision tree module (based on Quinlan's ID3 algorithm) [4], instance-based reasoning system iBARET [5], fuzzy system module, and neural network module. We plan to introduce further methods, as for example Bayesian reasoning, reinforcement learning, and support vector machine. These modules will be used for decision support (after learning from the training data sets).

A very important question is how the agents should recognize which one is to start the pre-processing or directly processing (without pre-processing step) of accepted input data. As it has been mentioned above, it is advantageous to use different methods for different data types, depending on the nature of the data, presence of noise, etc. In some cases, e.g. processing of anamnestic data or numerical values from laboratory tests, there is usually not necessary to apply any pre-processing and processing step can be initiated.

As it has been already mentioned and following this analysis we have decided to introduce a metaagent into the system architecture that would be able to learn from experience and thus to control cooperation of agents more efficiently. Since there is certain knowledge about advantages and disadvantages of application of individual methods to various data processing it is possible to formulate a set of initial rules (prior knowledge) for such a metaagent. If previously processed data type arrives and the metaagent informs the processing agents about it, they will know which sequence of pre-processing methods and successive processing and evaluation to use (because such sequence was successful in the past case). If unknown (unrecognised) data type arrives, then all agents should be given the opportunity to try to solve the problem. However, in this case we get a number of different results. The question is how to decide which of them are relevant, and which are the best. The metaagent serves as user interface as well. It enables visualization of the decision-making process and interaction with the user. If the suggested solution does not seem appropriate to the user, he/she can enter the decision making process and invoke certain agents manually.

Although the original 3bA model does not support proactive information search and retrieval, it is possible to equip the agents with this ability (adding relevant knowledge in the knowledge bases in agents' wrappers). This ability enables to locate and connect the ultimate service provider with the ultimate requesters in open environments such as the Internet. This is the case of connecting a human user with required processing agent. Since the user communicates with the agents using user interface agent we can equip this agent with necessary knowledge. Then it can ensure all functions that are realized, for example, by middle agents in the RETSINA system [9]. Let us describe one such scenario. The user has received ECG data from another source, thus he/she does not need any of the data collecting agents, but the services of preprocessing, processing and evaluating agents are required. The user interface agent starts communication with relevant agents and sends them data. When the whole processing and evaluation process is finished the user receives results.

CONCLUSION

This paper has described the possibility to use agent architecture for development of smart adaptive system. It has been illustrated by a case study: design of a multiagent system for supporting medical diagnostics. Cooperating agents provide a very natural means of automating pre-processing and (at least partially) evaluation of vast amount of medical data utilizing all available medical knowledge. The next phase of the project is targeted at implementing and putting the system into full use.

There are a number of issues that require further investigation. First, a more comprehensive set of pre-processing and processing techniques is required. Second, the agents should be able to adapt and learn from the social interactions they experience. Agents should learn which acquaintances give reliable results in which circumstances. Based on this knowledge they should be able to adapt their selection appropriately. At the moment we have decided not to equip each agent with this ability but to introduce the metaagent that will do this work for the whole community. This metaagent monitors and gathers all messages sent in the system. According to these observations it invokes desirable changes.

Of course, both versions (distributed and centralized learning) have their advantages and disadvantages. The distributed version is more effective and flexible because individual agents process smaller data sets (i.e. they process specialized data sets as ECG signals, EEG signal, anamnestic data, one agent does not process all data types), they can react faster to changes in the environment and the learning process may run faster. On the other hand, all information desirable for better learning may not be necessarily available to individual components. Another disadvantage is that learning components of each agent have to have essentially the same mechanisms available. That means implementing the same program several times. It is an advantage that agents do not have to monitor events in the whole system, but only in the part that is close to them. However, there is a hidden danger, namely they may not recognise undesirable phenomena in the system. Besides that, any intervention to the system structure may cause a necessity to modify all learning components. A single learning agent has disadvantages of a centralised solution, but on the other hand it has all available information from the whole agents' community at its disposal. In addition, all modifications and possible extension to other learning methods are much easier. The description of the system behaviour is more transparent.

As we have already mentioned, the multiagent system should be working in open environment and accessible through the Web. In such case, there are many issues to be solved, the most important ones security, reliability, accessibility, response time.

ACKNOWLEDGMENTS

The research described in the paper has been supported by the research program No. MSM 210000012 "Transdisciplinary Biomedical Engineering Research" of the CTU Prague, sponsored by the Ministry of Education, Youth and Sports of the Czech Republic and by MIRACLE - Machine Intelligent Research and Application Centre for Learning Excellence, EU project "Centre of Excellence", ICA1-CT-2000-70002.

REFERENCES

[1] Weiss G. (Ed.), 1999, "Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence", Cambridge, Mass. The MIT Press

[2] Wooldridge M.; Jennings N.R., 1995, "Agent Theories, Architectures, and Languages: A Survey", Intelligent Agents, Wooldridge M., Jennings N.R., eds. Heidelberg, Springer-Verlag LNAI 890

[3] Marik V.; Pechoucek M.; Stepankova O., 2002, "Organization of social knowledge in multi-agent systems". Integrated Computer-Aided Engineering 9 (2002), pp. 1-12

[4] Fejtova M.; Lhotska L.; Macek J., 2001, "ECG Event Detection and Classification Using Wavelet Transform and Decision Trees". EUNITE 2001 European Symposium on Intelligent Technologies, Hybrid Systems and Their Implementation on Smart Adaptive Systems. Aachen, Verlag Mainz

[5] Klema J.; Lhotska L.; Stepankova O.; Palous J., 2000, "Instance-Based Modelling in Medical Systems". Cybernetics and Systems 2000. Vol. 2. Vienna: Austrian Society for Cybernetics Studies, p. 365-370. - ISBN 3-85206-151-2

[6] Lhotska L., Marik V.; Vlcek T., 2001, "Medical applications of enhanced rule-based expert systems", International Journal of Medical Informatics 63, pp. 61-75

[7] Huang J., Jennings N.R.; Fox J., 1995, "An Agent-Based Approach to Health Care Management", Applied Artificial Intelligence 9(4)

[8] http://agents.felk.cvut.cz/main/index.php?home

[9] Sycara K., 2001, "Multiagent infrastructure, agent discovery, middle agents for web services and interoperation", ACAI 2001, Luck M. et al., eds. Heidelberg, Springer-Verlag LNAI 2086

Intelligent Multimedia Transmission for Back Pain Treatment

George Ghinea, George D. Magoulas and Andrew O. Frank*
Department of Information Systems and Computing
Brunel University, West London, U.K
Phone: +44-1895-203397, Fax: +44-1895-251686
email: {George.Ghinea,George.Magoulas}@brunel.ac.uk
*Department of Rehabilitation Medicine and Rheumatology
Northwick Park Hospital and Institute of Medical Research
Harrow HA1 3UJ, Middlesex
Phone: +44-208-8692104, Fax: +44-208-8692009
email:AndrewFrank@waitrose.com

ABSTRACT: Remote, multimedia-based, collaboration in back pain treatment is an option which only recently has come to the attention of clinicians and IT providers. The take up of such applications will inevitably depend on their ability to produce an acceptable level of service over congested and unreliable public networks However, although the problem of multimedia application-level performance is closely linked to both the user perspective of the experience as well as to the service provided by the underlying network, it is rarely studied from an integrated viewpoint. To alleviate this problem in the context of a multimedia application, a method is proposed in this paper for obtaining a priority order of low-level Quality of Service parameters, which would ensure that user-level Quality of Perception is maintained at an optimum level. Thus we present an approach that integrates technical concerns with user perceptual considerations for intelligent decision-making in the construction of tailor-made multimedia communication protocols. The proposed approach, based on multicriteria decision making, incorporates not only classical networking considerations, but, indeed, user preferences as well. Moreover, our approach also opens the possibility for such protocols to dynamically adapt based on a changing operating environment and user preferences.

KEYWORDS: Back pain, Adaptive transmission protocols, Quality of Perception, Quality of Service, Multi Criteria Decision Making, Fuzzy preference programming method.

INTRODUCTION

The deployment of Internet-based applications for patient care using advanced multimedia techniques aim to offer users of health services high-quality care over inexpensive communication pathways, using Internet-based, interactive communication tools However, the integrated use of telecommunications and information technology in the health sector leads to new challenges in data transmission.

Whilst the quality of multimedia transmission has traditionally been characterised using Quality of Service (QoS) parameters, the networking foundation on which current distributed multimedia applications are built either do not specify QoS parameters (also known as best effort service) or specify them in terms of traffic engineering parameters such as delay, jitter, and loss or error rates. However, these parameters do not convey application-specific needs such as the influence of multimedia content and informational load on the stakeholders. There is thus an architectural gap between the provision of network-level QoS and application-level user-centric requirements of the distributed multimedia applications. This gap causes distributed multimedia systems to inefficiently use network resources and results in poor end-to-end performance which in turn has a direct negative impact on the stakeholder experience of multimedia

In this paper, we present an approach to bridge the user-network gap by employing an intelligent decision making mechanism in the construction of adaptable transmission protocols, specifically tailored for the transmission of medical data. Accordingly, the structure of the paper is as follows: the next section introduces issues relating to back pain relevant to our work. The subsequent section presents the framework for adaptation used in our approach. Lastly, an application example is provided and concluding remarks are drawn.

BACK PAIN

According to a Department of Health survey, in Britain back pain affects 40% of the adult population, 5% of which have to take time off to recover [1]. This causes a large strain on the health system, with some 40% of back pain sufferers consulting a GP for help and 10% seeking alternative medicine therapy [1]. Due to the large number of people affected, backpain alone cost industry £9090 million in 1997/8, with between 90 and 100 million days of sickness and invalidity benefit paid out per year for back pain complaints [2]. Back pain is not confined to the UK alone, but is a worldwide problem: in the US, for instance, 19% of all workers' compensation claims are made with regard to back pain. Although this is a lot less than the percentage of people affected by backpain in the UK, it should be noted that not all workers are covered by insurance and not all workers will make a claim for backpain [3]. Moreover, back pain does not affect solely the adult population: studies across Europe [4] show that back pain is very common in children, with around 50% experiencing back pain at some time. Any improvement in the way that patients with backpain can be analysed (and subsequently treated) should therefore be viewed as one potentially capable of significantly saving both benefit expenditure and lost man-hours.

The problem with back pain is that "there exist no standardised clinical tests or investigations by which all people with low back pain can be evaluated" [5]. Nor will there ever be, as different people have different pain thresholds and will be affected differently. It is also difficult for medical personnel to know what has caused the backpain, as there are potentially many different causes behind it [2]. Not only is evaluation difficult, but, unfortunately, like most types of pain, back pain is also difficult to analyse, as the only information that can be used is suggestive descriptions from the patient. The need therefore for distributed, collaborative applications which allow communication and exchange of information between consultants, physiotherapists, and patients, becomes paramount.

The main medical work that is undertaken to resolve backpain tends to be with patients that have chronic backpain. However, these patients may have developed psychological and emotional problems, due to having to deal with the pain. Because of these problems, patients can have difficulty describing their pain, which can lead to problems during the treatment. In some patients, the psychological problems may have aided the cause of the backpain, by adding stress to the body, or the stress of the backpain may have caused psychological problems [6],[7]. It is because of this factor that patients suffering from backpain are usually asked to fill out questionnaires of different types in order to help the medical staff, not only to know where the pain is located, but also to identify the patient's mental state before treatment begins. In addition, the patient is usually required to mark on a diagram, usually of a human body, where the pain is located, and the type of pain. This type of diagram is known as a 'pain drawing' and is exemplified in Figure1, have been successfully used in pain centres for over 45 years [8] and act as a simple self-assessment technique, originally designed to enable the recording of the spatial location and type of pain that a patient is suffering from [7]. They have a number of advantages including being economic and simple to complete, and can also be used to monitor the change in a patient's pain situation [9].

INTELLIGENT TRANSMISSION OF BACK PAIN DATA

In the UK, the relative scarcity of back-pain rheumatology consultants, on the one hand, coupled with the widespread occurrence of back-pain in the general public, necessitates that technology, especially multimedia communication-related, be exploited in new ways. Whilst the idea of distributed collaborative environments for long-distance consultations and diagnostic is, by itself, not new, what is novel in our approach is the exploitation of multimedia perceptual results to optimise resource usage in data transmission. We are currently developing BPDRoPS, a distributed tool for remote e-collaboration in the treatment of back-pain. This tool, used by practitioners and patients alike comprises a two-way video link (over which interactions between the back-pain consultant, at one end, and GPs, physioterapists and/or patients, at the other, take place), a shared text box in which participants may specify various textual input (e.g. medication, details of rehabilitative exercises), as well as a further subwindow containing the details of the back-pain drawing as filled out by the patient. Moreover, BPDRoPS also integrates a repository of back-pain data, which can be accessed by either participant to the remote diagnostic session, a snapshot of which is given in Figure 2. BPDRoPS thus transmits video, audio, text and graphic content across inherently unreliable telecommunication networks. If this is done without regard for perceptual requirements, then chances are that transmission will be an ineffective, resource-consuming exercise, fraught with problems affecting multimedia transmission in general, such as delay, jitter and loss, and with consequently little take-up by stakeholders.

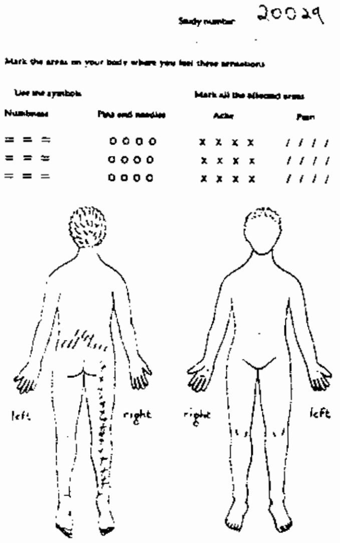
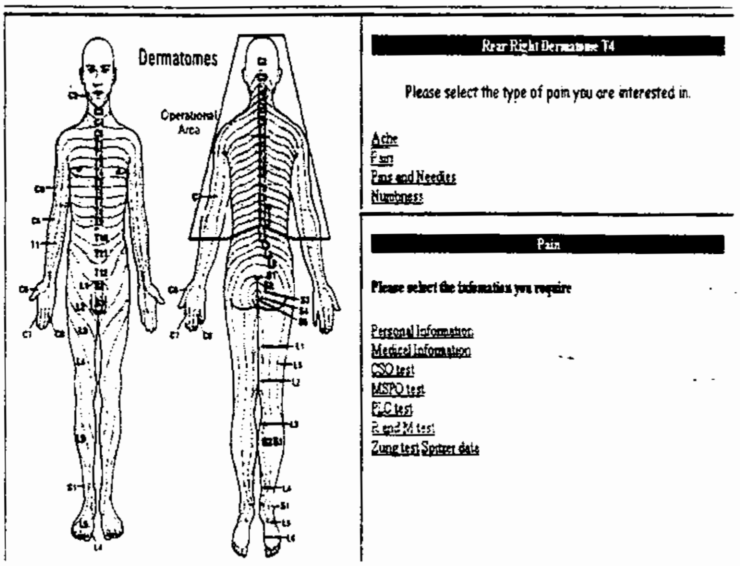


Figure 1: Pain drawing.



Rear Right Dermatome T4

Please select the type of pain you are interested in.

Ache
Pain
Pins and Needles
Numbness

Pain

Please select the information you require

Personal Information
Medical Information
CSQ test
MSPQ test
PLC test
R and M test
Zung test Spitzer data

Figure 2: Back pain data management in BPDRoPS.

QUALITY OF PERCEPTION

The concept of Quality of Service (QoS) in distributed multimedia systems is indelibly associated with the provision of an acceptable level of application performance. Ultimately this performance is itself dependent on:

1. the user's experience with the multimedia presentation which we define as Quality of Perception (QoP). QoP has two main components: a user's ability to analyse, synthesise and assimilate the informational content of multimedia applications, as well as his/her subjective satisfaction with the quality of such applications.
2. the service provided by the underlying network.

Whilst the focus in the telecommunications community has rested on the latter, it is our belief that it is indeed the former measure of quality which needs to be concentrated on in order for e-health applications to proliferate and gain increased acceptance in the medical community. Previous work on QoP [10],[11], based on extensive user tests, has shown that technical-oriented QoS must also be specified in terms of perception, understanding and absorption of content - Quality of Perception in short - if multimedia presentations are to be truly effective. Thus, for example, users have difficulty in absorbing audio, visual and textual information concurrently. In a multimedia based e-health environment (such as a remote video-based diagnostic system), if the user perceives problems with the presentation (such as synchronisation problems between different component media), users will disregard them and focus on the contextually important medium. This implies that critical and important messages in a multimedia presentation should be delivered in only one type of medium, or, if delivered concurrently, should be done so with maximal possible quality.

A FRAMEWORK FOR QoP ADAPTATION

Distributed guaranteed services need to incorporate capabilities for responding to QoP and QoS changes originating from the user/applications or the system/network. To achieve these changes, the networked multimedia system will require fast renegotiation protocols and adaptive mechanisms. The renegotiation protocols will rely on dependable and simple monitoring and recognition algorithms to detect requests for QoS changes or system degradations. The envisioned adaptive mechanisms should include update mechanisms for resource allocation in response to detection of system degradation. The challenge will be to make monitoring, detection and adaptation mechanisms efficient and fast.

The BPDRoPS tool is based on our previous work on the *Dynamically Reconfigurable Stacks Project* (DRoPS), which provides an infrastructure for the implementation and operation of multiple adaptable protocols [12]. The core architecture is embedded within the Linux operating system, is accessible through standard interfaces, such as *sockets* and the UNIX *ioctl* (I/O control) system calls, has direct access to network devices and benefits from a protected, multiprogramming environment. The architecture allows additional QoS maintenance techniques, such as flow shaping (to smooth out bursts in traffic), at the user or interface level, and transmission queue scheduling, at the device queue level.

BPDRoPS-based communication protocols are composed of fundamental mechanisms, called *microprotocols*, which perform arbitrary protocol processing operations. The complexity of processing performed by a microprotocol is not defined by BPDRoPS and may range from a simple protocol function, such as a checksum, to a complex layer of a protocol stack, such as TCP. In addition, protocol mechanisms encapsulated within a microprotocol may be implemented in hardware or software. If appropriate hardware is available, the microprotocol merely acts as a wrapper, calling the relevant hardware function. Microprotocols are encapsulated in loadable modules, allowing code to be dynamically loaded into a running operating system and executed without the need to recompile a new kernel. Each such microprotocol can be implemented via a number of adaptable functions. For instance, an acknowledgement protocol can be implemented either as an Idle Repeat Request or a Per Message Acknowledgement Scheme.

Whilst a protocol defines the structure and resources available for constructing a communication system, a *protocol stack* defines a unique instantiation assigned to a particular connection. In terms of microprotocols, a protocol stack is an ordered set combined to form a functional communication system. Each connection is assigned a protocol stack for its sole use, the configuration of which may vary according to the characteristics of the particular connection. Using this model, individual flows within individual sessions may be uniquely configured to provide an appropriate service. Thus, a connection between video client and server applications may use a semantically strong protocol for commands and a relatively weak one for bulk transfer of relatively loss tolerant graphical data

CONSTRUCTION OF COMMUNICATION PROTOCOLS IN BPDRoPS

In our work, transmission of back pain data incorporates intelligent decision-making mechanisms constructed using the Analytic Hierarchy Process (AHP), which is one of the most popular methods of Multi-criteria Decision-Making (MCDM). To this end, we have applied Saaty's (AHP) formalism [13] to obtain a method which, from combined user-, application- and network-level requirements, ultimately results in a protocol configuration specifically tailored for the respective user-needs. Our approach links perceptual considerations with low-level technical parameters, taking into account the end-user. Thus, within the QoP framework presented above, a BPDRoPS session can be characterised by the relative importance of the video (V), audio (A) and textual (T) components as conveyors of information. This is in accordance with experimental QoP results obtained which emphasise that multimedia QoP varies with: the number of media flows, the type of medium, the type of application, and the relative importance of each medium in the context of the application. Moreover, five network level QoS parameters have been considered in our model: bit error (*BER*), segment loss (*SL*), segment order (*SO*), delay (*DEL*) and jitter (*JIT*). Together with the V, A, T parameters these constitute, in AHP terminology, the *criteria* on the basis of which an appropriate tailored communication protocol is constructed. In BPDRoPS, the functionality of this protocol is realised through a number of 9 microprotocols, spanning 4 broad functionality classes [10],[11].

Saaty's methodology results in the construction of nine matrices. Eight of these matrices give the relative importance of the various microprotocols (*alternatives*, in Saaty's vocabulary) with respect to the criteria identified in our model, while the last of these matrices details *pairwise comparisons* between the criteria themselves.

$$A_{BER} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1/3 & 1/7 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1/3 & 1/7 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1/3 & 1/7 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1/3 & 1/7 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1/3 & 1/7 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1/3 & 1/7 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1/3 & 1/7 \\ 3 & 3 & 3 & 3 & 3 & 3 & 3 & 1 & 1/5 \\ 7 & 7 & 7 & 7 & 7 & 7 & 7 & 5 & 1 \end{bmatrix} \quad (1)$$

The formulation of the nine matrices has been based on QoP results described in [10]. An example of one of the former type of matrices, i.e. of the different alternatives with respect to one of the criteria (bit error rate in this case) is given in Relation (1). Each entry a_{ij} of the matrix A_{BER} defines the numerical judgement in comparing criterion to criterion.

Psychological experiments have shown that individuals cannot simultaneously compare more than 7 objects (±2) [14]. Thus, usually, pairwise comparisons are quantified by using a scale of nine grades, which describe the relative importance of the criteria [13]. If a_{ij} is a point on this nine-point scale, i.e. $a_{ij} \in \{1,2,3,...,8,9\}$, then $a_{ji} = 1/a_{ij}$ also holds [13]. For example, in Relation (1) the judgement *"microprotocol A is equally important as microprotocol B with respect to BER"* corresponds to a weighting of $a_{ij} = 1$, while the judgement *"microprotocol A is absolutely more important than microprotocol B"* would correspond to a value of $a_{ij} = 9$. Intermediate terms can also be assigned when compromise is needed between two adjacent characterisations. Note that in Relation (1), the considered microprotocols are, in order {*no sequence control, strong sequence control, no flow control, window-based flow control, IRQ, PM-ARQ, no checksum algorithm, block checking, full Cyclic Redundancy Check*}. For example, as far as bit error rate is concerned (see Relation (1)), the only microprotocols that have an impact upon it are the checksum algorithms. The strongest of these methods, the full CRC, has the highest weight (a value of $a_{ij} = 7$) in comparison with all the others, while a relatively weak block checking algorithm is considered to be *moderately more important* (a_{8j}=3, j=1,2,...,7) than microprotocols from other functionality classes.

While all the previous eight matrices considered have a constant form, the matrix of each criterion with respect to all the other criteria named C, shown below, is the only one whose values may fluctuate as a result of changes in the operating environment, as well as a consequence of changes in *user preferences* and *perceptions*. Relation (2) provides an instance of this matrix used in our model; the respective criteria are, in order, *BER, SO, SL, DEL, JIT, V, A,* and *T*.

$$C = \left[\begin{array}{ccccc|ccc} 1 & 1/2 & 1 & 1/4 & 1 & 2 & 4 & 4 \\ 2 & 1 & 1 & 1/4 & 1/3 & 5 & 4 & 5 \\ 1 & 1 & 1 & 1/3 & 1/2 & 4 & 6 & 4 \\ 4 & 4 & 3 & 1 & 5 & 6 & 7 & 6 \\ 1 & 3 & 2 & 1/5 & 1 & 4 & 6 & 6 \\ \hline 1/2 & 1/5 & 1/4 & 1/6 & 1/4 & 1 & 1/2 & 3 \\ 1/4 & 1/4 & 1/6 & 1/7 & 1/6 & 2 & 1 & 3 \\ 1/4 & 1/5 & 1/4 & 1/6 & 1/6 & 1/3 & 1/3 & 1 \end{array}\right] \tag{2}$$

An average user, though, would have difficulty in *a priori* judgement of varying technical parameters such as delay, jitter, error and loss rates on highly subjective attributes such as perception, understanding and satisfaction. Whilst this is true for QoS attributes at the level of the transport service, users are better able to quantify their requirements in terms of more abstract characteristics like the prioritisation of core multimedia components such as *V*, *A* and *T*. The matrix C reflects this situation and could conceptually be split-up into 4 sub-matrices, which are:

- A 5×5 matrix, in the upper left part of the matrix in Relation (2), giving the relative importance of the *BER, SO, SL, DEL* and *JIT* criteria with respect to one another. This matrix changes dynamically during the course of the transmission of a multimedia clip. An example illustrating this behaviour is presented below.
- A 3×3 matrix, located in the bottom right of the matrix given by (2). Here, user input can reflect personal choices of the relative importance of the video, audio and textual components in the context of the application, as well as a relative characterisation of the dynamism of the multimedia clip. Whilst, these values can be changed dynamically depending on the visualised scene, *a priori* values in this case could reflect the result of user- consultations, such as those conducted in our original QoP experiment [10].
- A 5×3 matrix and a 3×5 (one of which is the transpose of the other) which reflect designer choices of the relative importance of the five QoS parameters considered on *V*, *A*, and *T*. The elements of these matrices remain fixed at run-time, and, in our particular case, reflect the results of our previous work on QoP [10].

Following the AHP, the weights w_i, i=1,...8 denoting the relative importance of each criterion i among the p criteria (p=8) are evaluated using the formula:

$$w_i = \frac{\left(\prod_{j=1}^{p} a_{ij}\right)^{1/p}}{\sum_{i=1}^{p}\left(\prod_{j=1}^{p} a_{ij}\right)^{1/p}} \quad i = 1,2,...,p \tag{3}$$

and a higher priority setting corresponds to a greater importance. Pairs among alternatives are also compared with respect to the ith criterion and then a weight $w_{j,i}$, which denotes how preferable is the alternative j with respect to the criterion i, is derived. As previously, there is a total of p pairwise comparisons in the matrix and weights are calculated following Relation (3). The weighted sum model, [15], is used to find the preference of an alternative j with respect to all criteria simultaneously; preference is defined by P_j and denotes the overall priority, or weight, of action j:

$$P_j = \sum_{i=1}^{r} w_i \cdot w_{j,i} \tag{4}$$

In the maximisation case, the best alternative is the one that possesses the highest priority value among all others.

Due to the dynamic nature of our problem, where the technical information and the perceptual information could form inconsistent judgement matrices, the need to apply a weight determination technique suitable to handle inconsistencies was indispensable. Therefore, the Fuzzy Programming Method (FPM), which is a method capable to solve even high inconsistent matrices, was used. The FPM proposed by Mikhailov and Singh [16] is an approach capable to handle inconsistent pairwise comparison judgement matrices, where the judgements can be expressed either as crisp, intervals or fuzzy numbers. Each reciprocal pairwise comparison matrix, $A=[a_{ij}]\in \Re^{n\times n}$, can be represented as a system of $m = n\times(n-1)$ linear equalities:

$$\mathrm{R}\times w = 0, \tag{5}$$

where n is the number of elements compared, w is the vector of weights and $\mathrm{R}\in\Re^{m\times n}$. For the inconsistent cases, the method is finding a solution that approximately satisfies Equation (5), i.e. $\mathrm{R}\times w \approx 0$.

One of the most important advantages of the FPM is that the prioritisation problem is reduced to a fuzzy programming problem that can be easily formulated and solved as a standard linear program:

$$\begin{aligned} &\text{Obj.: } \max \lambda \\ &\text{s.t. } \lambda \mathrm{d}_k + \mathrm{R}_k w \le \mathrm{d}_k,\ k = 1,\dots,m,\ 1 \ge \lambda \ge 0 \\ &\sum_{i=1}^{n} w_i = 1\ ,\ \ w_i > 0,\ i = 1,\dots,n \end{aligned} \tag{6}$$

where the values of the tolerance parameters d_k represent the admissible interval of approximate satisfaction of the crisp inequalities $\mathrm{R}_k w < 0$. For the practical implementation of the FPM, it is reasonable all these parameters, d_k, to be set equal [16],[17].

After deriving the underlying weights from the comparison matrices through the FPM technique, the local weights are synthesised following the weight-sum model. The overall value V_j of each j^{th} alternative, A_j, is expressed as:

$$V_j = \sum_{i} w_i r_{ij} \tag{7}$$

where w_i is the weight assigned to the i^{th} criterion and r_{ij} is the relative score of the alternative j on criterion i. Obviously, the alternative with the maximum overall value V_j will be chosen.

The diagram of our architecture is given in Figure 3 and shows how both monitored QoS and user choices impact on the construction of the judgement matrix, which serves as the basis for the AHP to suggest a suitable protocol stack configuration under BPDRoPS ensuring that user QoP is maintained at an optimum level [11]. This is in contrast to traditional legacy protocols stack such as TCP IP and UDP, which make no allowance for user -related considerations in their functionality

APPLICATION EXAMPLE

In this section, we present experiments illustrating the ability of our approach to select appropriate micro-protocols and construct a suitably-tailored protocol stack depending on the prevailing operating network environment.

Priorities	micro1	micro2	micro3	micro4	micro5	micro6	micro7	micro8	micro9
Initial	0.0982	**0.1684**	0.0922	0.1361	0.0847	0.1279	0.0868	0.0674	0.1373
Updated	**0.1262**	0.1259	0.1154	0.1186	0.0819	0.1095	**0.1337**	0.0739	0.1251

Table I: Overall values of the alternative microprotocols for the experiment

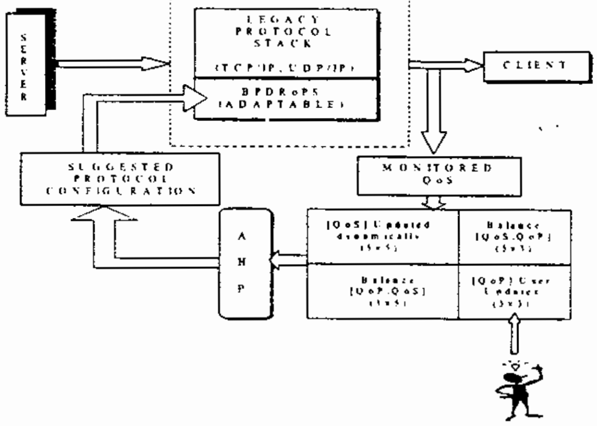


Figure 3. AHP-based architecture for QoP management

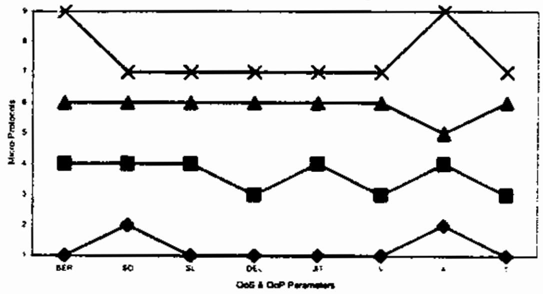


Figure 4: Resulting BPDRoPs protocol stack when QoS and QoP parameters are, in turn, of primary importance.

In Table I, our methodology has been applied to a situation where BPDRoPs is experiencing protracted delays due to network congestion. The notation adopted in Table I is as follows: *no sequence control* (micro1), *strong sequence control* (micro2), *no flow control* (micro3), *window-based flow control* (micro4), *IRQ* (micro5), *PM-ARQ* (micro6), *no checksum algorithm* (micro7), *block checking* (micro8), *full Cyclic Redundancy Check* (micro9).

As a result of a delay-intolerant audio transmission being subjected to a period of high network delays, the upper left sub-matrix of Relation 2 can reflect this situation by changing the numerical judgements to reflect a more radical bias in favour of the delay component. In Table 1 we can see that the priorities of the different microprotocols obtained through our approach change from the initial configuration, biased towards micro2 (an overall value of 0.1684 was assigned to that microprotocol initially), to an updated one in which micro7 and micro1 are top of the priority ordering. This means that the priority ordering of the microprotocols would change to one which favours microprotocols that do not lead to extra delays, as one would expect. In our case, these are represented by micro1 and micro7.

In Figure 4, we show the resulting protocol stack which is constructed using our approach in the BPDRoPS framework, when each of the QoP and QoS parameters becomes, in turn, of primary importance. This situation is not farfetched and can easily arise in real-life situations, particularly when component parts of networks fail or malfunction. Thus, for instance, if a link between two routers goes down, then connections using that link will experience a high degree of segment loss; alternatively, if there is a fault in router hardware, then connections involving that router might, for instance, experience high bit error rates. Thus, in the case where segment loss (SL) is of primary importance then, as can be seen from Figure 4, the BPDRoPS protocol stack is made up of micro1, micro4, micro6 and micro7. Whilst the choice of micro6 is to be expected, as it is the only microprotocol in the BPDRoPS framework explicitly able to handle losses, the choice of micro4 highlights the importance of flow control for segment losses, which would prevent, for instance, buffer overflows and the resulting loss of data. Otherwise, the choice of micro1 and micro7 reflect the streamlined functionality of the protocol stack, as these microprotocols, by not acting on sequence control and bit errors, respectively, reduce computational overhead.

Similar observations apply in the case when QoP parameters are of primary importance. Accordingly, all media component of multimedia presentations are tolerant to bit errors, except audio. Thus, the case when audio is considered of primary importance is the only one in which the resulting protocol stack includes in its configuration micro9, the most suited microprotocol to handle bit errors. The fact that most distributed multimedia applications have real-time constraints as well as being tolerant to bit errors, is reflected in the choice of the "no-frills" micro7 in all other cases, for this type of functionality. The delay-intolerant nature of distributed multimedia applications is also reflected in the choices of micro1 and micro3 in the suggested protocol stacks when video and text are of primary importance. The choice of micro6 for these two scenarios reflects, however, the importance of not losing segments of information, particularly in the case of compressed media, as any loss of information would propagate through subsequent media units, bearing in mind the widespread exploitation of differential characteristics in compression.

CONCLUDING REMARKS

Distributed multimedia e-health applications have a set of task-specific requirements which must be taken into account if effective use is to be made of the limited resources provided by public telecommunication networks. In this paper, we have addressed the problem of bridging the application-network gap, in a remote collaborative environment for back pain treatment, from a multi-attribute decision-making perspective. To this end, we have used the Analytic Hierarchy Process to integrate Quality of Perception- related requirements with the more technical characterisation of Quality of Service. We have shown how our framework is capable of suggesting appropriately tailored transmission protocols, by incorporating human-perceptual requirements in the remote delivery of e-health solutions.

REFERENCES

[1] Boucher, A, 1999, "The Prevalence of Backpain in Great Britain in 1998", Department of Health. Available at http://www.doh.gov.uk/public/backpain.htm.

[2] Frank, A. and De Souza, L.H., 2000, "Conservative Management of Low Backpain", International Journal of Clinical Practice, Vol. 55, No. 1, pp. 21 - 31.

[3] Jefferson, J.R. and McGrath, P.J., 1996, "Backpain and peripheral joint pain in an industrial setting", Arch Phys Med Rehabil , Vol. 77, pp. 385-390.

[4] Balague F., Troussier, B. and Salminen, J.J., 1999, "Non-specific low back pain in children and adolescents: risk factors", European Spine Journal, Vol. 8, pp. 429–438.

[5] Papageorgiou, A.C., Croft, P.R., Ferry, S., Jayson, M.I.V. and Silman, A.J., 1995, "Estimating the Prevalence of Low Backpain in the General Population: Evidence from the South Manchester Backpain Survey", Spine, Vol. 20, No. 17, pp. 1889-1894.

[6] Ginzburg, B.M., Merskey, H. and Lau, C.L.,1988, "The Relationship between Pain Drawings and the Psychological State", Pain, Vol. 35, pp. 141-145.

[7] Parker, H., Wood, P.L.R. and Main, C.J., 1995, "The Use of the Pain Drawing as a Screening Measure to Predict Psychological Distress in Chronic Low Backpain", Spine, Vol. 20, No. 2, pp. 236-243.

[8] Palmer, H., 1949, "Pain charts: A Description of a Technique whereby Functional Pain may be Diagnosed from Organic Pain", New Zealand Medical Journal, Vol. 48, No. 264, pp. 187-213.

[9] Ohnmeiss, D.D., Vanharanta, H. and Guyer, R.D., 1995, "The Association Between Pain Drawings and Computed Tomographic/Discographic Pain Responses", Spine, Vol. 20, No. 6, pp. 729-733.

[10]Ghinea G. and Thomas J.P., 1998, "QoS Impact on User Perception and Understanding of Multimedia Video Clips", ACM Multimedia '98, Bristol, UK., pp. 49-54.

[11]Ghinea G. and Magoulas G.D., 2001, "Quality of Service for Perceptual Considerations: An Integrated Perspective", IEEE International Conference on Multimedia and Expo (ICME2001), Tokyo, Japan, pp. 752-755.

[12]Fish R.S. and Loader R.J., 1998, "A kernel based adaptable protocol architecture", 24th Euromicro Conference, Vasteras, Sweeden, pp. 1029-1036

[13]Saaty, T.L 1977, "A scaling method for priorities in hierarchical structures", J Math Psych., Vol. 15, pp. 234-281.

[14]Miller G.A., 1956, "The magical number seven plus or minus two: some limits on our capacity for processing information", Psychological Review, Vol. 63, pp. 81-97.

[15]Triantaphyllou, E. and Chi-Tun Lin, 1996, "Development and Evaluation of Five Fuzzy Multiattribute Decision-Making Methods", International Journal of Approximate Reasoning, Vol. 14, pp. 281-310.

[16]Mikhailov, L. and Singh, M.G., 1999, "Fuzzy Assessment of Priorities with Application to the Competitive Bidding", Journal of Decision Systems, Vol. 8, No. 1, pp. 11-28.

[17]Mikhailov, L. and Singh, M.G., 1999, "Comparison Analysis of Methods for Deriving Priorities in the Analytic Hierarchy Process", IEEE International Conference on Systems, Man and Cybernetics, Tokyo, Japan, pp. 1037-1042.

The VPIS System: A New Approach to Healthcare logistics

Konstantinos K. Danas [1,2],A.V. Roudsari [2], Panayiotis H. Ketikidis [3]
1. Department of Mathematics and Computing,
ACT - The American College of Thessaloniki, Greece
cdanas@ac.anatolia.edu.gr
2. MIM Centre, City University, London,U.K.
a.v.roudsari@city.ac.uk
3. CITY Liberal Studies
Affiliated Institution of the University of Sheffield
Thessaloniki, Greece
ketikidis@city.academic.gr

Abstract: The escalated cost of healthcare has been under consideration and hospital management has devoted many efforts in order to control these costs. During a recent study, it was concluded that the costs produced by the inefficiencies of the hospital pharmacy inventory management could be reduced using the techniques that contemporary logistics and information systems provide. The Just In Time concept that has been successfully applied in other industries is not considered as a procurement technique for the hospital pharmacy due to irregularities that the demand for specific medicines may produce in cases of emergency. The aim of this study is to demonstrate the use of a software model that could support the concept of just in time in healthcare logistics. This model supports the virtual pharmacy inventory system that considers the inventories of the hospital pharmacies in the same geographical area as one virtual entity providing the way that hospitals support each other in pharmaceutical supplies. This model is based on the intelligent agent technology and the Internet is used as a communication medium.

Keywords: Healthcare Information Systems, Just in Time, hospital pharmacy inventory, unexpected demand support, pharmaceuticals supplies.

1. INTRODUCTION

The advances of medical and biological sciences have greatly improved the quality of healthcare but on the other hand have increased the costs of healthcare. Healthcare officials and managers are trying hard to find ways of reducing costs without affecting the quality of healthcare services [1],[2],[3]. The science of logistics that provides models for the optimization of supply chain systems must be seriously considered in order to eliminate the inefficiencies of the healthcare supply chain that lead to hidden cost increases. The use of information and communication systems tools within the supply chain could provide solutions that improve the way the healthcare supply chain is managed as such solutions have already been used in other industries. One such model of supply chain management is Just In Time (JIT), which has not been applied in the healthcare industry as hesitations about the security of supplies in emergency situations has concealed its effectiveness [1].

The focus of governmental projects in order to improve the healthcare supply chain is based on the supplies of each hospital individually and the way the hospital interacts with its suppliers. Logistics information systems that are based on Electronic Data Interchange (EDI) and bar coding systems are proposed to eliminate the supply chain inefficiencies

Intelligent e-Health Applications in Medicine
Special Session Proc EUNITE 2002 Albufeira Portugal Sept 19-21. 2002
Published by the University of the Aegean

and cut costs [1],[4];[5]. Another possible solution that could improve this proposed solution is to consider the hospitals of the same geographical area as one entity regarding their inventory and build an environment within witch these hospitals are able to cooperate and exchange medicines forming a virtual hospital pharmacy inventory that could store the emergency medicines required for their efficient operation providing ground at the same time for the application of a JIT model in the healthcare industry.

The aim of this study is to present a software model that could support the concept of just in time in healthcare logistics and demonstrate its use. This model supports the virtual pharmacy inventory system that considers the inventories of the hospital pharmacies in the same geographical area as one virtual entity providing the way that hospitals support each other in pharmaceutical supplies. The objectives are to:

- Investigate the solutions that are already proposed
- Explain the need of a new system
- Propose the new system
- Investigate the technologies that could support such a system
- Provide an outline of the system

2. HEALTHCARE LOGISTICS: PROPOSED SOLUTIONS

The logistics models that have been effectively applied in other industries, like JIT, supply chain management and partnerships using logistics information systems (LIS) for the integration of the supply chain, have attracted the attention of the public healthcare sector across the world and several projects have been produced aiming at the application of contemporary logistics in the healthcare industry. According to Jarrett, these projects incorporate the vision of having '*the right product in the right place at the right time in the most cost effective manner to serve efficiently the healthcare needs of the end consumer*'[1]. The projects that are proposed by governing bodies around the world are introduced, their limitations are outlined and a new system is proposed.

The first project under study is the EHCR, Electronic Health Care Requisitioning, which was initiated by the Clinton administration in the USA. Profitable and non-profit organizations participated in the project as well as consultants and other specialists that were working on the design, development and implementation of models. The Canadian government adopted the directives of the EHCR and applied them to some Canadian hospitals. The Australian government produced their own directives based on the EHCR, known as the PECC project and applied it in some model hospitals. The European Union has another project under development by a consortium of Universities around Europe, which is called DRIVE project also based on the EHCR directives [6],[7],[8].

The method of stockless hospital inventory eliminates the storage of medicines in the central storeroom and spreads the medicines according to the needs of each clinic in the hospital. The objective of the concept is to cut material handling procedures and the time spent in ordering systems. The suppliers make frequent deliveries in separate lots for each clinic of the hospital. Automated ordering systems as the one suggested by Danas and Ketikidis [11],[12] are required with the aid of EDI and bar coding, in order to minimize the time and effort spent on bureaucratic procedures. Many hospitals have successfully implemented the above concept like the University of Michigan Hospitals, Escambia Hights, Tempa General etc [3]. Baxter Healthcare Corporation, offering distribution management systems that help hospitals by holding their stock until it is needed and then deliver directly to the appropriate clinic [9], proves the fact that suppliers are willing to adapt to the new way of doing business with the hospitals.

To summarize, the current approach to the JIT concept in the healthcare industry is based on two axes:

i) The integration of the information system within the hospital and among the hospital and its suppliers based on the use of bar coding and EDI, and

ii) The elimination of the central hospital pharmacy as the stockless inventory system model

The current approaches provide system models aiming to reduce the costs of the logistics processes of the hospital supplies including the hospital pharmacy and to provide models for the JIT concept implementation.

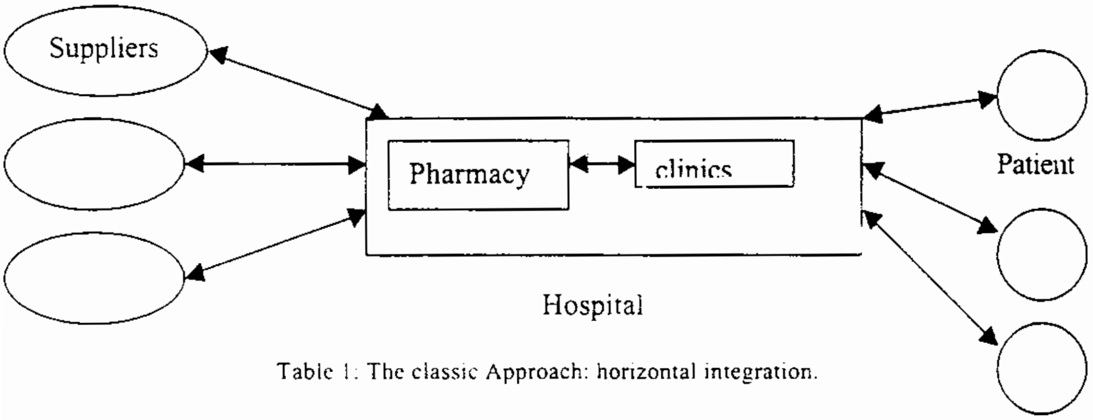


Table 1: The classic Approach: horizontal integration.

3. THE NEED FOR A NEW SYSTEM

The focus of the projects under consideration is on the integration of the processes within the supply chain especially from the hospital to the hospital's suppliers in order to reduce the administrative costs. The automation of the procurement procedure could lead to the reduction of inventory levels since e-procurement applications improve the transaction speed and accuracy. The control of the flow of medicines within the hospital could also lead to further reduction of inventory levels, as the number of medicines that are lost or wasted would be reduced. A JIT approach in the healthcare industry could be based in these initiatives given the fact that transaction accuracy and speed are improved. Given the fact that the stockless inventory system has been successfully applied in American hospitals, a combination of the initiatives proposed by these projects and the stockless inventory system could lead to the improvement of the JIT approach to the healthcare industry given that the pharmaceuticals could be ordered, checked by the pharmacist and delivered directly to the clinic that requires them. This way the central pharmacy inventory is eliminated and further cost reductions could be achieved.

A JIT system in the hospital pharmacy could be based on the combined implementation of the stockless inventory system and the proposed initiatives recommended by the projects under study. The limitation of such a combination is that the cooperation of the hospitals in the same geographical area is not considered. There are benefits that could be achieved if this cooperation is considered. These benefits include the circulation of pharmaceuticals within the hospital network in cases that a given medicine is not needed in a clinic that ordered it for a specific patient who for one reason or the other did not consumed it. In such a case this medicine would be held in stock in the clinic occupying inventory space and may be wasted because it is expired. Another benefit could be the support of the network regarding the pharmaceutical stock that each hospital must keep in order to face emergency situations of unexpected demand. The JIT concept could then be applied with greater security and better results [10].

A new approach must then be considered based on the cooperation of hospitals in the same geographical area in order to further improve the results of the project initiatives. This network must be based on an infrastructure that will support the cooperation of hospitals participating in the network. The EDI technology that is limited to the exchange of messages between computer systems [13] is not sufficient to support the hospital network. An information system based on intelligent agents that have the ability to act between computer systems could be the proper infrastructure for the hospital network.

A new system is therefore needed, based on the intelligent agent technology in order to form a virtual inventory of medicines within the hospital network. This virtual inventory could support the need for pharmaceuticals while at the same time follow a JIT approach in the procurement of pharmaceuticals directly to each clinic within the hospital eliminating the central hospital pharmacy inventory.

4. INTRODUCTION TO THE VPIS

In order to further improve the current system models and to reduce the fears of implementing a JIT model in healthcare, a new approach to further cost reduction in the logistical costs could be based on the cooperation between hospitals of the same geographical area in order to be able to exchange information and pharmaceuticals among them thus further reducing their inventory stock and improving the JIT system. Such a system could be based on EDI as current models suggest but such a solution would be much complicated if not impossible due to the fact that EDI is limited on exchanging messages and facilitating transactions. An information system based on intelligent agents that would facilitate and handle the exchange of information could support the evolvement of such a system model since the agent technology is not only limited to messages or information exchange but provide a means for certain actions that the agent can perform away from its base. This new system would be based on a virtual common pharmacy inventory that will support all the hospitals in the same geographical area.

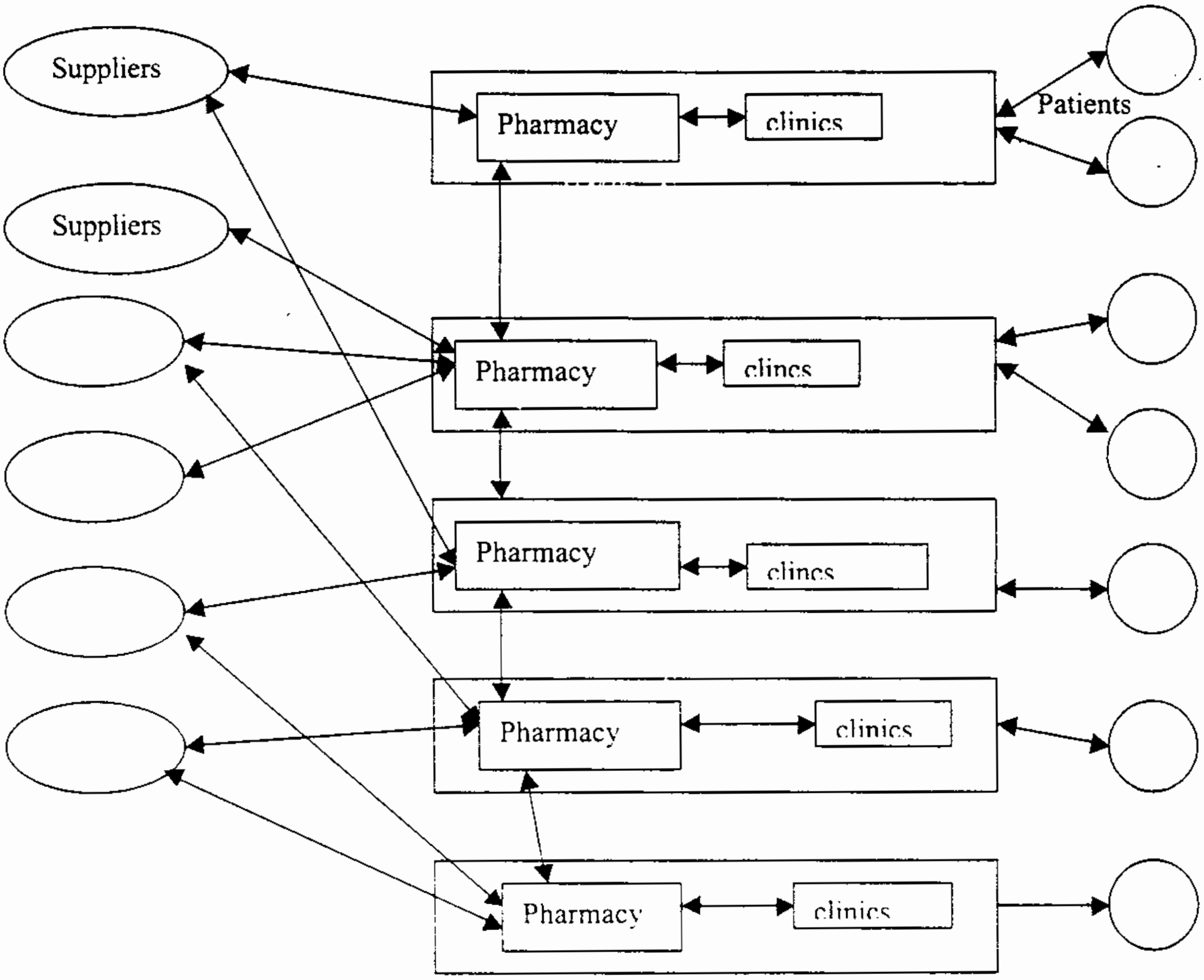


Table 2. Horizontal and vertical integration

5. A DESCRIPTION OF THE SYSTEM

5.1 FUNCTIONALITY

VPIS can be described as a web based procurement system that provides a platform for the cooperation among hospitals in the same geographical area. The hospital pharmacist will be the main user of the system even though the system provides the ability to doctors to enter the pharmaceuticals they require for their clinics. The pharmacist will be able to check the daily orders and then select to purchase the required material from a supplier or get them from another hospital in the area if they are in offer. The system will automatically check the stock of each clinic and then according to a rule based DSS will post them for offer to the central offerings blackboard system. Orders will be transmitted to the appropriate supplier and exchanges to the appropriate hospital.

5.2 INTEGRATION WITH THE HOSPITAL LOGISTICS INFORMATION SYSTEM (LIS)

Integration of the VPIS and the hospital LIS involves the cooperation of the two systems in order to exchange information regarding the stock that is kept in the system. The hospital information systems are not all identical. Each hospital has its own information system, which is either an LIS or a legacy information system. VPIS is based on an LIS in order to be able to keep a database updated with real time information. The database of each hospital would not have the same structure. VPIS performs queries in these databases and collects information regarding the stock of pharmaceuticals each hospital has on hand.

5.2.1 OBJECTIVES

The objective of the integration is to query each hospital database in order to get information regarding the status of pharmaceuticals stock collecting information like expiry dates.

5.2.2 SOLUTIONS

The integration could be achieved by placing a static agent in each local database in the network that would be able to query the local database. This means that each static agent knows the structure of the database that is attached to. The agent would accept messages from the system with specific information about queries would then perform the query and send back to the system the results.

5.3 OFFERINGS TABLE – THE BLACKBOARD

The offerings table, the blackboard of the system, hold information about the pharmaceuticals that are for offer inside the network. These pharmaceuticals include medicines that are about to expire that could either be returned to the suppliers or must be consumed before they expire, medicines that were ordered for a particular therapy and for some reason were not consumed and medicines that must be kept in stock for emergency situations and could be used by any hospital in the network.

5.3.1 OBJECTIVES

The blackboard is used to post information about stock that must be circulated in the network. Static agents send query information to an intelligent agent that is responsible to update the blackboard based on specific rules.

5.3.2 SOLUTIONS

The blackboard is actually a database holding data about the pharmaceuticals for circulation inside the network. An intelligent agent is responsible to update the blackboard. Other agents would perform queries in this database to locate medicines that must be circulated in the network or are used in emergency situations.

5.4 BLACKBOARD UPDATE

The system that updates the blackboard consists of an intelligent agent and a database that contains the knowledge to be used by this agent in order to decide whether to place a specific medicine in the blackboard.

5.4.1 OBJECTIVE

The objective of this module is to place the correct medicines for circulation in the network.

5.4.2 SOLUTIONS

The objective of the module could be achieved by an intelligent agent that will accept as input the results of queries from the static agents in hospital databases, it would then perform a checking using a rule based system with a database constructed specifically to act as a knowledge base for this agent.

5.5 ORDER SCHEDULER

The order scheduler would check the schedule for programmed pharmaceutical support for a specific period of time ahead, would then check the blackboard and produce a list of medicines that could be redistributed inside the network. The module is also responsible to produce a list of scheduled supplies to be sent to suppliers in order to be ready to supply in time the required medicines.

5.5.1 OBJECTIVE

The objective of the module is to check the blackboard for required medicines to be circulated inside the network and then produce a requirements list to be sent to suppliers.

5.5.2 SOLUTIONS

An agent could check for matches between the requirements list for each hospital and the blackboard. It would then produce a list to be send using EDI to suppliers in order to get ready for the scheduled demand of the whole network.

5.6 EMERGENCY STOCK LOCATOR

The emergency stock locator would be able to locate if a required medicine that is needed is held in stock in the network of hospitals. This would provide greater security for the pharmaceuticals support of the system.

5.6.1 OBJECTIVE

This module has the responsibility to locate a specific medicine within the network. Each hospital database must be queried and results must be sent back to the system to decide where to get the medicine from if it exists in the network.

5.6.2 SOLUTIONS

The static agents in local databases would perform the query based on a message that they would receive from the intelligent agent responsible to locate the required medicine. A list of the hospitals in the network that have in stock the required medicine would then be produced.

5.7 DEMAND SIMULATION

The demand simulator would simulate demand at random times to check whether the system is able to support the specific demand. The demand simulator would simulate an event that a number of patients are transferred to one hospital of the system suffering a specific emergency situation like a food poisoning. The system would then check the virtual inventory if it is able to support the specific event.

5.7.1 OBJECTIVE

The objective of the module is to simulate a specific emergency event, produce a list of the required medication to support this event and then check if the system is able to support the pharmaceutical requirements.

5.7.2 SOLUTIONS

An intelligent agent could simulate emergency events at random times out of a specific set of possible emergency situations that a hospital may face. Each specific event would have a list of pharmaceuticals requirements that the agent must check if the system is able to support. The agent would send a query message to each static agent in the local databases and produce a report based on the results that it would receive.

5.8 STOCK LEVEL SCHEDULER

This module would have the responsibility to check each local database and produce a report of the pharmaceuticals that are below the safety stock. A list of requirements could then be produced and sent to suppliers to schedule forthcoming supplies.

5.8.1 OBJECTIVE

The objective of this module is to keep a safety stock of specific pharmaceuticals that the pharmacy must always keep. The system is then secured against out of stock incidents.

5.8.2 SOLUTIONS

The stock level scheduler agent would be based on a specific database of medicines that must be kept in a hospital pharmacy in a specified stock level. If a stock of such a medicine is less than the required level than an order must be sent to the supplier. The agent responsible for this task must send a query to each hospital static agent to perform the query and send back a list of medicines that are under the security stock level.

6. TECHNICAL DESCRIPTION

The VPIS system will be based on a group of mobile intelligent agents that will be able to search the databases of the participating hospitals and manage another database that will include all pharmaceuticals that are on offer within the system. A brief outline of all participating agents is given below:

IA01: **Search the local db and post offerings**: The IA01 agent has the task to search the local hospital database and send results to other agents.

IA02: **Order scheduler**: The order scheduler will check the schedule of the following week and will include automatically in the orders all the prescribed pharmaceuticals that are needed for scheduled operations.

IA03: **Emergency stock locator**: This agent will locate in the local databases of all participating hospital a medicine that might be required for a certain emergency.

IA04: **Demand Simulator**: This agent will simulate emergencies in order to check the system if it is able to provide the proper pharmaceutical support.

IA05: **Search offerings table**: This agent will search the offerings table and then provide decision support based on a rule-based system for the pharmacist in order to decide whether to buy from the supplier or get the medication from another hospital.

IA06: **Stock Level Scheduler**: The stock level scheduler will decide weather or not to replace a medicine in stock that is consumed. Many medicines need to be in stock in order to be able to support all emergency situations.

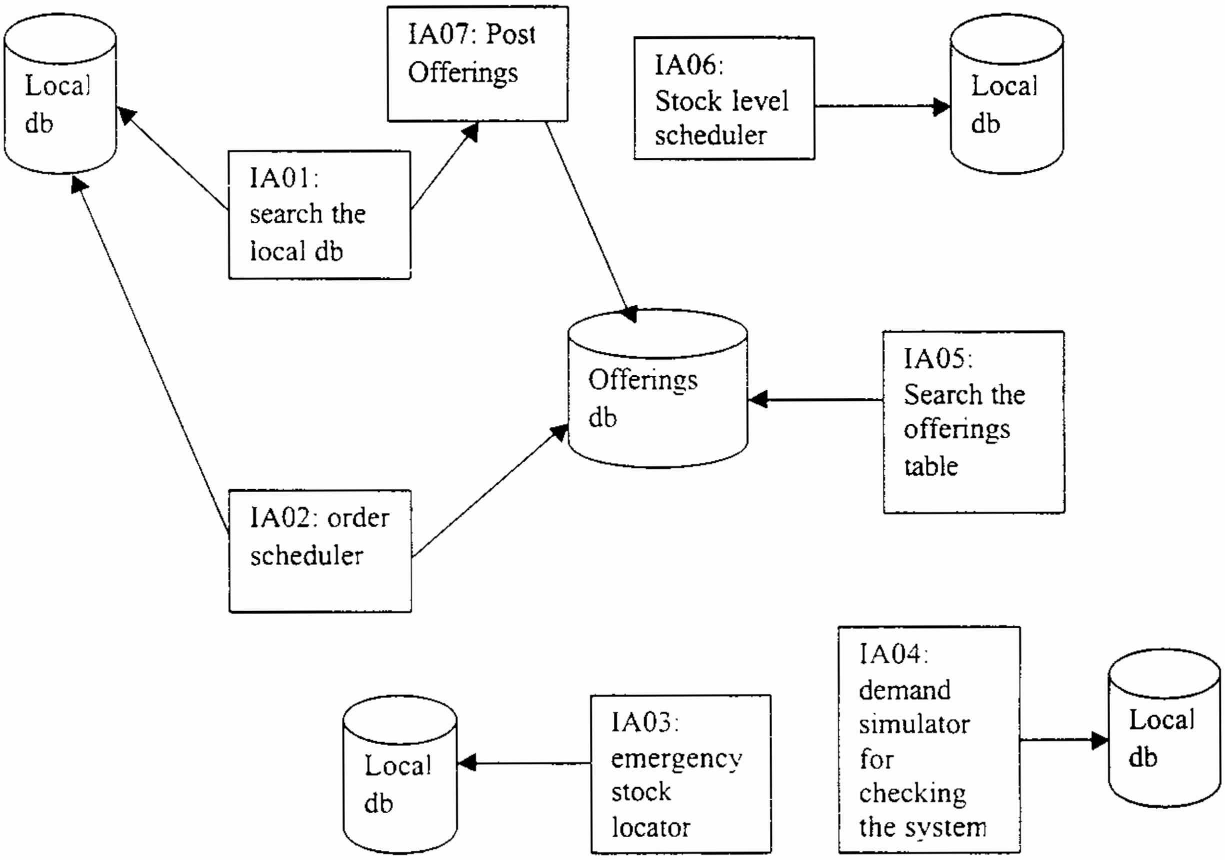


Table 3: The outline of the VPIS system

IA07:Post Offerings: This agent post the pharmaceuticals that are to be offered in the blackboard based on a rule-based system.

7. AN APPLICATION SCENARIO

The new JIT system model is based on the elimination of the central pharmacy inventory of the participating hospitals. More often supply orders that are based on scheduled operations are used and the VPIS information system is the backbone of the cooperation between hospitals and between hospitals and suppliers. The system does not attempt to change the process of ordering the pharmaceuticals or the persons responsible for the ordering process. The doctors at each clinic in the hospital can use the system to enter their orders. The hospital pharmacist then checks these orders for verification. The pharmacist then using the system can check if the required pharmaceuticals are in offer by other clinics of the hospitals in the network and then decide to get them from this list or order them from the appropriate supplier. In cases of emergency, the system can locate if the required pharmaceutical is in stock within the network. The pharmacists participating in the network must decide on the minimum level of stock of each medicine that is required in emergency cases and keep it in stock within the network. If a medicine that is required to be in stock is consumed then the system will produce an order for replacement. If a required medicine is near its expiration date then this medicine is also replaced. Pharmaceuticals are ordered based on a daily schedule that is supported by the system. A doctor must enter the scheduled operations and the required medication in the system for a certain time period and then the system will manage the pharmaceuticals orders. If a medicine that is ordered is not consumed and remains idle in the clinic then this medicine is offered for exchange within the network. Other medicines that offered for exchange are for example medicines that are near their expiration date so stock disposal is reduced within the hospital network.

8. CONCLUSIONS

The JIT approaches in the healthcare or other industries adopt models that are based on the horizontal integration of the supply chain. The healthcare industry in the public sector has the advantage that hospitals are not competing for profit. This advantage provides the ground for cooperation among hospitals in order to cut costs that will benefit the public healthcare sector since these amounts can be invested elsewhere for the benefit of the public health. In the governmental directions that are examined so far the approach of the vertical integration is not included. A vertical integration can provide the ground for the improvement of the JIT models that can be adopted in the healthcare sector reducing the fear of emergency stock outs since an emergency medication can be located among the hospitals of the same geographical area and be transported to the place that it is really needed. In order to support the JIT system model an information system is needed that will provide the tools for the cooperation of hospitals. This information system must be based on intelligent agents since the EDI approach is insufficient due to the limitation of EDI to exchange messages and not functionality between computers.

REFERENCES

[1] Jarrett, G. P. (1998) "Logistics in the health care industry". International Journal of Physical Distribution and Logistics Management Vol 28 No 9 pp 741-772. University Press.

[2] Whitson, D. (1997). "Applying just in time systems in health care", IIE Solutions, (USA), August, Vol 29 No 8 pp.33-38.

[3] Wilson, J.W. and Cunningham W.A. (1992) Stockless Inventory Systems for the Healthcare Provider: Three Successful Applications. Journal of Healthcare Marketing, Vol 12, No 2, p 39.

[4] Droge, C., Germain, R. (1998). "The just in time inventory effect: does it hold under different contextual, environmental, and organizational conditions?" Journal of Business Logistics, Vol 19, No 2.

[5] Germain R., Droge C.(1996) The Implications for Just-I-Time for Logistics Organization Management and Performance. Journal of Business Logistics, Vol 17 No 2 p 19.

[6] Healthcare Industry study on Efficient Consumer Response EHCR

[7] Blomberg G., Thawley C.,Blanchard P.(1998) PECC Project: NSW Central Coast Area, Health Service Scooping Study. Electronic Document [http://www.pecc.org.au/nswreport.html]. Site visited at 26/11/2001.

[8] DRIVE Project . www.cordis.lu/ist/projects/99-12040.htm

[9] Beier F. J. (1995) "The management of the supply chain for hospital pharmacies: A focus on inventory management practices" Journal of business Logistics. Oak Brook Vol 16 No 2 p153

[10] Danas,K., Roudsari, A.,V., Ketikidis, P. (2002). "A Virtual Hospital Pharmacy Inventory: An approach to support unexpected demand", International journal of Medical marketingVol 2 No 2 pp 125-129

[11] Danas, K., Ketikidis, P. (2000) "e-logistics: A framework for the hospital pharmacy", Eurolog 2000 Conference Proceedings, pp 114-121.

[12] Danas, K., Ketikidis, P.(2000). "Independent Trading Exchanges: a study in the healthcare industry".16th International Logistics Congress, Sole France, Paris, September 2000. Conference proceedings, section C-12 pp 1-8.

[13] Robeson J. and Copacino W. (editors) (1993) The Logistics Handbook. USA: Andersen Consulting, pp 698-757

Intelligent Alarms Integrated in a Multi-Agent Architecture for Diabetes Management

M. Elena Hernando, Enrique J. Gómez, Pedro Alvarez and Francisco del Pozo
Grupo de Bioingeniería y Telemedicina
E.T.S. Ingenieros de Telecomunicación. Universidad Politécnica de Madrid
Ciudad Universitaria s.n., E-28040 Madrid, Spain
Phone: +49-2408-94580, Fax: +49-2408-94582
email: elena@gbt.tfo.upm.es

ABSTRACT: This paper describes the development of an intelligent agent that interprets blood glucose monitoring data received by the m2dm telemedicine service for generating alarms when some deviations in the patient state are detected. The agent uses hybrid methods to generate data summaries and automatic alarms, such as rule-based techniques and Causal Probabilistic Networks (CPN) techniques. The rule-based analysis allows the detection of severe abnormalities using different time scales depending on the quality of the received information. The CPN analysis looks for deviations in the "insulin effectiveness" along a typical day using a physiological qualitative model. The advantages of an automatic response integrated in a telemedicine service are that it focuses doctors' and patients' attention on abnormal data and gives instantaneous feedback to patients reinforcing their education and motivation. The agent is being evaluated within the m2dm project at the Hospital Sant Pau of Barcelona.

KEYWORDS: Intelligent alarms, telemedicine, diabetes, causal probabilistic networks

INTRODUCTION

Diabetes Mellitus is a chronic disease where the process of care is complex and requires the patient to take an active role. The daily management of diabetes mellitus implies that the patient has to be able to react in a proper manner under any situation that can appear during his/her daily life. Telemedicine is an effective way to support patients' decisions providing them with a "supervised autonomy" [1] carried out remotely from the hospital. But telemedicine increases the physicians workload because it increases the amount of information they have to process and it demands a higher interaction with patients. It is crucial to complement telemedicine services with automatic data processing tools to focus doctors' attention on abnormal data optimising the time needed for data interpretation.

This paper describes the development and integration of an intelligent agent that interprets blood glucose monitoring data received by the a telemedicine service. The agent generates automatic summaries and alarms when some deviations in the patient state are detected.

The agent has been integrated in the m2dm project [2], which aim is to provide new telemedicine services for diabetes care emphasizing the provision to physicians and patients of personal health services 24 hours a day using a multi-access concept considering a full range of non-expensive and widely accepted information technologies. The multi-access services are implemented with a multi-agent architecture (see Figure 1) where communication agents and Knowledge Management agents are coordinated by a middleware module called Organizer. One of the goals of the project is to provide users with Knowledge Management tools to give the right knowledge to the right people in the right moment. The Knowledge management tools considered in the Spanish implementation of the m2dm telemedicine service are of three types: 1) On users' demand: it includes the graphical data representation; 2) Triggered by the system: it includes the dynamic pre-programmed notification of events to users, both patients and physicians; 3) Triggered by data reception: it includes the automatic data processing both for generating summaries for patients and intelligent alarms for physicians and patients.

The Knowledge Management Agent described in this paper belongs to the third category and uses hybrid methods to generate data summaries and automatic alarms, such as rule-based techniques and Causal Probabilistic Networks (CPN) techniques [3]. The rule-based analysis allows the detection of severe abnormalities in the patient state using different

Intelligent e-Health Applications in Medicine
Special Session Proc. EUNITE 2002 Albufeira Portugal Sept 19-21, 2002
Published by the University of the Aegean

time scales depending on the quality of the received information. The CPN analysis looks for deviations in the "insulin effectiveness" along a typical day using a physiological qualitative model and managing uncertainty and data incompleteness that are two of the main characteristics in the ambulatory monitoring of patients.

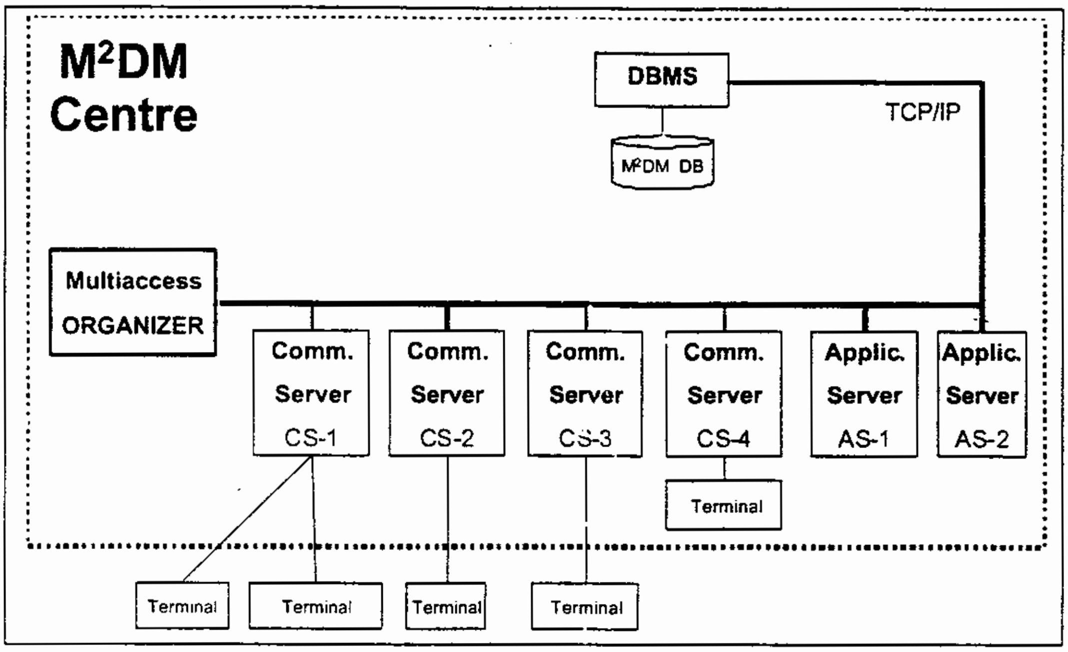


Figure 1: The Multiaccess Server Architecture composed by the common database, the Multiaccess Organizer, the agents (CS: Communication Server; AS:Application Server) and the user applications.

Some of the "M2DM Agents" are in charge of communications with the different user terminals integrated within the m2dm Centre (named Communication Server agents) and other agents are in charge of data analysis and data processing (named Application Server agents). The number of agents present in a M^{2}DM Centre is not limited a priori and it only depends on the kind of terminals that will be used at each site.

AUTOMATIC DATA PROCESSING METHODS

The KM Agent is activated whenever new monitoring data is received at the central multi-access server from any patient. When triggered, the KM agents performs the automatic blood glucose data processing both for generating summaries for patients and for detecting deviations in the patient state.

GENERATION OF DATA SUMMARIES

The main objective of data summaries is to give feedback to patients after sending his/her monitoring data to the hospital. Additionally, we take advantage of this system interactio : with the patient to provide him/her with more added-value information extracted both from the data received and from the previous data stored in the database. The processed information allows patients to compare their actual metabolic state with the past and motivates them to continue registering information.

The analysis performed by the KM agent is based on basic statistics that are presented to the patient in a summary text message. The analysis is done both considering all the blood glucose measurements within a period and classifying the readings as pre or postprandial measurements associated to the four main daily intervals (breakfast, lunch, dinner and night). The content of the test message is structured as follows:

- Header: It includes the general m2dm service identification

- Subheader: It includes the nature of the message -"automatic message generated by the system"- and the reason - "Reception of patient's monitoring data". The subheader is important to clearly distinguish this type of messages from others generated by humans.
- Body: The message body has two different sections
    - Information about the blood glucose data received: Date and time of the analysis and data summary (no. of measurements; global statistics; no. of hypoglycemia values; no. of hyperglycemia values; statistics along daily periods; etc).
    - Information about the patient's whole monitoring data available at the hospital: it includes a blood glucose summary similar to the previous one for the whole period and information about other monitored variables, such as intakes, insulin daily adjustments, illness, etc)
- Footer: It includes information to contact the m2dm responsible at the hospital

RULE-BASED ANALYSIS

The rule-based analysis allows the detection of severe abnormalities in patient control using different time scales depending on the quality of the received information. This analysis is triggered by multiple blood glucose data reception and is performed immediately after generating the data reception summary.

The severe problems that are considered are:
- Hypoglycaemia status. Presence of very low blood glucose values (< 70 mg/dl)
- Hyperglycaemia status. Presence of very high blood glucose values (> 200 mg/dl)
- Oscillating status. Simultaneous presence of hypo and hyperglycaemia values.
- Hyper-insulinitation: Administration of insulin under the dose prescribed by the diabetologist (> 100%).

The rules run in three different time scales that define the analysis time period:
- Transmitted data period: it covers the data received in the last communication process.
- Pre-fixed periods: 2 weeks / 1 month periods having as end point the communication date.
- Therapy related period: includes all the data from the last therapy update.

When any severe problem is detected an alarm is triggered and two different text messages are composed to notify the alarm to the patient/physician. Additionally, if the quality of data available in the pre-fixed periods is acceptable, the model-based analysis is triggered for doing a deeper data processing. The acceptable data quality is defined according to data incompleteness as a combination of the amount of blood glucose values (> 3 readings per day) and the percentage of days where the required values are available (> 80% of the days in the analysis period).

MODEL-BASED ANALYSIS

The model-based analysis uses a physiological model for diabetic patients represented with a Causal Probabilistic Network that can detect deviations in the "insulin effectiveness" along a typical day .

The qualitative model

The model represents the day divided into four time intervals ('breakfast', 'lunch', 'dinner' and 'night' intervals). The inputs into the metabolic model are divided into two categories: self-monitoring patient data, that can be considered as the "symptoms" of the metabolic state (blood glucose and ketonury measurements); and the "causes" affecting these measurements (insulin doses, meals, meal time and time span between insulin injections and their associated meals). The physiological parameters of this model are qualitative, this means that, for example, an intake is catalogued as "delayed" or "excessive in carbohydrate content" but no quantitative values in minutes or in grams are managed. The relations between the model parameters were obtained from experts and were represented with a qualitative Causal Probabilistic Network (CPN)

The CPN-model was previously developed and evaluated as the core of DIABNET [4], a decision support system for therapy planning that integrated the CPN-model in a hybrid architecture where the qualitative results were later on converted into quantitative insulin dose modifications adapted to the patient characteristics. The DIABNET system was evaluated with real patients [5] demonstrating the suitability of the qualitative CPN-model for therapy planning. One reason that makes a qualitative model suitable in the education process is that the knowledge used by professionals is also expressed to patients in qualitative terms, as for example: "If a main intake is delayed, then it is possible that it will produce ketonuria and a low blood glucose measurement". In addition, the choice of CPNs facilitates the model reutilization mainly because it supports both diagnostic and predictive reasoning. The previous characteristics greatly benefit any educational tool aiming to implement interpretation and simulation of data.

Detection of the insulin effectiveness

To interpret patient's data, the KM Agent exploits not only blood glucose readings but also the knowledge about patient's habits (i.e diet and schedule) and the therapy compliance. This information has been registered by patients and transmitted to the telemedicine server using any of the multi-access terminals provided by the project (Web forms, glucometers, PDA's, conventional phones, etc). When triggered, the KM agent feeds the CPN with the data stored for the specific patient along all the days included in the analysis period. The network infers the insulin effectiveness associated to each of the four daily periods giving qualitative information about the lack or the excess of exogenous insulin. The KM agent generates new alarms when any of the nodes takes the value "insufficient" or "excessive". This information focus physician attention on the points where the insulin therapy should be modified to improve patients' metabolic control.

RESULTS

The KM agent has been integrated in the m2dm multi-access architecture and cooperates with the existing communication agents to provide the analysis results to the recipient users. The communication of the KM agent with the Organizer follows the messages protocol defined in the m2dm multi-agent architecture for the coordination between the Organizer and the multi-access agents [6]. None modification in the previous communication protocol was required.

The integration requires the m2dm Organizer to be pre-programmed with an additional activity-event relationship that activates the KM agent whenever new monitoring data is received from a patient. The process at the Multi-Access server is as follows (see 0): 1) The glucometer agent receives data from a patient, it stores them in the database and notifies the "Data reception event" to the Organizer providing information about the author; 2) The Organizer consult the pre-programmed activity-event relationships for this specific event and author; 3) the Organizer notifies the event to the KM agent, that performs the data analysis, detect the alarms and composes the messages for the users. Finally it sends the "New message event" to the Organizer; 4) The Organizer consults again the pre-programmed activity-event relationships for this new event looking for the notification way preferred by each recipient (Web, conventional e-mail or SMS); 5) The Organizer activates the correspondent/s notification agent/s that will deliver the message/s.

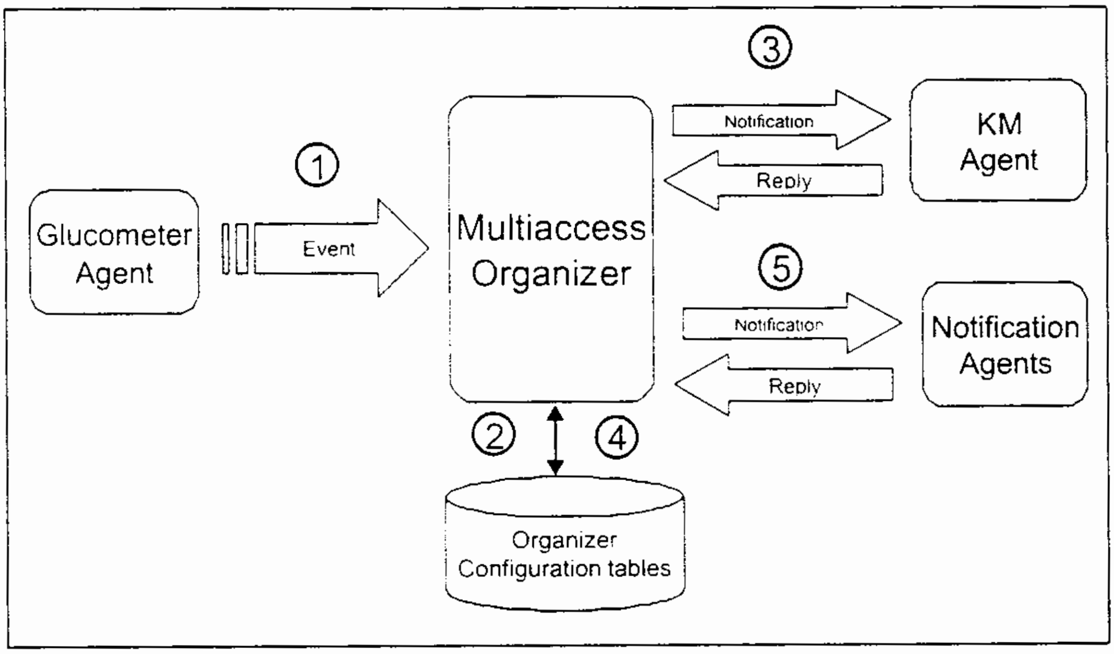


Figure 2: Interoperability between the KM agents and the Organizer

Alarms are recorded in the m2dm database together with the information to identify the patient, the date of the analysis, the analysis period and other relevant information for alarm interpretation. Users can be aware of alarms in two scenarios: 1) Monitoring data review scenario: Alarms are displayed in the graphical visualization of monitoring data; 2) Messaging scenarios: Alarms are notified to users as text messages through Web, e-mail or mobile SMS if they activate this features in their personal "user's notification preferences".

CONCLUSIONS

This KM Agent has been successfully integrated in the m2dm multi-access server. The advantages of integrating automatic response tools in a telemedicine service are clear because they provide two features that are very relevant in patients' care: 1) they focus doctors' and patients' attention on abnormal data preventing future risky situations and 2) they give instantaneous feedback to patients promoting patient interaction with the tele-care organisation and reinforcing their motivation to use the telemedicine system. For this reason, the architecture of any telemedicine service should be designed to allow the integration of any kind of knowledge management tools in a seamless manner and without interfering with the existing functionalities. In this sense, the current work demonstrates the suitability of the m2dm architecture for the coordination of heterogeneous agents developed for different purposes such as multi-access communications and knowledge management. The KM Agent is currently under evaluation at the m2dm Spanish pilot in the Hospital Sant Pau of Barcelona.

ACKNOWLEDGEMENTS

This work has been partially funded by the European Project M2DM (IST-1999-10315) and the Spanish Research Council (TIC2000-2491-CE).

REFERENCES

[1] E.J. Gomez, F. Del Pozo and M.E. Hernando, 1996, "Telemedicine for Diabetes Care: the DIABTel approach towards Diabetes Telecare", Med. Inform. 21(4), pp.283-295.

[2] M2DM. multi-access services for the management of diabetes mellitus. EU Project. IST-1999-10315. http://aim.unipv.it/projects/m2dm

[3] Pearl, J., 1988, "Probabilistic reasoning in intelligent systems: networks of plausible inference". Morgan Kaufmann, San Mateo/California, USA

[4] Hernando, M.E., Gómez, E.J., Del Pozo, F. and Corcoy, R. DIABNET: A qualitative model-based advisory system for therapy planning in gestational diabetes. Medical Informatics 21 (1996) 359-374.

[5] Hernando, M.E., Gómez, E.J., Corcoy, R and del Pozo, F. Evaluation of DIABNET, a decision support system for therapy planning in gestational diabetes. Computers Methods and Programs in Biomedicine 62 (2000) 235-248

[6] E. Hernando, E. J. Gómez, A.I García, A. Jalón, R. Bellazzi, S. Montani, F. Ortiz and J. M. Viejo, 2000, "Deliverable 5. Technology, services and systems specification". M2DM European Union Project (EU. IST-1999-10315) Brusels.

CNS Damage Classification in Newborn Infants by Neural Network Based Cry Analysis.

Mannes Poel and Taco Ekkel
Dept. of Computer Science
University of Twente
P.O. Box 217, 7500 AE Enschede, The Netherlands
email:{mpoel, taco}@cs.utwente.nl

ABSTRACT: The central nervous system (CNS) of the human body is the whole system of brain, spinal marrow and nerve cells throughout the body that correlates and regulates the internal reactions of the body and controls its adjustment to the environment [BOS70]. It controls muscles and processes sensory information originating from visual, aural, and other sensorial systems. Apart from that, it constitutes important human properties such as the ability to learn, control, think, feel, have self-awareness et cetera. The central nervous system not only directs the body on a conscious level but also on a subconscious level, making sure that we keep breathing, for example.

Considering all these functions, it is easy to understand that malfunctions of the central nervous system can easily result in very severe complications, ranging from diminished control over movements and actions to abrupt cessation of life.

One of the possible ways for the central nervous system to become damaged is by a shortage of oxygen. Normally, oxygen is extracted from the air by the lungs, which have a huge surface that is rich with capillaries. This surface allows for diffusion of oxygen from the air into the blood. After this, the heart pumps the oxygen-rich blood through the body, where the oxygen is used for various purposes, including the function of the central nervous system.

Hypoxia (or hypoxaemia) is the general phenomenon in which the oxygen level in the blood is too low for the body to perform normally. In hospitals, the oxygen level is usually monitored because it is a critical parameter in surgical anesthesia. Apart from the measurement tools available, hypoxia can often be diagnosed visually, because the subject's skin tends to turn bluish. In the case of newborns, however, this diagnosis is often more difficult. It is possible for the neonate to experience episodes of hypoxia before and during birth, in which case the hypoxic episode may very well go by unnoticed. When this happens, damage may have been inflicted to the central nervous system. This damage is often hard to recognize, as it may not show until a much later stage in the life of the newborn.

In this paper we explore a neural network approach for classifying infant cry in order to detect hypoxia related CNS damage. The data set consists of 35 recorded cases of infant cries episodes, called the FCU data set. Each cry episode is split into valid cry units resulting in a data set of 183 elements, called the ACU data set.

Relevant numerical features are determined by the research on features in the area of infant cry analysis and speech recognition. After determining the relevant features an ensemble of Radial Basis Neural Networks was constructed for both data sets (ACU and FCU) using bootstrap aggregation. Testing the classification performance resulted in a performance of 85 ±7 % (99 % confidence level) on the ACU data set and 76 ±19 % (99 % confidence level) on the FCU data set.

These classification rates should be approached with caution, however. The 78% is the lower bound of the confidence interval for the results of the classification is based on a small data set. It should therefore be stressed that the results from this investigation are not to be taken any other than 'preliminary'. Much more data and testing needs to be done

In conclusion, the results that were found are encouraging in the sense that a statistically significant relationship was found. The feature selection and classification methods developed in this investigation can be a sound basis for continuation. Further research should divert its attention to the (frequency) measurement and noise issues and to the data availability issues.

KEYWORDS: infant cry analysis, neural networks, hypoxia-related CNS damage.

INTRODUCTION

The central nervous system (CNS) of the human body is the whole system of brain, spinal marrow and nerve cells throughout the body that correlates and regulates the internal reactions of the body and controls its adjustment to the environment [BOS70]. It controls muscles and processes sensory information originating from visual, aural, and other sensorial systems. Apart from that, it constitutes important human properties such as the ability to learn, control, think, feel, have self-awareness et cetera. The central nervous system not only directs the body on a conscious level but also on a subconscious level, making sure that we keep breathing, for example.

Intelligent e-Health Applications in Medicine
Special Session Proc. EUNITE 2002. Albufeira, Portugal Sept. 19-21, 2002
Published by the University of the Aegean

Considering all these functions, it is easy to understand that malfunctions of the central nervous system can easily result in very severe complications, ranging from diminished control over movements and actions to abrupt cessation of life.

One of the possible ways for the central nervous system to become damaged is by a shortage of oxygen. Normally, oxygen is extracted from the air by the lungs, which have a huge surface that is rich with capillaries. This surface allows for diffusion of oxygen from the air into the blood. After this, the heart pumps the oxygen-rich blood through the body, where the oxygen is used for various purposes, including the function of the central nervous system.

Hypoxia (or hypoxaemia) is the general phenomenon in which the oxygen level in the blood is too low for the body to perform normally. In hospitals, the oxygen level is usually monitored because it is a critical parameter in surgical anesthesia. Apart from the measurement tools available, hypoxia can often be diagnosed visually, because the subject's skin tends to turn bluish. In the case of newborns, however, this diagnosis is often more difficult. It is possible for the neonate to experience episodes of hypoxia before and during birth, in which case the hypoxic episode may very well go by unnoticed. When this happens, damage may have been inflicted to the central nervous system. This damage is often hard to recognize, as it may not show until a much later stage in the life of the newborn.

In this article we discuss the possibility to differentiate between normal cases and abnormal cases as diagnosed by a medical expert. In practice, these abnormal cases that are dealt with are all somehow related to hypoxia in the sense that hypoxic episodes may have occurred, such as in case of a transpelvian euthocic delivery or a cesarean dysthocic delivery. It is the physician's professional opinion (which will be assumed to be correct) that these hypoxic episodes have had their effect on the central nervous system, which then should be reflected in the infant crying sound.

FEATURES FOR INFANT CRY ANALYSIS

Various studies have been undertaken to discover possible acoustical features in the infant crying sound that can be correlated to a certain state of the central nervous system. Since the goal is to discriminate between normal and abnormal cases, we need to find these features to perform this classification, in which we could view upon the features as 'decision variables'. Specific types of occurrences of features (or combinations of features) possibly relevant to the diagnosis will be referred to as *indicators*.

A starting point for building a candidate set of features is, of course, extensive research done in the field of hypoxia-induced disorders by the Grupo Procesamiento de Voz at the Universidad de Oriente. Cano et. al. [CAN95, CAN98, CAN02] annotate the following phenomena as possibly being indications for hypoxia-related disorders. For an explanation of the speech recognition related terminology see the book of Jurafsky & Martin [JM00] or the master thesis of Ekkel [EKK02]

The first formant is more variable in abnormal cases than in normal cases. The fundamental frequency is found to be, on average, significantly higher in abnormal cases. Instant changes in the fundamental frequency of four or more octaves occur significantly more often in abnormal cases. The latency of the first cry tends to be larger in abnormal cases than in normal cases. The melody type in test cases is, more often than in normal cases, of the rising type.

In order to expand the set of possibly useful features, literature offers the following hints in relation to either CNS-specific abnormalities or to abnormal cries in general [HIR99, MIC99]:Both the number of turbulent noises (stridors) and biphonations are increased in abnormal cases. In general, the number of aberrations tends to be higher in abnormal cases.

Furthermore, cries of non-healthy infants are, in general, more often considered to be either hyperfunctional (higher, louder and shriller) or hypofunctional (lower, weaker and thinner) [MIC99]. This can be expressed in terms of energy (higher or lower than usual), voicedness (lower than usual), fundamental frequency (higher or lower than normal), and formants (higher or lower than normal).

Considering all of the above, the following features will need to be extracted from the crying sound in order to measure all of the aforementioned indicators: fundamental frequency over time, first formant over time, energy over time and voicedness over time

From these values, the remaining features can be calculated: number of occurrences of shift, number of occurrences of stridor, number of occurrences of biphonation, amount of jitter, melody type and latency.

THE DATA SET AND FEATURE SELECTION

THE DATA SET

Initially, the research team of the Universidad de Oriente considered the measured values over a whole cry *episode* (one case, or 12 seconds of sound, often containing many cry units) as one pattern. For one episode, values like average, minimum, maximum, and standard deviation of the measured valued were assembled to form a pattern. Thus, the number of data points was equal to the number of recorded cases: 38.

Another approach was used by Mulder [MUL01] who split a cry episode into cry units and considered the average values for one cry unit as a single data pattern. Hence, this approach yields more data points: depending on the definition of a valid cry unit, this may range from some 120 to almost 200 patterns. Having more patterns is desirable in terms of classification.

The drawback of this approach is that the cry units after the first one are increasingly less relevant to the infant's status, due to the fact that they represent a transitory process from pain cry to normal cry; the development of this process over time differs from case to case. Therefore, by using all units, the amount of information does not grow linearly with the amount of patterns. Still, it makes sense to prefer a much larger amount of data points even if the amount of information added is not in proportion, because the additional data points are not expected to *decrease* the total amount of information.

Since the first cry unit is most relevant to the status of the infant, we will develop subsequent classification efforts by using *two* sets of data points. The first one contains only information from the first cry unit, coupling a highly relevant informational content to a very small data set of only 35 patterns. The second one contains information from all cry units, resulting in 183 patterns, at the price of an expected decrease in relevance.

From here on, the data set with 35 patterns will be referred to as 'FCU' (first cry unit) and the set with 183 patterns will be called 'ACU' (all cry units).

For the ACU data set, there is a representation issue to be solved: the value of the first latency not present for all cry units, as this value applies only to the first cry unit of a case. However, the classification algorithm needs the same format for each input pattern. To work around this, we will assign the second and further cry units from a certain case the *same* latency value as the first cry unit. This decision is based simply on the fact that all cry units are part of the same case. They hence represent the same physiological status, to which the first latency is supposedly related.

Given the large set of possible features and the small amount of data, the ACU data set consists of 183 examples and the FCU data set of 35 examples, there is a need for feature reduction in order to avoid the curse of dimensionality.

FEATURE SELECTION

The first approach we took to reduce the number of features was Principal Component Analysis. Applying the PCA algorithm to our data set yields a set of principal components, of which the first two (having the greatest variance) are used to project the whole data set onto two dimensions in Figure 1 (ACU) and Figure 2 (FCU).

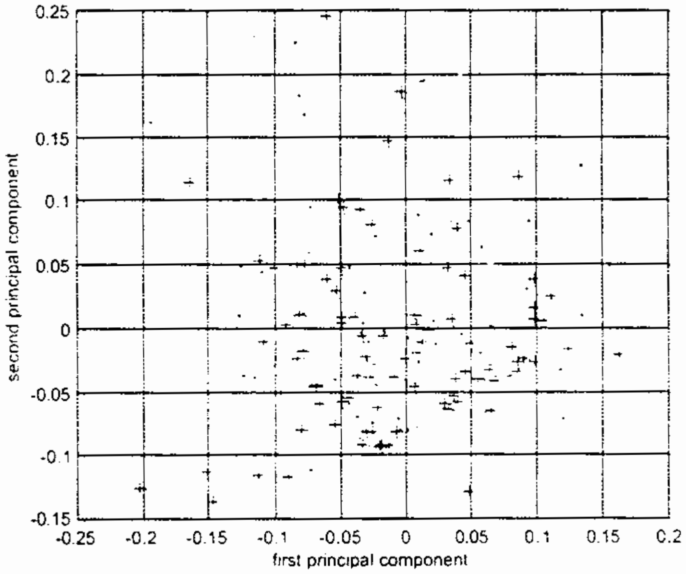


Figure 1 : Projection of cry data (ACU) on first two principal components
(− : normal; • : disorder)

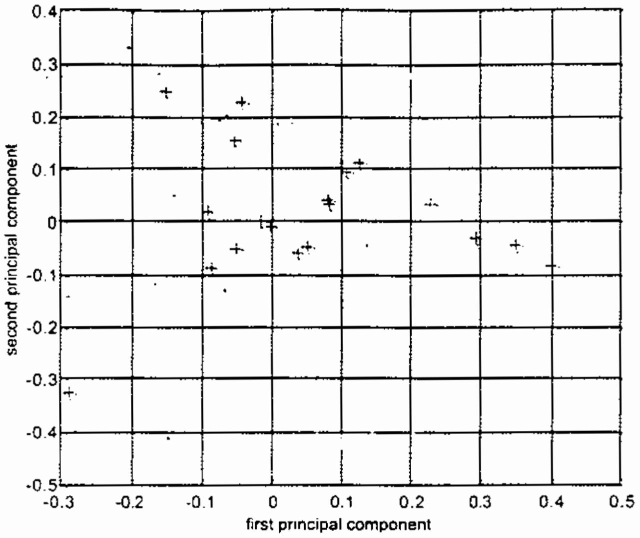


Figure 2 : Projection of cry data (FCU) on first two principal components
(+ : normal; • : disorder)

Although the algorithm finds the axes along which the variance in the data is largest, there is no visible separation between target classes. This can mean either one or both of the following:

1. The variance caused by class difference is much smaller than the variance caused by other factors such as inherent between-infant and between-cry differences.
2. The linear analysis discards the majority of the information because it is hidden in nonlinear relationships.

At any rate, we cannot be sure that the information discarded by the PCA algorithm is *not* important in separating the classes, so this form of PCA cannot be used.

Another option is to simply *try* every possible subset of features that can be constructed from the whole feature set, and then selecting the subset yielding the best test results. This process is called 'feature selection' and is actually the simplest and most logical method for feature selection or dimensionality reduction.

To compare different subsets, we need an algorithm that provides with a certain criterion describing the quality of the subset. In this application, for example, the whole feature set contains 25 features. So, for all possible subset sizes, the total number of possible features combinations is 33.554.431. Since each combination would have to pass through a nonlinear training (optimization) and testing process, the computational requirements would become far too great to simply try every combination. If a nonlinear network would need one second to be trained and tested for a subset, testing all features would take over a year.

The approach we took was to combine radial basis neural networks and sequential backward elimination for feature selection. To select those features that are best *in general*, the whole feature search was repeated many times, each time with a different random partition of the pattern data into training- and test set. The results are then averaged, and those features that appear in optimal subsets most often can be considered the overall best features.

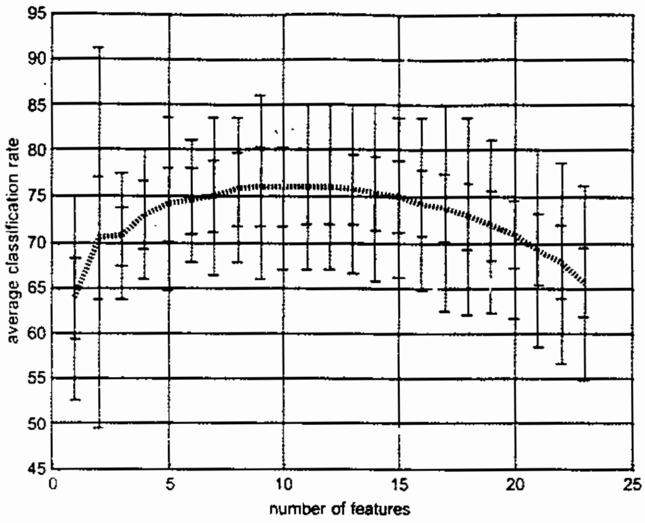


Figure 3 : Averaged best classification rate: 100 times for all 465 subsets (ACU set)
(with bold error bars: standard deviation; thin error bars: extreme values)

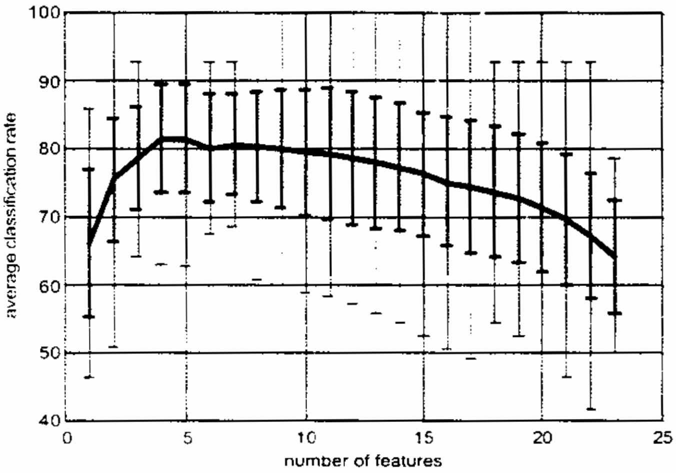


Figure 4 : Averaged best classification rate: 100 times for all 465 subsets (FCU set)
(with bold error bars: standard deviation; thin error bars: extreme values)

The best features are given in Table I and Table II.

Feature	Appears in best feature set
First latency	71 %
Voicedness mean	60 %
F_1 minimum	59 %
F_1 mean	52 %
F_0 mean	51 %
Voicedness variability	41 %
Energy maximum	33 %
Voicedness standard deviation	30 %
F_0 maximum	30 %
Energy mean	26 %

Table I : Best features (ACU set)

Feature	Appears in best feature set
First latency	47 %
Energy standard deviation	38 %
Voicedness standard deviation	36 %
Voicedness variability	35 %

Table II : Best features (FCU set)

In subsequent classification efforts, the features shown in Table I and Table II are the features that will be used.

NEURAL NETWORK BASED CLASSIFICATION

For constructing a neural classifier, an ensemble of RBF networks is constructed and combined using bootstrap aggregation. Bootstrap aggregation (bagging) is a form of *arcing*: adaptive re-weighting and combining. This general term refers to reusing or selecting data in order to improve classification [DUD01]. In bagging, several versions of a training set, each creating by drawing samples with replacement from the data set, are created. Each training set is then used to train a classifier (RBF network). Then, in the classification stage, the outputs of all those networks are combined to form the final classification based on the vote of each individual network (also called component classifier). Since the classifier is now essentially composed of many neural networks, it is referred to as a multi-classifier system, and the individual networks are called component classifiers.

Sampling with replacement means picking one data point at random without removing it from the data set it came from, so that it may be selected again. A bootstrap sample from a data set may therefore contain the same data point more than once. By constructing many versions of the same data set, each component classifier will create a different 'view' on the mapping that's at the base of the classifier's data set.

In this multi-classifier system, all component classifiers are of the same type and share the same output format. The output of one component classifier may simply be seen as the probability that some case belongs in a certain class (or the probability that the case does not belong in the other class). We may therefore simply linearly add the probabilities provided by the component classifiers and average them. After this, rounding the average value directly provides us with the classification result.

A multi-classifier system of 10 networks was constructed. On the FCU data set, the performance of the classifier system was 76%. The performance on the ACU data set was 85%. Confusion matrices are shown in Table III and Table IV; they show a slight inclination towards type 1 errors ('false alarms') which is a 'good' thing (in medical terms, a false alarm could be considered less damaging than unjustly *not* reporting an abnormality).

		predicted class	
actual class	group size	0	1
0	15	79 %	21 %
1	21	12 %	88 %

Table III : Confusion matrix for multi-classifier system on ACU data set (0: normal; 1:abnormal)

		predicted class	
actual class	group size	0	1
0	9	70 %	30 %
1	6	11 %	89 %

Table IV : Confusion matrix for multi-classifier system on FCU data set

EVALUATION OF THE CLASSIFICATION RESULTS

LINEAR DISCRIMINANT ANALYSIS

We also applied a linear discriminator to the data set and found the following results. Although this is probably a little bit off bias due to small differences in the prior probabilities (the amount of patterns in each class) we use this discriminator function to classify all cases and then compare the results to the *real* class memberships. For the FCU data set, this resulted in a correct classification of no less than 97%; Table V shows the corresponding confusion matrix. The algorithm performs less on the ACU data set, resulting in a classification rate of 74% (see Table VI for the confusion matrix).

		predicted class	
actual class	**group size**	0	1
0	19	95%	5%
1	16	0%	100%

Table V : Confusion matrix for linear classification (FCU) on the training set, resulting in 97% correct classification

		predicted class	
actual class	**group size**	0	1
0	90	74%	26%
1	93	27%	73%

Table VI : Confusion matrix for linear classification (ACU) on the training set, resulting in 74% correct classification

The result for the ACU data set is not very satisfying, but the good results for the FCU set demand further investigation. If this linear classification mechanism is to be used in practice, it should be tested on test data ('real life data') instead of the training data it was constructed with. To get the idea of what would happen in such a case, we start by analyzing the sensitivity of the linear classifier to new data. This is done by dividing the data set in a training set which is used to find the discriminator function and a testing set, which is used to test its performance. For each different test set size, the classification rate is the average of 100 different test runs each having a randomly selected (set of) example(s) from the data set as test examples.

Figure 5 shows what happens in this testing scheme, in a graph, which shows the classification rate as the number of examples taken from the data set increases at the cost of decreasing the number of training examples[1]. When one sample is kept apart from the rest of the data set and used as testing example (so about 97% of the data is used for training), the performance drops to 82%. Adding another example to the test set makes the performance drop to only 65%. The two 'bumps' in the figure (one around 90% and one towards 80% on the horizontal axis) are results of coincidentally varying influences of individual patterns due to the small size of the data set.

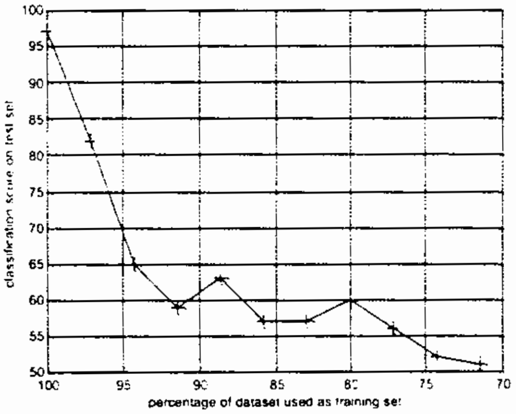


Figure 5 : Performance of the linear classifier on new data

---

[1] The first value is the 97% from the experiment in which all examples used both for training and for testing. It is shown here only for reference.

Therefore this classifier is not feasible for any practical use since its performance on unknown data is not at all good: 65% correct classification at best. The poor performance of the linear classifier can be the result of two-reasons. First, it is possible that the data set does not or hardly contain the information necessary to distinguish between the classes. Secondly, it is very well possible that the relation between feature data and class membership is of a nonlinear nature. This would explain the non-optimal classification by a linear discriminator, as it is not capable of modeling this nonlinear nature.

CONFIDENCE INTERVALS

Given the classification results in the previous section of the neural network based classifier we calculated the confidence intervals [MIT97].

Data set	**Measured performance**	**Lower bound**	**Upper bound**
ACU	85 % ± 5 %	80%	90%
FCU	76 % ± 14 %	62%	90%

Table VII : 95%-Confidence intervals for the neural network classifiers.

Data set	**Measured performance**	**Lower bound**	**Upper bound**
ACU	85 % ± 7 %	78%	92%
FCU	76 % ± 19 %	57%	95%

Table VIII : 99%-Confidence intervals for the neural network classifier

In words, with a probability of 95%, the ACU classifier's performance (correct classification rate) is between 80% and 90%. Since we are dealing with an application in which the reliability of the system is especially important, it's best to use the lower bound – 80% for the 95% confidence interval – to specify the expected performance of the system so as to prevent too high expectations in the medical community.
Then, there's the data quality, or rather the amount of *noise* present in the data set. It does not directly influence the confidence interval itself, but it is sure to be of influence on how representative the classifier system is for the real world situation. Noise is present in three stages of the collection of data:

1. Physical noise in the recording: ambient sounds from the hospital room, tape deck wow&flutter, humming sounds as a result of poor groundwiring.
2. Noise in the measurement of features in the sound: imperfections in the frequency and voicedness measurement algorithms, limited sampling rate, limited dynamic range.
3. Noise in the class information: however unlikely, a physician *may* have established a faulty diagnose.

It would be reasonable to assume that the responsible physicians have checked their diagnosis later on, thereby virtually excluding the possibility of faults resulting in a 0 % noise level in the class information. Even then however, the first two noise components cannot be determined, or even be reasonably estimated, at this time. Especially the imperfections of the frequency estimation algorithm pose a threshold in the reliability of the total system. Although it is known that, in general, LPC-based pitch determinations yield performances of more or less 95% correct determination [OSH00], this only goes for adult speech, which is, from the analysis viewpoint, less intricate than infant crying.

CONCLUSIONS

As can be seen from the results in the previous section the classification system that was devised is capable of distinguishing between normal and abnormal cases with up to 85 % accuracy. Given the confidence interval (± 5% on the 95% confidence level), this result can be regarded as statistically significant. Hence is possible to improve the classification of newborn infant's cry into normal and abnormal (CNS damage) by using an ensemble of Radial Basis Neural Networks
It is interesting to see that, as compared to previous investigations in classification of hypoxia-related infant cries, especially the newly deployed features such as voicedness, energy and latency are most critical for reaching this classification rate. It is assumed that those features, coupled with the deployment of kernel-based neural networks for both feature selection and classification, made it possible to get to the aforementioned classification rate.
Regarding the two data sets (FCU and ACU), we note that using all recorded cry units seems to work better than using just the first cry units of each cry episode. It is difficult to say however whether this should be contributed to a very small amount of decrease in information in the non-first cry units (as opposed to a severe decrease in information), or to

a rather severe consequence of the size of the FCU data set. It is recommended to keep the options open at this time, e.g., to continue the investigation keeping both sets in mind.
The classification rates should be approached with caution, however. As discussed, the 80% is the lower bound of the confidence interval for the results of the classification are based on a small data set and the assumption of *noise-free data*. We are pretty sure however that the data is *not* noise-free, and therefore conclusions about how the classification is made by the system may not reflect real-world interdependencies. It should therefore be stressed that the results from this investigation are not to be taken any other than 'preliminary'.
In conclusion, the results that were found are encouraging in the sense that a statistically significant relationship was found. The feature selection and classification methods developed in this investigation can be a sound basis for continuation. Further research should divert its attention to the (frequency) measurement and noise issues and to the data availability issues.

REFERENCES

[BOS70] Bossy J., *Atlas of Neuroanatomy and special sense organs*, W.B. Saunders, Philadelphia (1970), ISBN 0-7216-1875-8

[CAN95] Cano S. D. & Escobedo D. I., *Spectral Analysis of the Infant Cry: An initial approximation*, Eurospeech 95 proceedings, Madrid, 1995

[CAN98] Cano S. D. & Escobedo D. I., *Clasificación de unidades de llanto infantil mediante el mapa auto-organizado de kohonen*, AIRENE, Grupo de Procesamiente de Voz, Universidad de Oriente, Santiago de Cuba, Cuba, 1998

[CAN02] Cano S. D., Escobedo D. I. & Coello E., *Some acoustical features of the cry signals in infants with epilepsy*, to be published

[DUD01] Duda R., Hart P. & Stork D., *Pattern Classification, 2nd edition*, John Wiley & Sons, Inc., ISBN 0-471-05669-3, 2001

[EKK02] Ekkel P.T., *Neural network-based classification of cries from infants suffering from hypoxia-related CNS damage*. Master thesis. Dept. Computer Science. University of Twente, 2002.

[HIR99] Hirschberg J., *Dysphonia in infants*, International Journal of Pediatric Otorhinolaryngology 49 suppl. 1 (1999), pp. S293-S296

[JM00] Jurafsky D. & Martin J.H. *Speech and Language Processing. An introduction to Natrual Language Processing, Computational Linguistics, and Speech Recognition*. Prentice Hall, ISBN 0-13-095069-6, 2000.

[MIC99] Michelsson K & Mischelsson O., *Phonation in the newborn, infant cry*, International Journal of Pediatric Otorhinolaryngology 49 suppl. 1 (1999), pp. S297-S301

[MIT97] Mitchel T.M., *Machine Learning*. McGraw-Hill, ISBN 0-07-042807-7, 1997.

[MUL01] Mulder M., *Infant Cry Analysis a clue using self-organizing maps*. Dept Computer Science, University of Twente, 2001.

User-Dependent Adaptability for Implementation of an Intelligent Dental Device

Detlef Mitzschke[1], Jens Strackeljan[1], Fiamma Ferrero[2], Bernd Oehme[2]

[1]Technische Universität Clausthal
38678 Clausthal-Zellerfeld, Graupenstr.3, Germany
Phone: +30-821-37323, Fax: +30-821-69410
email: {detlef.mitzschke, jens.strackeljan}@tu-clausthal.de
[2]Sirona Dental Systems GmbH
Fabrikstr. 11, D-52076 Bensheim, Germany
Phone: +49-2408-94580, Fax: +49-2408-94582
email: bernd.oehme@sirona.de

ABSTRACT: The prevention, diagnosis and therapy of periodontal diseases are relevant aspects of dental care. Calculus is a mineralised biomass with porous microstructure which harbours bacterial plaque and causes inflammation of the supporting structures of teeth (periodontium). The very hard substance calculus can only be removed during a professional cleaning. Therapy consists therefore in a careful professional removal of the subgingival bacterial and calculus accumulations from the crown and root surfaces of the teeth and below the gumline. Potential risks arise during the therapy because either some root surfaces may be damaged by excessive treatment (overtreatment) or some calculus may be left on place (undertreatment. The objective is to integrate a verification tool, which can classify the kind of the solicited dental surface, into the therapy instrument, the ultrasonic scaler. Automated detection of hard substances (enamel, cement, dentin) and hidden calculus is theoretically feasible, since they differ with respect to elasticity, density, hardness and surface properties. The surface recognition feasibility relies on is the analysis of the impulse response: if opportunely solicited, different surfaces show different impulse responses. The problem which has to been solved is that the vibrational signal which is the basis of the feature extraction algorithm depends not only on the surface, but depends strongly from the user handling. The signal pre- processing has to consider these changing environmental effects because otherwise the classification accuracy increases significantly.

KEYWORDS: Pattern recognition, feature extraction and selection, dental device, subgingival calculus

1. INDRODUCTION

Although many ignore, prevention, diagnosis and therapy of periodontal diseases are relevant aspects of dental care. In the United States more than one in three people over age 30 have a form of periodontal disease that has advanced beyond gingivitis and periodontal disease together with dental decay are the primary causes of adult tooth loss. Recent research has even found a relationship between periodontal infection and more serious health problems, such as cardiovascular diseases, diabetes, respiratory diseases and premature births[1].

Figure 1 presents a front view of a two rooted tooth detailing some substances in the root region: dentin and cement. Some pathological processes are also shown: gingiva recession and deposits of plaque and subgingival calculus. Calculus, also known as concrement, is a mineralised biomass with porous microstructure which harbours bacterial plaque and causes inflammation of the supporting structures of teeth (periodontium), namely gingiva, periodontal ligament, cement and alveolar bone.

If plaque is not removed, it turns into calculus in less than two days. The very hard substance calculus can only be removed during a professional cleaning. Therapy consists therefore in a careful professional removal of the subgingival bacterial and calculus accumulations from the crown and root surfaces of the teeth and below the gumline ("Clean it

---

[1] Destructive Periodontal Diseases in Adult 30 years of Age and older in the United states. 1988-1994*Journal of Periodontology*, January 1999,13-29.

Intelligent e-Health Applications in Medicine
Special Session Proc EUNITE 2002 Albufeira Portugal, Sept 19-21, 2002
Published by the University of the Aegean

all"). This technique is known as "scaling and root planing" (*Figure 2*). Scaling therapy is performed using oscillating instruments such as ultrasonic driven scalers.

Potential risks arise during the therapy because either some sound root surfaces may be damaged by excessive treatment (overtreatment) or some calculus may be left on place (undertreatment). Nowadays, to prove whether a sufficient quantity of calculus has been removed, a verification phase by means of a mechanical probe is performed. If not, a further treatment session is newly started. It would be practical if verification and treatment could be done with the same instrument and if the instrument itself could help in recognising hidden concrement deposits during verification. The objective is to integrate a verification tool, which can classify the kind of the solicited dental surface, into the therapy instrument, the ultrasonic scaler [KRM97].

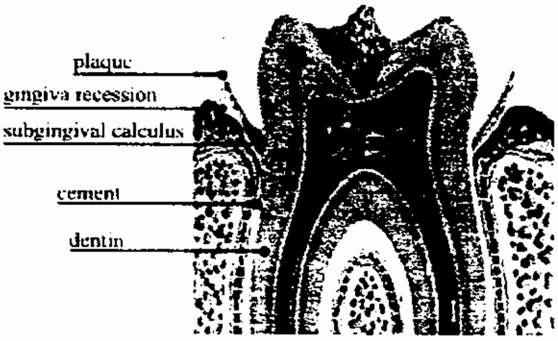


Figure 1: A tooth affected with periodontal disease

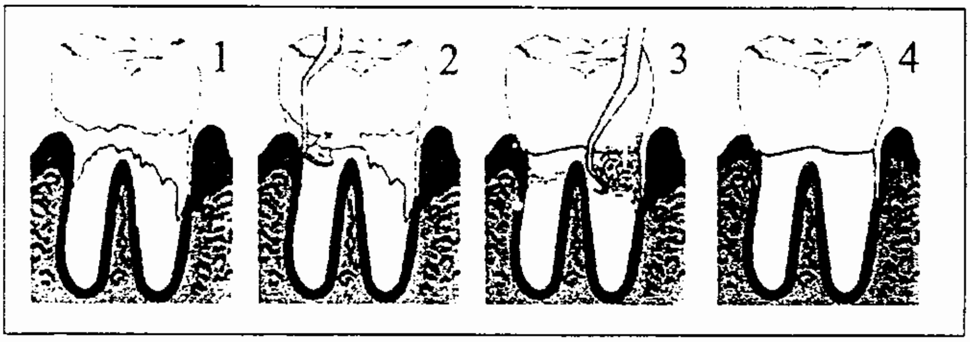


Figure 2: Scaling therapy. Rough root surfaces are cleaned and smoothed.

2. PRINCIPLES OF AUTOMATED SURFACE RECOGNITION

Automated detection of hard substances (enamel, cement, dentin) and hidden calculus is theoretically feasible, since they differ with respect to elasticity, density, hardness and surface properties. The surface recognition feasibility relies on is the analysis of the impulse response: if opportunely solicited, different surfaces show different impulse responses (*Ultrasonic Contact Impedance Principle*).

The system includes the scaler used both as actor and sensor, an exciting subsystem and a recording system (*Figure 3*) [SKB00]. The working principle of ultrasonic scalers used as sensors is innovative and relies on the application of the converse piezoelectric effect. The excitation system produces the pulse train and the high voltage: squared voltage pulses with an amplitude of 200V and a pulse duration of 10^{-4} to 10^{-5} sec excite the piezosystem at ca. 10 times/sec. The surface response to this stimulus induces a mechanical deformation of the piezoceramic which is transferred back to the recording system. The recording system filters the high voltage and prepares the signal for further processing.

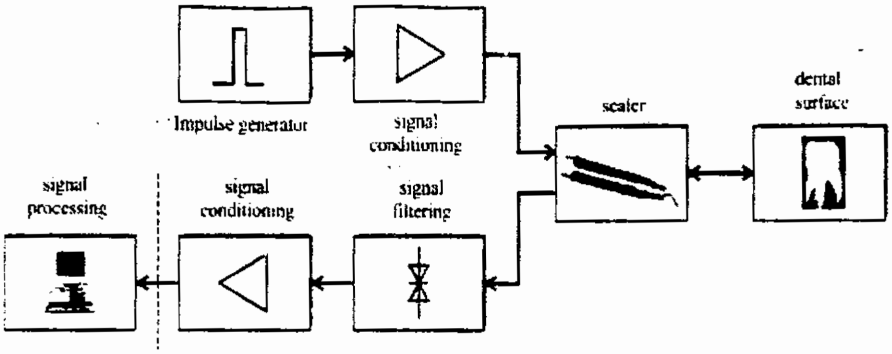


Figure 3: Measuring system blocks

A measuring session can be described as follows. An experienced operator places the tip of a commercially available ultrasonic scaler on the tooth, and the system as a whole is stimulated to oscillate. The surface specific pattern of oscillation is measured during this time by re-transmitting oscillations to the piezoceramic system.

3. DENTASOFT DESIGN ISSUES

3.1 GENERAL DESCRIPTION

Although the short excitation impulse stimulates the global system to oscillations in the upper kHz range, the difference among various surfaces is hardly visible in time signals. Visual analysis of power spectra[2] demonstrates that signals differ in few frequencies, while wide frequency ranges possess almost identical magnitudes (Figure 4). Figure 5 shows a zoom of the power spectra in the frequency range between 76 and 84 kHz. In spite of the considerable variations in the amplitude values the use of derived feature like the gradient in selected frequency bands could improve the partition between both classes (Figure 11). The classification task could not be solved by taking as features those amplitudes where a slight difference between the two classes is remarkable. As it can be conveniently illustrated by sample plotting in two dimensional feature maps (Figure 6), "clouds" due to scattering between samples of the two classes consistently overlap, thus not allowing the definition of a decision boundary.

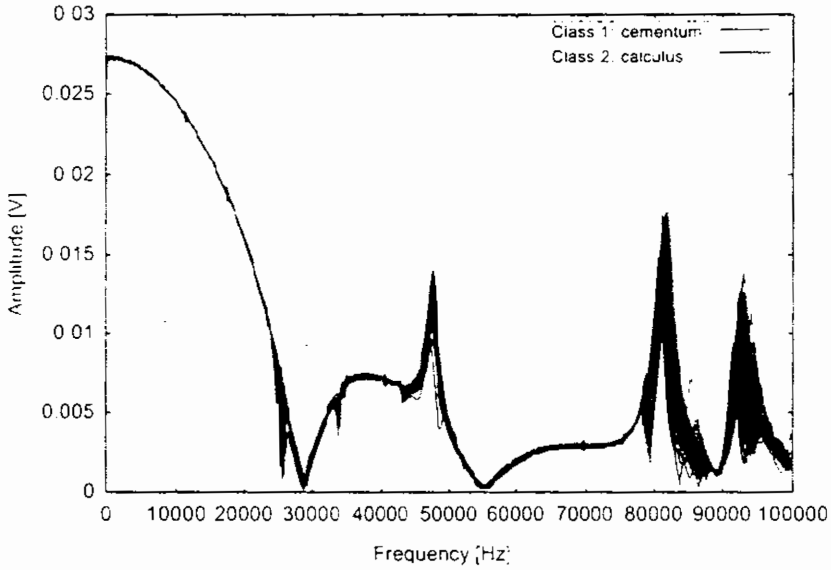


Figure 4: Root and concrement frequency spectra after pre-processing.

---

[2] Power spectra are FFT magnitudes of the time signals. Phase information is neglected.

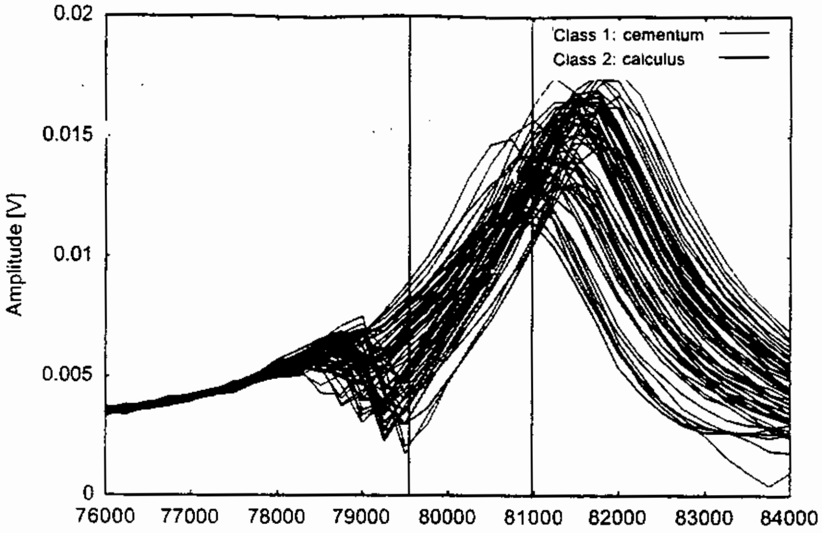


Figure 5: Illustration of the derived feature gradient.

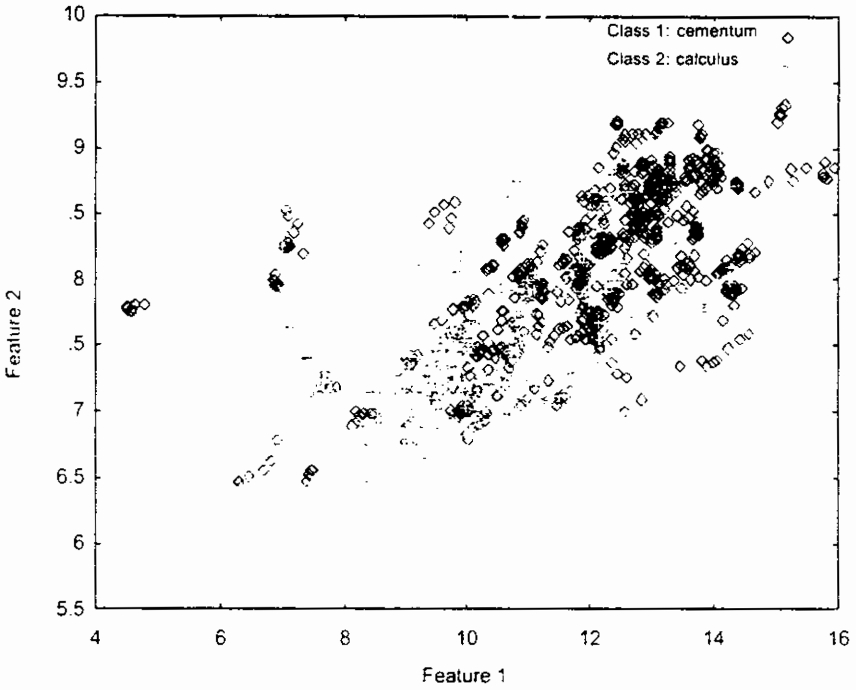


Figure 6: Task complexity in the feature map: clouds overlap even if visually discriminating features are chosen.

All this motivates the design of a principled classification strategy in the form of *Dentasoft*, a software digital signal processor based on soft computing techniques (neural networks, evolutionary strategies) and able to recognise hidden dental surfaces and deposits (e.g. root and concrement). The pattern recognition system should take as inputs the signals recorded by the ultrasonic scaler and should output the code corresponding to the surface which has been solicited.

Dental surfaces are described by means of their impulse responses (patterns); patterns are described using the following characteristics: frequency features (=samples of power spectra) as primary features, time signals and parameters calculated from time signals as secondary features.

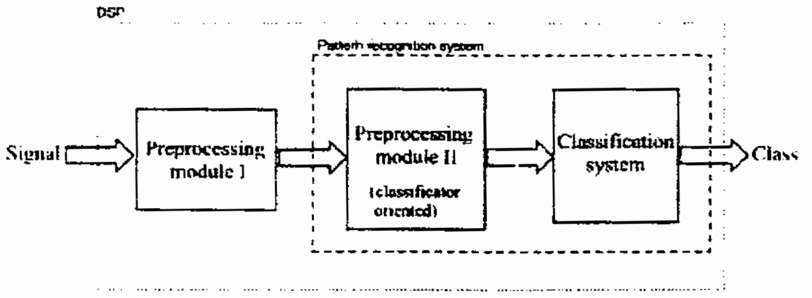


Figure 7: Software signal processor for dental surface pattern recognition.

To enhance flexibility and reusability, *Dentasoft* has been divided in three main blocks: two pre-processing modules and the actual pattern recogniser (Figure 7).
Pre-processing module I takes as inputs the information recorded by the ultrasound scaler and gives as outputs primary and secondary feature vectors. It accomplishes all pre-elaborations not concerned with a specific classification system: impulse response extraction, time feature[3] definition, FFT calculation, general class encoding (e.g. class "1" for root and class "2" for concrement).
Pre-processing module II takes the outputs of pre-processing module I and prepares input feature vectors for the classification system. Facilities provided by the pre-processing module II include the averaging of all examples of a class (e.g. to evaluate the influence of scattering between samples), the frequency adaptation (a different frequency resolution may be obtained by averaging neighbour frequency samples)[4], the noise reduction (features with minor amplitudes may be eliminated), the signal filtering (frequency bands where influences of the measuring systems are more relevant than the difference between the classes can be a priori eliminated)[5] and the integration of time features[6].
The classification system itself is also split into a feature extractor (an evolutionary strategy performing the task of extracting the most discriminant feature combinations) and a classificator (implemented using a neural network).

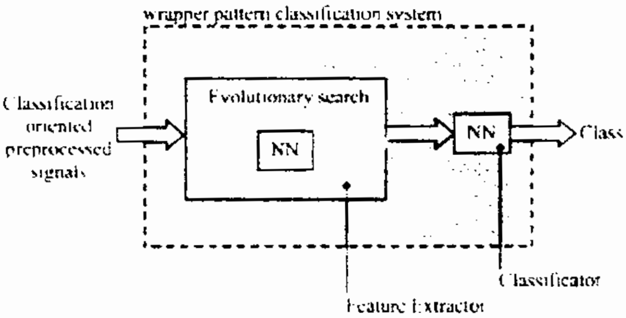


Figure 8: An overall concept for dimensionality reduction and classification.

The objective of finding a combination of features well suited for a given classifier is pursued according to the *wrapper approach* (*Figure 8*). This means that the performance of the feature set is assessed using the chosen classification algorithm, so that the applied optimising criteria are directly related to the classifier. Consequently, the error criteria

---

[3] Time signals are filtered, if frequency bands are known not to be significant.

[4] Results achieved with different frequency resolutions chosen by trial and error support the final choice of the most appropriate frequency resolution. This can be considered as an elementary feature reduction procedure inglobated in the pre-processing stages.

[5] Windowing is done to speed up the process of feature extraction. It may also follow the genetic search used as signal processing tool if, observing the search results, bands are found where the two classes do not differ.

[6] It is necessary to know their position in the vector because time features are not allowed to be modified by operators.

used for driving the learning phase of the network are the same ones building the fitness function associated to each evaluated feature combination. The reference neural network structure consists in a Multi Layer Perceptron trained by means of the Backpropagation training algorithm modified with momentum [Pao89]. The complexity of the net structure has been determined by trial and error, reaching the final form of a two layered MLP with 6 logistic neurons in the first hidden layer and 3 in the second one.

3.2 A MULTI-TASKING EVOLUTIONARY SEARCH IN THE FEATURE SPACE

The reduction of feature vector dimensionality aiming at excluding irrelevant and redundant features (*feature extraction*) may be accomplished either by selecting some of the existing features (*selection*) or creating new ones (*creation*). *Dentasoft* performs both techniques in parallel, in a mutual proportion which is flexible but fixed during the learning process. Feature selection is randomly performed, while new features are created transforming original features by a set of non-linear pre-operators (such as log, exp, ^2) and then combining them by simple binary arithmetic operations (such as +,-,*,/) or using step and derivative approximations.[7] Objective of such a feature extraction technique is twofold: finding useful projecting feature planes where clouds of features can be more easily separated, but also discovering relationships between non-neighbour frequency amplitudes.

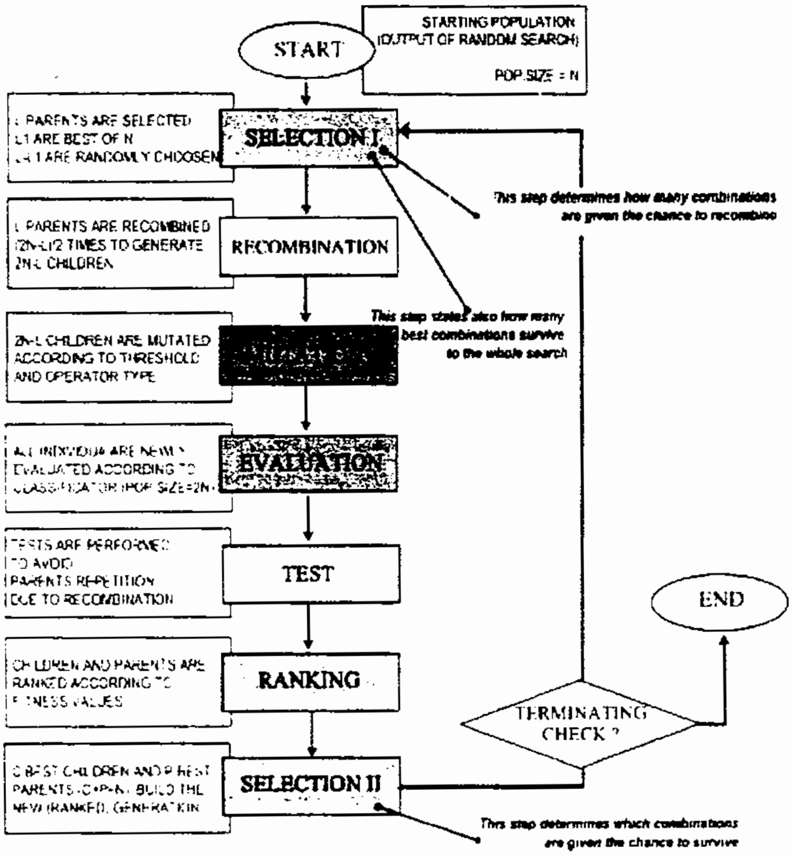


Figure 9: Evolutionary search block scheme.

The search in the feature space is based on an evolutionary strategy using an evaluation function measuring the classification accuracy and the average system error as feedback to guide the search. The best feature subset found is then output as the recommended set of features to be used in the final design step of the classification system. Genetic operators, such as mutation, have been designed regarding the individuum representation in order to enhance the search

---

efficiency and to prevent it from being trapped in local minima. *Figure 9* illustrates how the search finds best feature combinations starting from a random population.

3.3 TEST RESULTS

In order to obtain a robust and effective classification system and to make optimisations more likely to work well also if applied to real data, *Dentasoft* has been firstly tested on a classification task reproducing the real one in a simplified way. Two power spectra have been simulated by means of 100 features. Features are noisy values between 0 and 1, except for 2 ranges where amplitudes up to 4 are generated at one and two third of the total range. The two classes differ from the derivative with which the amplitude values grow to reach peaks of 3.5-4 Volts. Class 1 has derivatives around 2-2.5 in both ranges while class 2 has either low values in both regions or low derivatives around the two-third region and high values in the other one (*Figure 10*). Then only combinations of 4 selected features or combinations of 2 newly created features including the operator "derivative" allow a 100% separation between the classes (*Figure 11*).

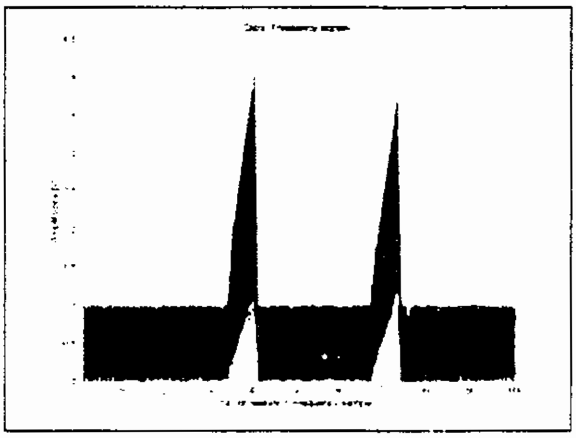


Figure 10: Test session by means of simulated data (100 features).

Dentasoft has a good performance as it finds a set of combinations which obtain a reclassification rate of 1.0 for train, test and recall sets.

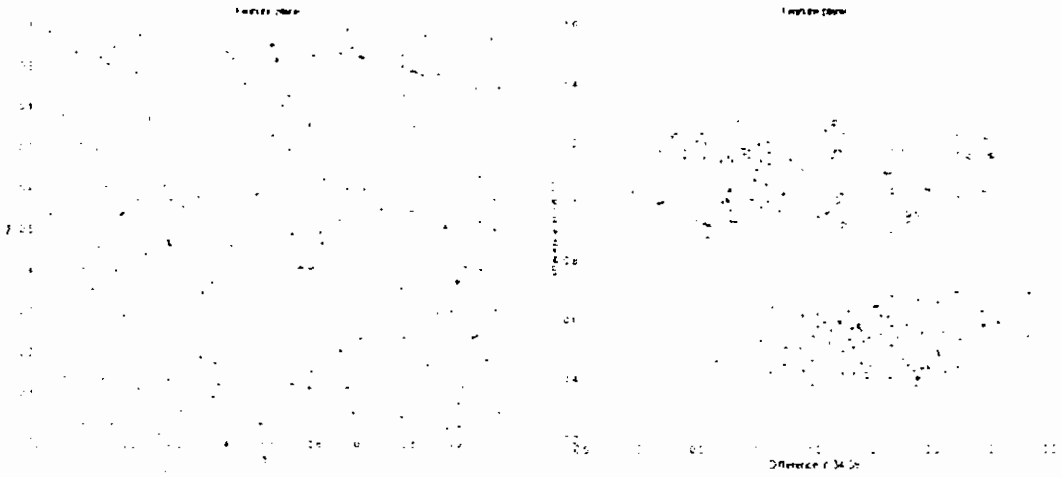


Figure 11: Clouds migrate in opposite directions if the right operators and feature combinations are chosen.

Dentasoft has also been tested with real data. Approximately 3700 measurements has been taken using ca. 70 fresh extracted teeth conserved in a cool NaCl solution. Tests with real data reached promising reclassification rates (0.95 for train data and 0.90 for test data). During real tests, subtraction and non linear operators such as derivative and division seemed to be more effective than others and particularly derivatives in frequency regions near to the peaks seemed to be useful to discriminate between the classes. Combinations of newly created features contained also features lying where no difference in the spectra is visible, probably thanks to the discovering of hidden relationships between non neighbour features.

4. CONCLUSION

The objective of the present study was the design of an unconventional, soft computing based classification system. The ultimate purpose of such a system was to be combined with an ultrasonic scaler for automated detection of different tooth substances. The experimental results obtained during test with simulated data demonstrated that the presented strategy is reliable and that classification-oriented modifications are worth considering, as they seemed to improve the efficiency of the classificator in discriminating between good and bad combinations of features. Experimental results obtained during test with real data were meant to assess the diagnostic ability of the system and showed promising results in exploiting unconventional techniques of feature construction by means of operators. In comparison with other approaches, the reported experiments showed an improvement in using non linear and derivative operators.

Foreseeable improvements and extensions may concern alternative neural network specifications (encoding, activation functions, training algorithm) and study of the influence of random processes on results. On the whole, the very promising results of this work state that investigation of unexplored techniques for feature extraction and classification is an exciting subject, which offers concrete possibilities for the implementation of new diagnostic and therapeutic instruments - particularly (but not exclusively) in periodontology practice.

REFERENCES

[Bis97] Bishop, C.M. (1997) *„Neural Networks for Pattern* Recognition", Clarendon Press Oxford.

[Cha91] Chang E.I., Lippmann, R.P (1991) *„Using Genetic Algorithms to select and create features for pattern recognition*", Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts.

[CRS98] Caruana,R. R.de Sa,V (1998) *„Using features: Selection to find inputs that work better as Extra Outputs"* , Proceedings of the 8th International Conference on Artificial Neural Networks (ICANN 1998).

[Fer02] Ferrero,F. (2002). *„Unconventional signal processing for dental surface pattern recognition - A bioinformatic, classification oriented approach*" Technische Universität Darmstadt, Sirona Dental System GmbH, Bensheim, Diplomarbeit.

[KRM97] T.Kocher, A. Rühling , H. Momsen , H.C. Plagmann (1997) *Effectiveness of subgingival instrumentation with power-driven instruments in the hands of experienced and inexperienced operators.* A study on manikins. J Clin Periodont 24, 498 - 504 (1997)

[Pao89] Pao, Y.H. (1989) *„Adaptative pattern recognition and neural networks"*, Addison-Wesley ISBN 0-201-12584-6 Chapters 1 (3-23),5 (111-140), Appendix A.

[Rip96] Ripley, B.D. (1996) *„Pattern Recognition and Neural Networks"*, Cambridge University Press.

[Sch01] Schubert, A. (2001): *Der Einsatz von intelligenten Methoden zur Diagnose von Maschinen- und Prozeßzuständen und ihre Integration in die Instandhaltungsstrategie chemischer Anlagen*, Dissertation, TU Clausthal.

[Seb99] Sebastian, M.T. (1999) *„Parodontologie, Stand der Wissenschaft"* (from the Annals of the American Academy of Periodontology, 1996), Phillip Verlag.

[SKB00] Strackeljan, J., Kocher, Th. , Behr D. (2000) *„Feasibility of computer assisted recognition of different dental hard tissues"*. Journal of Dental Research, Vol.79, No.3, 829-834.

[Str02] Strackeljan, J. (2002): *Einsatzmöglichkeiten von Softcomputing Methoden zur Auslegung, Optimierung und Konstruktion von Rotorsystemen.* Habilitationsschrift TU Clausthal, 2002.

[The89] Theilade, J. (1989) *„Dental plaque and dental calculus."* In: Textbook of clinical periodontology. Lindhe J. editor. 2.nd ed. Copenhagen: Munksgaard.

A Scheme for the Evolution of Feedforward Neural Networks using BNF-Grammar Driven Genetic Programming

Athanasios Tsakonas and Georgios Dounias
University of the Aegean, Dept. of Business Administration,
8 Michalon Str., 82100 Chios, Greece,
Phone: +30-271-35165, Fax: +30-271-93464
e-mail: tsakonas@stt.aegean.gr, g.dounias@aegean.gr

ABSTRACT: This paper presents our attempt to automatically define feedforward neural networks using genetic programming. Neural networks have been recognized as powerful approximation and classification tools. On the other hand, the genetic programming has been used effectively for the production of intelligent systems, such as the neural networks. In order to reduce the search space and guide the search process we employ grammar restrictions to the genetic programming population individuals. To implement these restrictions, we selected to apply a context-free grammar, such as a BNF grammar. The proposed grammar extends developments of cellular encoding, inherits present advances and manages to express arbitrarily large and connected neural networks. Our implementation uses parameter passing by reference in order to emulate the parallel processing of neural networks into the genetic programming tree individuals. The system is tested in two real-world domains denoting its potential future use.

KEYWORDS: genetic programming, context free grammars, cellular encoding, neural networks, thyroid, Pima Indians diabetes

INTRODUCTION

Genetic Programming has been proved nowadays as a flexible methodology and a valuable tool in data mining [1]. The prime advantage of genetic programming over genetic algorithms, is the ability to construct functional trees of variable length. This property enables the search for very complex solutions. These candidate solutions, in standard genetic programming, are usually in the form of a mathematical formula. This approach is commonly known as symbolic regression process. Later paradigms extended this concept to calculate any boolean or programming expression. Thus, complex intelligent structures, such as *fuzzy rule-based systems* or *decision trees* have already been used as the desirable intention in genetic programming approaches [2,3,4,5]. The main qualification of this solving procedure is that the feature selection, and the system configuration, derive in the searching process and do not require any human involvement. Moreover, genetic programming, by inheriting the genetic algorithms' stochastic search properties, does not use *local search* -rather uses the *hyperplane search*-, and so avoids driving the solution to any local minimum. The potential gain of an automated feature selection and system configuration is obvious: no prior knowledge is required and, furthermore, not any human expertise is needed to construct an intelligent system. Nevertheless, the task of implementing complex intelligent structures into genetic programming functional sets in not rather straightforward. The function set that composes an intelligent system retains a specific hierarchy that must be traced in the GP tree permissible structures. This writing offers two advantages. First, the search process avoids candidate solutions that are meaningless or, at least, obscure. Second, the search space is reduced significantly among only valid solutions. Thus, a *genotype* - a point in the search space- corresponds always to a *phenotype* - a point in the solution space. This approach -known as *legal searchspace handling method* [6]- is applied in this work using a context-free grammar. Implementing constraints using a grammar can be the most natural way to express a family of allowable architectures. While each intelligent system -such as a fuzzy system- has a functional equivalent -by means of being composed by smaller, elementary functions-, what defines and distinguishes this system is its grammar. Although mapping decision trees or fuzzy rule-based systems to specific grammars can be relatively easily implemented, the execution of massively parallel processing intelligent systems -such as the neural networks- is not forthright. In order to explore variable sized solutions, usually a kind of *indirect encoding* is applied. The most common one is the *cellular encoding* [7] in which a genotype can be realized as a descriptive phenotype for the desired solution. More specifically, within such a function

Intelligent e-Health Applications in Medicine
Special Session Proc. EUNITE 2002 Albufeira Portugal Sept. 19-21 2002
Published by the University of the Aegean

set, there are elementary functions that modify the system architecture together with functions that calculate tuning variables. Current implementations include encoding for feedforward and Kohonen neural networks [7,14] and fuzzy Petri-nets [8]. In his original work, Gruau also used a context-free grammar - a BNF[1] grammar- to encode indirectly the neural networks. On the other hand, in [8] a logic grammar - a context-sensitive one- is adapted to encode fuzzy Petri-nets. In our work, we show that as long as the depth-first execution of the program nodes of a GP tree is ensured -which is the default-, a context-free grammar such as a BNF grammar is adequate for expressing neural networks. Gruau's original work has been faced some skepticism [9] on the ability to express arbitrarily connected networks. Later developments [10] seem to offer less restrictive grammar, though the *cut* function[2] still maintained bounded effect. In our approach, we inherit present grammar advances proposed in [8] in his logic grammar for fuzzy Petri-nets and we suggest a BNF grammar for neural networks that is more descriptive than previous works. In the next section we present the design and the implementation of our system. The section that follows the design paragraph covers the results and our discussion in two real-world data sets: the Pima Indians diabetes data and the thyroid data. Finally, our conclusions are drawn in the last section as well as our suggestions for further research.

DESIGN AND IMPLEMENTATION

In order to incorporate the architecture of feedforward neural networks into genetic programming one has to decide whether he will use direct or indirect encoding. Direct encoding, although it describes fixed sized neural networks effectively, in most problem cases is inefficient while the *a-priori* knowledge of the neural network's best architecture is not available. Thus, it seems normal to prefer a learning process that uses variable sized neural networks such as cellular encoding, a variation of indirect encoding. Our work is consisted of the application of a BNF grammar to encode indirectly variable length feedforward neural networks. The idea behind cellular encoding is that each individual program in the population is a specification for developing a neural network. As it can be seen from *Figure 1*, to implement neural networks we consider three types of places. First, *input places*, such as N_1 and N_2, associated with the system input values during run. Second, *intermediate places*, such as N_3 and N_4, and third, *output places*, such as N_5 that represents the system outputs. The structure of a neural network for a particular problem must be restricted, in terms that if the problem requires a binary classification, then the network will have only two output places. The final output can follow the *winner-takes-all* concept -the output with the larger value is assumed as the system's output. *Table I* presents the manipulating functions. Manipulating functions initialize the inputs and/or insert additional places into the developing network. These functions are classified into two types, since there are two types of modifiable places, the input places and the intermediate places.

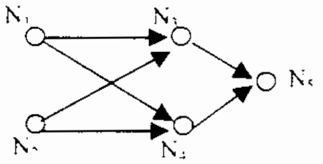


Figure 1. A simple feedforward neural network.

Table I. Manipulating functions

	Name	Description	Number of arguments
Input place	sp1	Sequential division	3
	pp1	Parallel division	2
	in	initialize the value	1
Intermediate place	sp2	Sequential division	3
	pp2	Parallel division	3
	stop	Terminate the modification	1
	lnk	Modify synapse (link)	4
	act	Activate	0

The implementation of a parallel processing system -such as neural networks- in a tree like the GP-tree that is executed *depth-first*, requires special handling of variables in order to emulate a *breadth-first* execution and simulate the parallel

---

[1] BNF: Backus-Naur-Form

[2] for the definition of the *cut* function see the paragraphs that follow.

processing. Thus we selected to apply to all functions a parameter *passing by reference* for two variables: a parameter array Q and a parameter value V. The array Q keeps synapse values and the value V mostly handles activation results. In our implementation the parallel execution of each individual is ensured, by the proper handling of these variables. In order to avoid confusion, we should define explicitly the meaning of the *arguments* shown in *Table I*. These arguments correspond to the permissible *children* a parent *GP-node* may have in a *GP tree*. On the other hand, the array Q and the variable V, correspond to the *internal implementation* of the neural network's elementary functions into the GP tree and they enable *variable sharing* throughout an individual's execution, that is needed to simulate the parallel processing. In the following paragraphs, a short presentation is given for each of the functions together with an explanation of the implementation. Each function calls its arguments, passing the array Q and the variable V by reference, unless otherwise stated. For example, function *pp2* creates copies of the (input) array Q before passing them to its arguments.

- The *sp1* function takes three arguments. The first and the third argument can be *pp1* or *sp1* (or *in*) functions. The second argument is a *weight value*. It calls sequentially the three arguments. Its application to the developing neural network is to add sequentially a node next to the node that is applied.
- The *pp1* function has two arguments. They can be *pp1* or *sp1* (or *in*) functions. It feeds the arguments with copies of the array Q. It then saves the concatenation of them to array Q. It does not affect explicitly the variable V. Its modification to the developing neural network is to create a node in parallel to the node that is applied.
- The *in* function has one argument. This argument is one of the network inputs. It initializes the array Q and the variable V to this value.
- The *sp2* function has three arguments. The first and the third argument can be *pp2* or *sp2* (or *stop*) functions. The second argument is a *weight number*. It calls sequentially the three arguments. Its application to the developing neural network is to add sequentially a node next to the node that is applied similar to *sp1*.
- The *pp2* function takes three arguments. The first and the third argument can be *pp2* or *sp2* (or *stop*) functions. The second argument can be a *lnk* or an *act* function. It feeds the three arguments with copies of the array Q. It then saves the concatenation of the first and the third argument to array Q. It does not affect explicitly the variable V. Its modification to the developing neural network is to create a node in parallel to the node that is applied, similar to *pp1*.
- The *stop* function takes one argument. This argument is a *bias number*. It adds the bias number to the value V and performs hyperbolic tangent activation. The result is saved to variable V. The array Q is also initialized with the value V.
- The *lnk* function takes four arguments. The first argument is a number that calculates the number of the synapse to be processed. The number derives by the application of the formula Z *mod* N, where Z is the number of existing synapses (kept in array Q) and N is the value of the argument. The second argument is a *weight number* and updates the weight of the selected synapse. The third argument is a *cut number* and cuts the selected synapse if and only if the cut value is 1 and the number of inputs to this place is greater than 1. The fourth argument can be again a *lnk* function or an *act* function.
- The *act* function summarizes the elements (synapse inputs) in the array Q and returns the result to V. It takes no arguments.

Table II. BNF grammar for the neural network

Symbol	Rule
<CLAUSE>	::= <ANN>
<ANN>	::=<PROG> <PROG>
<PROG>	::=<PLACE1><WT>
<PLACE1>	::=**SP1**<PLACE1><WT><PLACE2> | **PP1**<PLACE1><PLACE1> | <IN>
<IN>	::=<ATTR>
<WT>	::=float in [-1,1]
<PLACE2>	::= **SP2** <PLACE2><WT><PLACE2> | **PP2** <PLACE2><SYNAPSE><PLACE2> | <STOP>
<STOP>	::= <BIAS>
<SYNAPSE>	::= <LNK> | <ACT>
<BIAS>	::= float in [-1,1]
<LNK>	::=**LNK** <NUM> <WT> <CUT> <SYNAPSE>
<ACT>	::=**ACT**
<NUM>	::=integer in [1,256]
<CUT>	::=integer in [0,1]
<ATTR>	::=data attribute (system input)

As seen from the previous, there are additional functions that assist the selection of a number into a given range and a given precision. These functions are presented below.

The *weight number* function returns a float in the range [-1,1] with a precision of 0.00390625. The *bias number* function returns a float in the range [-1,1] with a precision of 0.000244140625. The *cut number* function returns an integer from the set {0,1}. The *number* function returns an integer in the range [1,256]. The selection of the precision for the *weight number* and the *bias number* is similar to the existing literature [10]. The BNF grammar production rules used for implementing the neural network are summarized in Table II. This example grammar corresponds to a binary decision neural network - with two independent outputs. A tree starts with the <CLAUSE> symbol. As it can be seen from the design of this grammar, there is no limit on how many *cut* functions could be applied in a node. This is an upgrade from previous implementations - where the number of effective *cut* functions was limited - and enables to a larger set of neural networks to be expressed within the search process.

RESULTS AND DISCUSSION

The model is tested in two data sets from the medical domain. These data sets have been taken unmodified by a collection of real-world benchmark problems, the *Proben1* [11] that has been established for neural networks. Table III shows the problem complexity of these data sets. Each data set is separated into a training set, a validation set and a test set. The training set is consisted of 50% of the data and the rest 50% is divided equally between the validation set and the test set. During the training phase, the validation set is used to avoid overfitting. For example, a solution which has better classification score in the training set, is adapted as new best solution if and only if the sum of classification scores of both training and validation sets is the same or better than the best solution's respective score. The genetic programming parameter settings are presented in Table IV. For comparison reasons, we selected to keep the same settings in both problems.

Table III. Problem complexity for two real-world benchmarking data sets

Problem	**Attributes**	**Inputs**		**Classes**	**Examples**
		continuous	*discrete*		
diabetes	8	8	0	2	768
thyroid	21	6	15	3	7200

Table IV. Genetic Programming Settings

Parameter	**Setting**
Population Size	4000 individuals
GP implementation	Steady-state grammar driven GP
Selection	Tournament of 6 with elitist strategy
Maximum number of generations	200
Crossover probability	65%
Mutation probability (overall)	35%
Shrink mutation probability (relative to overall)	40%
Node mutation probability (relative to overall)	30%
Number mutation probability (relative to overall)	30%
Maximum size of individual	650 nodes
Function Set and Terminal Set	See *Table I* and *Table II*
Fitness function	Sum of correct classifications

PIMA INDIANS DIABETES DATA

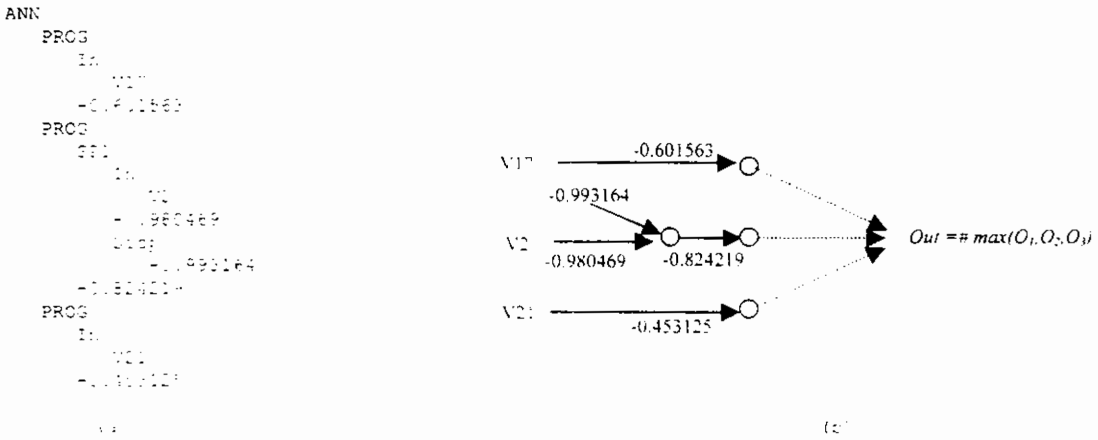
The first data we apply the system is the Pima Indians diabetes data. It is a binary classification problem. The aim is to diagnose whether or not a patient has diabetes. The class probabilities are 65% and 35% for no diabetes and diabetes correspondingly. The attributes are presented in *Table V*. Although there are not missing values in this data, there are several senseless zero values that introduce some errors into the dataset. In [12], a linear genetic programming approach is applied, and for the same dataset *-diabetes1-* an average classification score of 76.04% is referred for the test set. In [11], this average classification score for the test set is 75.9% using neural networks trained with RProp [13]. The solution presented in *Appendix*, was achieved after 92,000 iterations (23 generations) and has accuracy 76.96% (147/191) in the test set. The accuracy in training and validation sets is 83.76% and 70.68% respectively. Although the results seem very promising, extended experimentation has to be completed in order to obtain average training results.

Table V. Diabetes attributes

Name	Attribute
T1	Number of times pregnant
T2	Plasma glucose concentration a 2 hours in an oral glucose tolerance test
T3	Diastolic blood pressure (mm Hg)
T4	Triceps skin fold thickness (mm)
T5	2-Hour serum insulin (mu U/ml)
T6	Body mass index (weight in kg/(height in m)^2)
T7	Diabetes pedigree function
T8	Age (years)

THYROID DATA

The second data we selected to apply our model is the thyroid database. The aim is to diagnose overfunction, normal function, or underfunction of the thyroid. The class probabilities are 5.1%, 92.6% and 2.3% respectively; entropy is 0.45 bits per example. There are 21 attributes, 6 of them of continuous value and 15 of them of discrete (0 or 1) value. A good classifier has to have accuracy better than 92% [11]. The solution presented in *Figure 2* was accomplished after 4,000 iterations (1 generation) and has accuracy on the test set 94.44% (1699/1799). The classification score in the training set and the validation set is 93.8% (3375/3598) and 93.71% (1686/1799) accordingly. As it can be observed, even very early results in the search process of this problem, may produce systems with satisfying accuracy. This result verifies that this data may be classified adequately by linear systems [11]. It is seen also, that a relatively small number of attributes (3 out of 21) are enough for obtaining such classification accuracy. The same dataset has been used as benchmarking set in literature. In [12], the average classification score, using linear genetic programming, reaches, for the test set, the 98.09%. In [11], a neural network application, using RProp, achieves 97.62% average accuracy on the test set.



```
ANN
    PROG
        IN
            V17
        -0.601563
    PROG
        [illegible]
            IN
                V2
            -0.980469
            [illegible]
                -0.993164
        -0.824219
    PROG
        IN
            V21
        -0.453125
```

Figure 2. Solution after first generation (4,000 iterations): (a) program individual, (b) network representation.

CONCLUSIONS AND FUTURE WORK

In this work, we presented our approach to the evolutionary development of feedforward neural networks using genetic programming. While our intention was to enable arbitrarily large and arbitrarily connected neural networks, we selected to implement an indirect encoding variation for the neural networks, the cellular encoding. The latter is expressed, within this work, using a BNF grammar that considers current grammar advances found in literature. Our implementation makes use of parameter passing by reference among genetic programming functions, in order to simulate the parallel processing of neural networks. Two real-world datasets from the medical domain are used to test the system. These two databases were preferred since they are characterised by different type of complexity. The first set, the Pima Indians diabetes data, represents a binary classification task and is assumed to involve noise. The second set, the thyroid data, is a classification between three classes, it has a relatively large number or records and classification scores found in literature are high. Our first results for the Pima Indians diabetes data seem very promising and highly comparative to those found in literature. For the second data, the thyroid database, very early results are considered good, though they are definitely less accurate than those found in literature. However, more experimentation has to be accomplished in both domains, in order to secure transparent results.

Further research should be accomplished in the testing on different domains. A number of separate runs should be included for each data in order to attain average classification scores. Moreover, within the same data set testing, different subsets for the training, validation and test sets should be considered, in order to avoid any data selection bias. On the other hand, the system should be tested with various genetic programming parameter settings. A comparison with neural networks derived by context-sensitive grammars [14] could present valuable results. Finally, a comparison with other genetic programming-derived intelligent systems -such as fuzzy rule-based systems and decision trees- could offer a clear view on the efficiency of this approach among similar ones.

REFERENCES

[1] Koza J.R., Bennett III F.H., Andre D., Keane M.A., "Genetic Programming III", Morgan Kaufmann Publ., Inc.,1999

[2] E.Alba, C.Cotta, J.M.Troya, "Evolutionary Design of Fuzzy Logic Controllers Using Strongly-Typed GP", In *Proc. 1996 IEEE* Int'l Symposium on Intelligent Control., pp. 127-132. New York, NY., 1996

[3] Koza J.R., "Genetic Programming: On the Programming of Computers by Means of Natural Selection", Cambridge, MA, MIT Press., 1992

[4] Tsakonas A., Dounias G., "Hierarchical Classification Trees Using Type-Constrained Genetic Programming", in Proc. of 1^{st} Intl. IEEE Symposium in Intelligent Systems, Varna, 2002

[5] Tsakonas A., Dounias G., Axer H., von Keyserlingk D.G., "Data Classification using Fuzzy Rule-Based Systems represented as Genetic Programming Type-Constrained Trees", in Proc. of the UKCI-0, 2001, pp 162-168

[6] Yu T., Bentley P., "Methods to Evolve Legal Phenotypes". In Lecture Notes in Comp.Science 1498, Proc. of. Parallel Problem Solving from Nature V, 1998, pp 280-291

[7] Gruau F., "Neural Network Synthesis using Cellular Encoding and the Genetic Algorithm", Ph.D. Thesis, Ecole Normale Superieure de Lyon, anonymous ftp:lip.ens-lyon.fr (140.77.1.11) pub/Rapports/PhD PhD94-01-E.ps.Z

[8] Wong M.L., "A flexible knowledge discovery system using genetic programming and logic grammars", Decision Support Systems, 31, 2001, pp 405-428

[9] Hussain T., "Cellular Encoding: Review and Critique", Technical Report, Queen's University, 1997, http://www.qucis.queensu.ca/home/hussain/web/1997_cellular_encoding_review.ps.gz

[10] Gruau F., Whitley D., Pyeatt L., "A Comparison between Cellular Encoding and Direct Encoding for Genetic Neural Networks", in Koza J.R., Goldberg D.E., Fogel D.B., Riolo R.L. (eds.), Genetic Programming 1996: Proceedings of the First Annual Conf., pp 81-89, Cambridge, MA, 1996, MIT Press

[11] Prechelt L., "Proben1 - A set of neural network benchmark problems and benchmarking rules", Tech.Rep. 21/94, Univ. Karlsruhe, Karlsruhe, Germany, 1994

[12] Brameier M., Banzhaf W., "A comparison of Linear Genetic Programming and Neural Networks in Medical Data Mining", in IEEE Trans. on Evol.Comp., vol.5,no.1, Feb 2001,pp 17-26

[13] Riedmiller M., Braun H., "A direct adaptive method for faster backpropagation learning: The RPROP algorithm", in Proc. IEEE Intl. Conf. Neural Networks, San Francisco, CA,1993,pp 586-591

[14] Hussain T., Browse R., "Attribute Grammars for Genetic Representations of Neural Networks and Syntactic Constraints of Genetic Programming", in AIVIGI'98:, Workshop on Evol.Comp., Vancouver BC, 1998

APPENDIX

```
(ANN (PROG (PP1 (In T8) (PP1 (In T3) (SP1 (PP1 (In T2) (PP1 (PP1 (PP1 (PP1 (In T6) (In T6)) (In T7))
(SP1 (PP1 (In T2) (PP1 (In T1) (PP1 (In T6) (PP1 (PP1 (In T6) (SP1 (PP1 (In T2) (PP1 (PP1 (In T6)
(PP1 (PP1 (PP1 (PP1 (In T6) (In T6)) (In T7)) (SP1 (PP1 (In T2) (PP1 (In T1) (PP1 (In T6) (PP1 (SP1
(PP1 (In T2) (PP1 (In T2) (In T7))) -0.078125 (PP2 (PP2 (PP2 (Stop -0.976563) ACT (PP2 (Stop -
0.972412) ACT (Stop -0.951660))) ACT (Stop -0.975830)) ACT (PP2 (Stop -0.992676) ACT (PP2 (Stop -
0.972412) ACT (Stop -0.951660))))) (PP1 (In T1) (SP1 (In T8) -0.773438 (Stop -0.992676)))))) -
0.078125 (PP2 (Stop -0.999756) ACT (Stop -0.999756)))) (PP1 (In T1) (SP1 (In T8) -0.773438 (Stop -
0.992676))))) (In T7))) -0.078125 (PP2 (Stop -0.975830) ACT (Stop -0.976563)))) (PP1 (In T1) (SP1
(PP1 (In T6) (PP1 (In T3) (SP1 (PP1 (PP1 (In T1) (PP1 (In T6) (PP1 (SP1 (PP1 (In T2) (PP1 (PP1 (In
T2) (PP1 (In T2) (PP1 (In T8) (PP1 (SP1 (PP1 (In T2) (PP1 (In T1) (PP1 (In T6) (PP1 (SP1 (PP1 (In
T2) (PP1 (PP1 (In T1) (In T8)) (In T7))) -0.078125 (PP2 (PP2 (PP2 (Stop -0.976563) ACT (PP2 (Stop -
0.972412) ACT (Stop -0.951660))) ACT (Stop -0.975830)) ACT (Stop -0.951660))) (SP1 (PP1 (In T2) (PP1
(In T2) (In T7))) -0.078125 (PP2 (Stop -0.975830) ACT (Stop -0.951660)))))) -0.078125 (PP2 (Stop -
0.992676) ACT (Stop -0.992676))) (PP1 (In T1) (SP1 (In T7) -0.773438 (Stop -0.992676)))))) (In
T7))) -0.078125 (PP2 (Stop -0.976563) ACT (PP2 (Stop -0.972412) ACT (Stop -0.958252)))) (PP1 (In T1)
(SP1 (In T8) -0.773438 (Stop -0.992676)))))) (PP1 (PP1 (PP1 (PP1 (In T6) (In T6)) (SP1 (PP1 (In T2)
(PP1 (In T2) (In T7))) -0.078125 (PP2 (PP2 (PP2 (PP2 (Stop -0.972412) ACT (Stop -0.951660)) ACT (PP2
(Stop -0.972412) ACT (Stop -0.951660))) ACT (Stop -0.975830)) ACT (Stop -0.958252)))) (SP1 (PP1 (In
T2) (PP1 (In T1) (PP1 (In T6) (PP1 (SP1 (PP1 (In T2) (PP1 (In T1) (PP1 (In T6) (PP1 (SP1 (PP1 (In
T2) (PP1 (In T2) (In T7))) -0.078125 (PP2 (Stop -0.975830) ACT (Stop -0.951660))) (PP1 (In T1) (SP1
(In T8) -0.773438 (Stop -0.992676))))))) -0.078125 (PP2 (Stop -0.992676) ACT (Stop -0.992676))) (PP1
(In T1) (SP1 (In T7) -0.773438 (Stop -0.992676)))))) -0.078125 (PP2 (Stop -0.999756) ACT (PP2 (Stop
-0.999756) ACT (Stop -0.999023))))) (In T5))) -0.078125 (PP2 (Stop -0.999756) ACT (Stop -
0.992676))))) -0.773438 (Stop -0.992676))))))) -0.078125 (PP2 (Stop -0.999756) ACT (Stop -
0.999023)))) (In T7))) -0.078125 (PP2 (Stop -0.951660) ACT (Stop -0.992676))))) -0.019531) (PROG
(PP1 (SP1 (PP1 (In T3) (In T6)) -0.859375 (PP2 (Stop -0.972412) ACT (Stop -0.951660))) (In T3)) -
0.625000))
```

Figure 3. Neural network description for the diabetes data.

Fuzzy-Controlled Glaucoma Monitoring Based on Arden Syntax

Sven Tiffe*
Siemens Medical Solutions, Business Development/Knowledge-Based Biomedical Services
Henkestraße 127, 91052 Erlangen, Germany
and
Department of Medical Computer Sciences, Section on Medical Expert and
Knowledge-Based Systems, University of Vienna Medical School
Spitalgasse 23, 1090 Vienna, Austria

Abstract

Monitoring glaucoma-related changes within the eye status of a patient requires the following steps: a) to decide whether the patient's ophthalmic data set is indicative of critical or suspicious situations by way of differential diagnosis, and b) detect changes in the status over time. A knowledge base founded on an artificial neuronal network and a set of fuzzy control rules was defined in order to aid in differential diagnosis in this setting. The fuzzy control rules were implemented using an extended Arden Syntax, and are currently being used in a glaucoma monitoring telemedicine project.

KEYWORDS: Arden Syntax for Medical Logic Systems, medical expert systems, knowledge representation, fuzzy logic, glaucoma classifier, telemedicine

Introduction

According to the World Health Organization, glaucoma is one of the three major causes for blindness worldwide [1]. Early detection of glaucomatous changes in the eye status may help in the prevention of a significant risk factor. This purpose can be achieved by measuring the ophthalmic parameters of a patient, which are monitored by an expert system that classifies the data sets and generates alerts when registering data that indicate a suspicious or critical status of the eye.

One means of classifying the eye status of a patient is to monitor the intraocular pressure (IOP), which is frequently elevated in patients with glaucoma, as the ability of the eye to drain the intraocular fluid is reduced. Whenever the current IOP value exceeds a threshold value (the normal eye maintains an internal pressure of 12 to 22 mm of mercury), such a system would issue a warning. However, a more sophisticated classifier would include additional parameters in the classification process.

Structure of the classifier

In recent work, a knowledge base for such an extended classifier was defined on the basis of an artificial neuronal network (ANN) and a fuzzy rule set (FRS) [2]. The ANNs generate a classification of perimetry data sets, which is then used as one input source for the FRS during the monitoring process.

Perimetry data describe the status of the visual field of the patient and are measured by perimeter devices that detect the loss of light sensitivity at various stimulus points of the retina. As the patient presses a response button to report whether or not he sees a stimulus, the results are subjective and generally uncertain. The ANNs that classify perimetry data were trained with training sets defined by four medical experts.

Further input values for the FRS are estimations of papilla description parameters. These are measured by observing the papilla during an eye examination, and consist of the cup-disc-ratio (CDR), the location of the excavation, and the difference between the CDR values of the two eyes. Whereas the CDR and the difference are given as real numbers, the location is given in linguistic terms such as 'central', 'inferior', or 'superior'.

The last parameter that is used as input for the FRS is the intraocular pressure. This value is measured in constant time ranges, such as daily or weekly. The remaining parameters are measured once or twice a year at the beginning of a monitoring period, and are stored in the data base. Thus, perimetric classification by the ANNs is done after each new measurement, whereas the fuzzy rules have to be evaluated each time the IOP is measured.

---

[*] e-mail: Sven.Tiffe@siemens.com

Intelligent e-Health Applications in Medicine
Special Session Proc. EUNITE 2002, Albufeira, Portugal, Sept 19-21, 2002
Published by the University of the Aegean

Fuzzy production rules

The rule sets used for the classification of glaucoma are structured in a similar way as fuzzy control rule sets. Such a set usually consists of one or more linguistic variables as input values, a set of production rules, and one or more linguistic variables as output values (see figure 1).

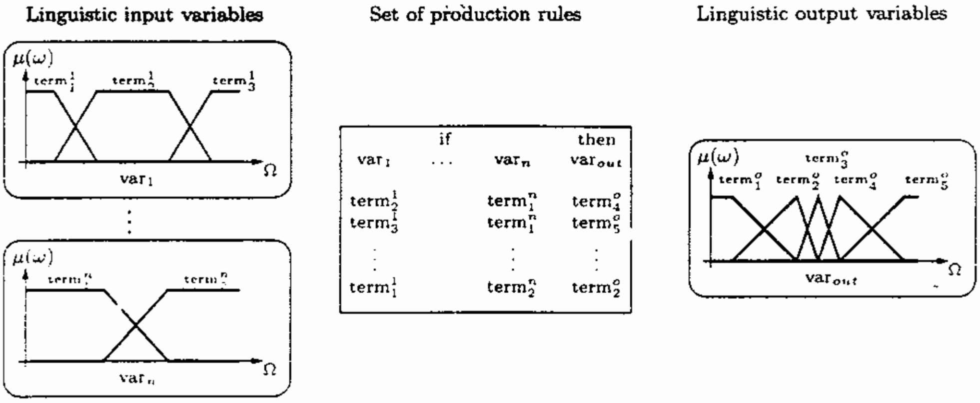


Figure 1: Structure of fuzzy control rule sets

A linguistic variable defines characteristics (values) of an abstract concept, such as 'normal' or 'increased' intraocular pressure. Mathematically, a linguistic variable can be defined by the quintuple

$$V_L = \{X, T, \Omega, G, B\} \tag{1}$$

where X is the name of the variable, T a set of terms representing the values of X, Ω the universe of discourse, G a set of syntactical rules which generate T, and, finally, B a set of semantic rules which define the linguistic discretization of Ω.

For example, intraocular pressure may be defined as $X = \text{IOP}$, $T = \{\text{normal}, \text{increased}\}$, and $\Omega = [0, 70]$[1]. The linguistic discretization B is defined by two fuzzy sets, which define the relationship of $t \in T$ to Ω (figure 2). A fuzzy set is characterized by a compatibility function which defines the degree of compatibility of $\omega \in \Omega$ to a term: a degree of 0.0 indicates no compatibility whether a degree of 1.0 indicates full compatibility [3].

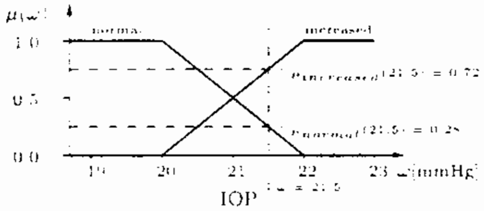


Figure 2: Linguistic variable 'IOP' (intraocular pressure)

Fuzzy rule inference

Before a numeric value can be used as input for a production rule, it has to be represented by a linguistic variable. Usually, for every term $t \in T$ the degree of compatibility to a given numerical value is determined by applying the value to the corresponding fuzzy set. This step is known as the *fuzzification* of the input value. Based on the fuzzy sets defined in the last example and a numeric input value of $\omega = 21.5$, the linguistic variable 'IOP' has the value 'normal' by a degree of 0.28 and the value 'increased' by a degree of 0.72.

As the next step the, input variable is used by the set of production rules as condition on the left side. Every rule consists of a condition and a conclusion. The conditional expression can be made from one or more comparisons of a linguistic variable and a linguistic value. If the condition consists of more than one comparison, they may be combined using logical operators such as "and", "or", and "not".

---

[1] In this example, G is not defined separately.

The conclusion usually consists of one or more assignments of a value to a linguistic variable. The degree of truth of the condition determines its degree of validity. If the conditional expression is only partially fulfilled. the output variable also partially set to the value.

Thus, a resulting linguistic variable may have one or more values, each by a certain degree. As the result of the glaucoma classifier represents different characteristics of a diagnosis ('normal', 'pathological', 'suspicious', or 'glaucomatous'), it can be communicated in textual form as a description of the single values.

A more detailed introduction to fuzzy set theory and fuzzy control can be found in the literature [3].

Fuzzy Arden

The knowledge base of the FRS is represented by an extended version of the Arden Syntax for Medical Logic Systems. The Arden Syntax is a standard for the representation of medical knowledge and is maintained by the Health Level Seven (HL7) organization [4]. An Arden Syntax knowledge base consist of individual. independent rules known as Medical Logic Modules (MLMs). Each MLM is structured by three categories that contain entries ("slots"). They provide information about the author and the status ("maintenance" category), explanatory information, links to external knowledge sources ("library" category), and the necessary decision logic ("knowledge" category). In particular, the last category includes information about data base access ("data" slot), separated from information about the decision making process ("logic" slot).

In recent work, we defined extensions that can be used to represent fuzzy truth values and to model fuzzily defined conditions [5]. Further extensions allow the definition of linguistic variables as Arden Syntax MLMs [6]. Such an MLM provides in the knowledge category, instead of a decision logic, the name of the variable, the values, the fuzzy sets that define them, and some additional information such as the universe of discourse or optional data sources for the fuzzification of the variable.

Linguistic variable MLMs are referenced by other MLMs in the data slot. The reference statement can be used to define whether the variable is meant to be an input variable and whether it has to be initialized (fuzzified). The variable may be used in the decision slot, for instance within linguistic conditional expressions.

Technical structure of the monitoring application

The Arden Syntax knowledge base is used by a Java based Arden Syntax rules engine that is connected to the Siemens MedStage[2] communication platform. Patients participating in the monitoring program have to measure their IOP at home using a tonometric device with an attached communication unit. This unit stores the measured values as well as the time stamps and communicate them to the MedStage server, where the values are stored in the data base. Data base triggers fire an event that is communicated to the rules engine. The rules engine evokes an Arden Syntax MLM that reads the new IOP values as well as all other needed input values and evokes the fuzzy rules. Finally, a message is generated and communicated by the message server.

Methods

The fuzzy rules set was initially modelled by using the FuzzyTech[3] software, which is a commercial tool for the definition and evaluation of fuzzy control rule sets. The knowledge base can be saved as an ASCII text file in a proprietary format and can additionally be converted into a program code, such as a Java class file, which may then be directly used by a software application.

Using the program code representation of the FRS would imply linking the code directly to the program, which therefore would have to be recompiled every time the rules are altered. A more flexible way to implement the rules is to use the Arden Syntax rules engine. We generated a set of Arden Syntax MLMs, which represent the FRS, and used them for the glaucoma monitoring program.

The FuzzyTech file format is based on a structured text format. We created a parser based on JLex[4]/Jay[5], which directly generates MLMs from a FuzzyTech project file. The parser has generated a total of 28 MLMs.

Linguistic variables are represented by 17 MLMs, including seven input variables, one overall output variable, and nine intermediate variables. Each set of production rules is represented by its own MLM; in all there are

---

[2] http://www.medstage.com

[3] http://www.fuzzytech.com

[4] http://www.cs.princeton.edu/~appel/modern/java/JLex/

[5] http://www.informatik.uni-osnabrueck.de/bernd/jay/

11 production rules MLMs. Although the MLMs were created automatically by the parser, some linguistic variable MLMs had to be modified and two additional control MLMs had to be written. One such linguistic input variable is shown in figure 3.

```
maintenance:
  title: linguistic variable IOP;;
  mlmname: LV_IOP;;
  arden: Version 2flv;;
  version: 1.0;;
  institution: Siemens Medical Solutions;;
  author: Sven Tiffe;;
  specialist: ;;
  date: 2002-01-03;;
  validation: testing;;

library:
  purpose: This linguistic variable represents an input value of
           the TOSCA glaucoma classification rules.
           It is used as input by the rule blocks: RB2, RB3, RB4.;;
  explanation: ;;
  keywords:  ;;
  citations: ;;
  links:  ;;

knowledge:
  type: linguistic variable;;
  values: 'normal', 'increased';;
  input: read { %event.iop:date% };;
  range: 0.0, 70.0;;
  unit: 'mmHg';;
  sets:
    'normal'    := linear((0.0, 1.0), (20.0, 1.0), (22.0, 0.0), (70.0, 0.0));
    'increased' := linear((0.0, 0.0), (20.0, 0.0), (22.0, 1.0), (70.0, 1.0));;
end:
```

Figure 3: Arden Syntax linguistic variable MLM: IOP

Modifications of the linguistic variable MLMs affected the fuzzification of the seven input variables. As the FuzzyTech project file does not provide any information about data sources for the numerical input values, all necessary information about data base queries had to be added to the input slots.

The intermediate linguistic variables are the result of intermediate production rule sets. Thus, the input slot is already generated by the parser, as the variables simply reference those MLMs that represent the intermediate production rule sets in their input slots. The final result of the FRS is one output variable that is defined analogously.

In an Arden Syntax fuzzy control knowledge base structured like the present one, the classifier can be called in two ways, namely by backward and forward inferencing.

The backward inference process is started by referencing the final output variable and initializing it automatically[6]. The variable would then call the production rule MLMs that are referenced in the input slot. These MLMs would initialize their linguistic variables, which would have to resolve their input slots, and so on. Those input variables that are not the result of intermediate rule sets have to define a data base query to retrieve the initial numerical input values.

The forward inference process is controlled by an additional MLM that first has to query all required data from the data base and initialize the input variables manually by passing the numerical values as arguments. Then, every production rule MLM is evoked with input variables as arguments. These MLMs return the intermediate variables which are used as input variables for the next production rules MLMs, and so on.

---

[6] This presumes, that every linguistic variable defines the input slot

As some input parameters had to be prepared before we could use them for the linguistic variables, we used such additional MLMs for forward inferencing. Each ophthalmic data set is classified twice, first based on values of the right eye, then based on values of the left eye. If the patient only measured the IOP of one eye, the contralateral data set is not classified. The result is coded in an XML message, which can be used to generate a text message.

Results and discussion

Medical aspects

The evaluation of the Arden Syntax classifier was based on the data sets of 31 patients who were considered suspicious and were selected by a medical expert. The data sets were classified by the FuzzyTech software and the Arden Syntax-based; the performance of the two systems was identical in terms of results.

If the intraocular pressure was within the range that is considered to be normal (up to 21 mmHg), the fuzzy rules correctly classified the data as non-glaucomatous. The test data comprised 13 such cases; three of them strongly indicated a pathological state of the eye, which however is not glaucomatous.

Further, the test data included 11 cases in which the threshold-based classifier would have generated a serious warning message due to a significantly increased IOP value. However, only four of them were rated glaucomatous and three of them were classified as being not very suspicious. Even when the IOP was increased, the perimetry data were rated to be nearly normal.

The IOP values of the remaining cases were in the range of 21mmHg to 25mmHg and were rated suspicious to different degrees. In particular, six cases that were close to the lower threshold of 21mmHg were rated suspicious *and* glaucomatous, and would not have been classified in this manner by a simple IOP threshold classifier.

In summary, perimetry data and CDR parameters may be additionally used to detect glaucoma related changes in intraocular pressure from other pathological influences on this parameter. Additional pathological states can also be detected and communicated to the patients or physicians.

Technical aspects

The fuzzy control rule sets of the glaucoma classifier have been successfully represented by an extended version of the Arden Syntax. Every linguistic variable and every production rule set has been represented by its own MLM. Two control MLMs read data from the data base, initialize the input variables, and control the inference process. As a result, a textual message that includes the results for both eyes is generated.

The performance of the Arden Syntax-based system is good enough to return results instantly, even if it requires a little more time to evaluate the data compared to the generated native Java code. This advantage of the native approach has to compete the flexibility of a system based on Arden Syntax in terms of maintaining the knowledge base and to linking it to the running decision support system.

Every time the rules are altered, they have to be imported or linked to the system. Both representations had to be modified and individualized manually, as data base queries and evocation procedures had to be implemented separately. In contrast to the classifier based on a program code, when using Arden Syntax MLMs no programmer was needed to integrate the rules into the system and no program code had to be directly linked into the information system. In other words, the knowledge base of a telemedicine system can be dynamically modified without stopping, recompiling, or restarting the software.

The reusability of the rules and linguistic variables is improved by the modular representation in Arden Syntax. The use of the linguistic variable MLMs is not limited to this particular classifier or fuzzy control production rules in general. Common linguistic variables, such as body temperature or blood pressure, could be used in fuzzy control rules as well in classic Arden Syntax rules to improve the readability of the code, by encouraging the author to use linguistic expressions.

However, the general overview of complex fuzzy rule sets is easier to comprehend when a graphical user interface, such as the one supplied with the FuzzyTech software, is used. Thus, a graphical environment for developing Arden Syntax MLMs could not only improve the development of Medical Logic Modules in general, but also that of complex Arden Syntax-based fuzzy rules in particular.

Conclusion

When a large number of patients have to be monitored for glaucoma-related changes in their eye status, the fund of data tends to increase rapidly due to the ongoing measurement of IOP values for each patient. Our classifier can assist the physician in daily routine by sending alerts in the event of critical or suspicious eye states in their patients.

The fuzzy rule based classifier has some advantages over a simple classifier that only monitors the IOP and uses crisp thresholds. Compared to such a classifier, the fuzzy rules classify the data sets more precisely and are additionally able to detect non-glaucomatous states with an increased IOP as well as glaucomatous states with IOP values close to the normal range. The use of linguistic variables based on the concept of fuzzy sets and fuzzy production rules avoids unintuitive changes of the classification results of borderline cases.

The explicit representation of knowledge by the Arden Syntax based rules engine has certain advantages over systems that represent the knowledge implicitly in their program code. It provides flexibility in terms of adding, removing, or modifying knowledge without having to modify the system.

The extension used in this project is currently being discussed with the Arden Syntax Special Interest Group of HL7.

Acknowledgements

I gratefully acknowledge Prof. Dr. Klaus-Peter Adlassnig from the Department of Medical Computer Sciences, Section on Medical Expert and Knowledge-Based Systems, University of Vienna Medical School, and Dr. Gudrun Zahlmann from Siemens Medical Solutions BD KBS, for their contributions to this work on theoretical and practical fuzzy issues.

References

[1] World Health Organization. *The world health report.* World Health Organization, Avenue Appia 20, 1211 Geneva 27, Switzerland, 1998.

[2] G. Zahlmann, M. Scherf, and A. Wegner. A neuro-fuzzy-classifier for a knowledge-based glaucoma monitor. In *Artificial Intelligence in Medicine*, pages 273–284, 1997.

[3] G.J. Klir and T.A. Folger. *Fuzzy sets, uncertainty, and information.* Prentice Hall, London, 1988.

[4] Health Level Seven. *Arden syntax for medical logic systems.* Health Level Seven, Inc., 3300 Washtenaw Ave, Suite 227, Ann Arbor, MI 48104, 1999.

[5] S. Tiffe and K.-P. Adlassnig. Representation of medical knowledge by Arden Syntax with fuzzy extensions. In K.-P. Adlassnig, editor, *Intelligent Systems in Patient Care*, pages 80–88. Österreichische Computer Gesellschaft, Wien, 2001.

[6] S. Tiffe. Defining medical concepts by linguistic variables with Fuzzy Arden Syntax. *Submitted to AMIA*, 2002.

Application of Learning Vector Quantization to detect drivers dozing-off

David Sommer, Thilo Hink, Martin Golz
University of Applied Sciences Schmalkalden,
Department of Computer Science
D-98574 Schmalkalden, Germany
{sommer, golz}@informatik.fh-schmalkalden.de

ABSTRACT: EEG-signals and EOG-signals of eleven subjects were recorded during an overnight driving simulation task. By scoring the recorded videos clear microsleep events and clear non-microsleep events were picked out and small segments of the measured EEG- and EOG-signals before and during the events were analyzed. The spectral densities of these segments were classified using three methods of Learning Vector Quantization. Best classification results, up to 91%, were obtained with inclusion of all used EEG and EOG channels.

KEYWORDS: EEG, EOG, Microsleep, Learning Vector Quantization

INTRODUCTION

Reliable methods to discriminate dozing-off events from continuously measured signals of a driver will be an important milestone in the development of drowsiness warning systems. Dozing-off periods are characterized by sudden intrusions of sleep into wakefulness, lasting for 2 to 30 sec [1], and often called as microsleep events (MSE). As pointed out by several authors, it is very difficult to get a drowsiness measure; for a review see [2]. On the other hand, drowsiness measures are often calculated on a minute-scale, while signal processing to discriminate MSE has to be on a scale of seconds. As stated in [2], the only method that measures continuous fluctuations of sleepiness should be polysomnography, mainly by analyzing the electroencephalogram (EEG) and the electrooculogram (EOG). Later on, some success was made by pupillography [3-5]. By processing the pupil diameter signal, only, a discrimination performance of over 80% is achievable, as recently reported [6].
In this paper we are regarding the following questions:

1) Do small segments of EEG and of EOG immediately before and during a MSE contain enough information to discriminate them from segments during non-microsleep episodes (NMSE), when the driver is drowsy but still attentive?
2) Is it possible to detect MSE without a noticeable decrement by analyzing only one channel of the multi-channel EEG- /EOG- recordings?
3) Which segment length and which time offset related to the MSE starting time are optimal?

EXPERIMENT

Our experimental setup is comparable to [7]. Four EEG-signals and two EOG-signals were recorded of eleven young subjects during driving simulation sessions lasting 25 min and repeated every hour between 1 a.m. and 7 a.m. Two video cameras were utilized to record drivers portrait and right eye region for visual off-line scoring of MSE, typically recognizable by closed eyes or by drop of the head. Scoring was performed by two experienced persons under the guideline to take only undoubtable cases into account. Disadvantageously, attention losses with open eyes and with stare gaze are not detectable by this scoring method. The number of MSE was very different between subjects and was increasing with time of day for all subjects. All in all 1,675 MSE and 1,286 NMSE were scored.

Intelligent e-Health Applications in Medicine
Special Session Proc. EUNITE 2002, Albufeira Portugal, Sept 19-21, 2002
Published by the University of the Aegean

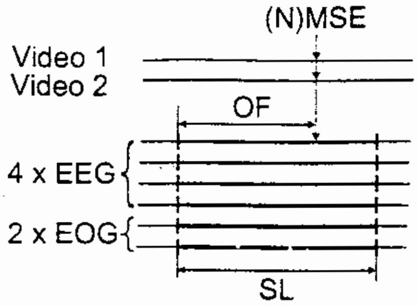


Figure 1: Segmentation of EEG- / EOG- signals with relation to visually scored microsleep- (MSE) or nonmicrosleep- (NMSE) events by two parameters offset (OF) and segment length (SL).

ANALYSIS

EEG and EOG were lowpass filtered with 40 Hz cut-off frequency and sampled at 64 Hz. After synchronization between both video recordings and EEG- / EOG- recordings, segments were stored for further processing using variable time offsets (OF) and variable segment length (SL) (Fig. 1).
After linear trend removal and applying Hanning window, spectral estimation was done by discrete Fourier transform. Spectral power densities were averaged by summing up in frequency bands of width of 0.5 Hz, 1 Hz and 2Hz and in conventional bands of EEG research: delta (0.5 ... 3.9 Hz), theta (4.0 ... 7.9 Hz), alpha (8.0 ... 11.9 Hz), sigma (12.0 ... 13.9 Hz) and beta (14.0 ... 29.9 Hz).

DISCRIMINATION ANALYSIS

For each feature vector consisting of averaged absolute spectral power densities of all six signals a label "MSE" or "NMSE" was scored, thus we have a two-class classification problem. Analysis was done by Learning Vector Quantization (LVQ) [8]. All three methods OLVQ 1, LVQ 2.1 and LVQ 3 were processed sequentially preceded by an initialization phase [9]. The learning set was partitioned into 80% training set and 20% test set. Partitioning was repeated 50 times for each parameter setting of the LVQ network. The classification rate was estimated as the ratio between the number of correct classifications to the number of all classifications using the test set. This kind of discrimination analysis is comparable to that in [10].

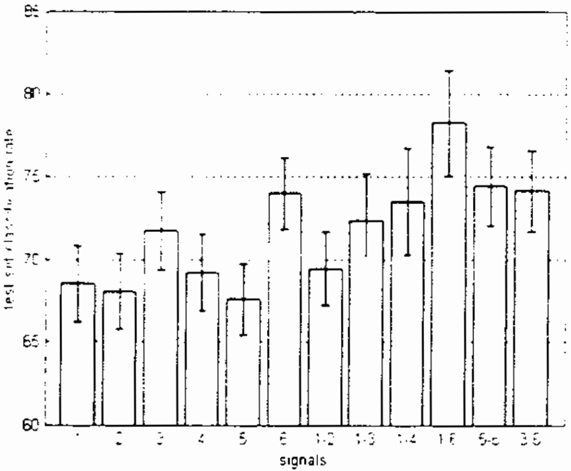


Figure 2: mean values and standard deviations of classification rates for different selected signals. Label "1-4" e.g. means inclusion of signal 1, ..., 4 (all EEG-signals) and label "3,6" means inclusion of signal 3 and 6.

Elimination of dead neurons after training, slightly improves classification rates, therefore it is recommended.

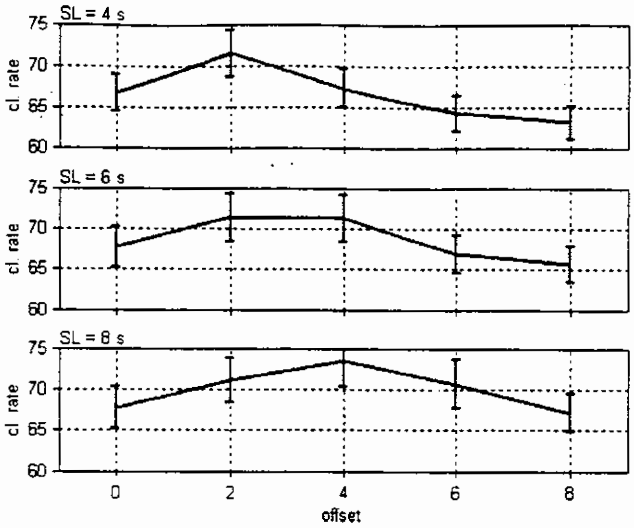


Figure 3: mean values and standard deviations of classification rates for different offsets (OF) and three different segment lengths (SL). Best results are obtained with SL = 8 sec and OF = 4 sec.

At first, we come to question 2, concerning the number of signals necessary for optimal classification. The LVQ network was trained and tested by different selections of signals (Fig. 2). Every signal was selected alone for training and testing. Signal 3, an EEG signal, and signal 6, an EOG signal, were much more suitably than the others. Both EOG signals (5-6) were more suitably than all EEG signals (1-4). But most success was gained by inclusion of all six signals (1-6).

Best classification results were obtained for averaging the spectral power densities in frequency bands. In general, a compromise has to be found between the number of features, because of curse of dimensionality, and the lack of refinement in spectral domain, i.e. extend of information loss. For band width of 1Hz results were slightly better than for 0.5 Hz and for 2 Hz. Surprisingly, decrease of classification rates were only about 3% when using the five frequency bands of EEG research mentioned above.

At second, we varied two preprocessing parameters (question 3), segment length (SL) and temporal offset (OF), mentioned above (Fig. 3). Larger segment lengths were better than shorter. The optimal offset was 4 sec. That means segments of all EEG- and EOG-signals beginning 4 sec before MSE- or NMSE-starting points and 4 sec after starting of an event were optimal.

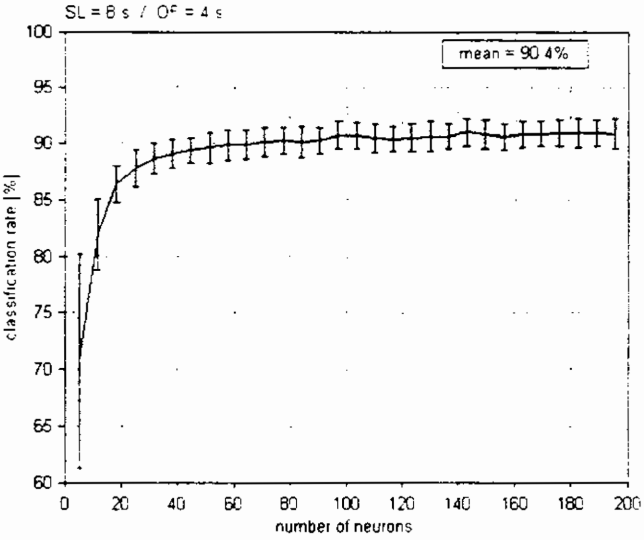


Figure 4: Classification rates versus number of neurons. 5,000 LVQ networks were trained and tested with different training set / test set partitions and different initializations. Mean classification rate (see inset) was computed between 30 and 200 neurons.

Further improvements in classification were reached by transformation log(x) of spectral power density as proposed in [14] and by choosing the number of neurons. Fig. 4 shows a typical plot of classification rate versus number of neurons. Classification success is increasing up to 90% in the mean if more than 50 neurons were used, and up to 91% with more than 140 neurons. With more neurons no significant changes are gained. Standard deviations are nearly constant.

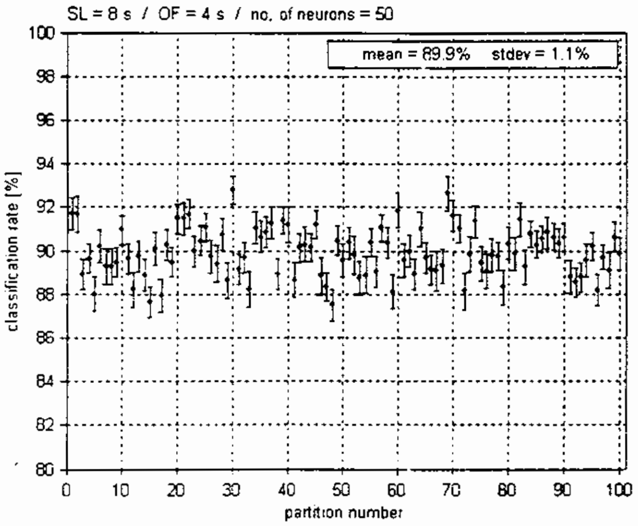


Figure 5: Classification rates of 100 fixed training set / test set partitions and different initializations. The grand mean and the standard deviation of mean classification rates is shown in the inset.

One might assume, that these results were obtained for a picked out test set. Therefore, we repeated partitioning many times. Classification results for 100 random partitions are shown in Fig. 5. For each partition, training and testing was repeated 500 times with different weight matrix initializations. The standard deviation caused by different initializations of the LVQ network and caused by training progress due to randomly applied input vectors is shown by error bars in Fig. 5. The deviations of the mean values are caused by different partitions and also caused by training progress and are slightly higher than the deviations caused by initialization.

DISCUSSION

Some authors pronounced that a combination of EEG and EOG measures should be most successful in predicting MSE [7,12,15]. Our results give further support for this statement. Classification rates of 90.4% in the mean were unexpectedly high and standard deviations of 1.4% were moderate. The high discrimination ability is also important in a general sense, because sudden behavioral transitions might be detectable by analyzing cortical potentials after some digital signal processing.

Estimation of spectral densities as a simple feature extraction method was applied because of a lack of prior knowledge for special patterns. Many authors reported an occurrence of alpha bursts, an increased EEG activity in the alpha frequency band, preceding MSE or during MSE [2,7]. But some subjects do not produce such an activity. Another characteristic pattern during wake / sleep transitions are slow eye movements [12,13], detectable in the EOG. Slow eye movements as well as alpha bursts don't occur during every MSE and their temporal relationship to the moment of starting MSE seems to be loose [2,11,16]. Therefore, we refrained from enforcing pattern specific analysis.

Further investigations are necessary to validate these results on additionally subjects. Many authors reported of large inter-individual differences of the EEG- and EOG- characteristic [2,11,15-17]. It is conceivable that the LVQ networks are representing individual characteristics of each of the eleven subjects under investigation. Furthermore, it should be investigated if the combination of EEG- / EOG- features and features of pupillography [4-6] might improve the discrimination of MSE.

Another topic of interest on the way to drowsiness warning systems must be investigation on discrimination of MSE from continuously measured and analyzed signals. Besides of robust artifact elimination techniques additional driver status estimation techniques, like sleepiness estimation, are strongly requested.

REFERENCES

[1] Thorpy, M.J., Yager, J.; The encyclopedia of sleep and sleep disorders; NewYork: Facts on File, 1991.

[2] Akerstedt, T., Torsvall, L. and Gillberg, M. (1987) Sleepiness in shiftwork, a review with emphasis on continuous monitoring of EEG and EOG. Chronobiology International 4, 129-140.

[3] McLaren, J., Erie, J.C. und Brubaker, R.F. (1992) Computerized analysis of pupillograms in studies of alertness. Investigative Ophthalmology & Visual Science 33, 671-676.

[4] Morad, Y., Lemberg, H., Yofe, N. and Dagan, Y. (2000) Pupillography as an objective indicator of fatigue. Current Eye Research 21, 535-542.

[5] Wilhelm, B., Giedke, H., Lüdtke, H., Bittner, E., Hofmann, A. and Wilhelm, H. (2001) Daytime variations in central nervous system activation measured by a pupillographic sleepiness test. J Sleep Research 10, 1-8.

[6] Golz, M., Sommer, D., Seyfarth, A., Trutschel, U. and Moore-Ede, M. (2001) Application of vector-based neural networks for the recognition of beginning microsleep episodes with an eyetracking system. In: Kuncheva, L. I. (ed.) Computational Intelligence: Methods & Applications (CIMA 2001), pp. 130-134. Alberta: ICSC Academic Press, Canada / The Netherlands.

[7] Torsvall, L. and Akerstedt, T. (1988) Extreme Sleepiness: Quantification of EOG and spectral EEG parameters. Int. J. Neuroscience 38, 435-441.

[8] Kohonen, T. (1988) Learning Vector Quantization. Neural Networks 1, 303-309

[9] Golz, M., Sommer, D., Lembcke, T. and Kurella, B. (1998) Classification of the pre-stimulus-EEG of k-complexes using competitive neural networks. In: Zimmermann, H. and Lieven, K., (eds.) Proc. of the 6th European Congress on Intelligent Techniques and Soft Computing (EUFIT98), Vol. 3, 1767-71. Aachen, Germany: Verlag Mainz.

[10] Peters, B.O., Pfurtscheller, G. and Flyvbjerg, H. (2001) Automatic differentiation of multichannel EEG signals. IEEE Trans Biomed Eng 48, 111-116.

[11] Torsvall, L. and Akerstedt, T. (1987) Sleepiness on the job: continously measured EEG changes in train drivers. Electroencephalography and clinical Neurophysiology 66, 502-511.

[12] Santamaria, J. und Chiappa, K.H. (1987) The EEG of Drowsiness in Normal Adults. J Clin Neurophysiol 4, 327-382.

[13] Sommer, D. and Golz, M. (2001) Clustering of EEG-segments using hierarchical agglomerative methods and self-organizing maps. In: Dorffner, G., Bischof, H. und Hornik, K., (eds.) Artificial Neural Networks - ICANN 2001, pp. 642-649. Berlin: Springer.

[14] Gasser, T., Bächer, P. and Möcks, J. (1982) Transformations Towards the Normal Distribution of Broad Band Spectral Parameters of the EEG. Electroenceph. clin. Neurophysiol. 53, 119-124.

[15] Hyoki, K., Shigeta, M., Tsuno, N., Kawamuro, Y. and Kinoshita, T. (1998) Quantitative electro-oculography and electroencephalography as indices of alertness. Electroenc. Clin. Neurophysiol. 106, 213-219.

[16] Ogilvie, R.D., Mc Donagh, D.M., Stone, S.N. and Wilkinson, R.T. (1988) Eye movements and the detection of sleep onset. Psychophysiology 25, 81-91.

[17] Jung, T.-P., Stensmo, M., Sejnowski, T. and Makeig, S. (1997) Estimating Alertness from the EEG Power Spectrum. IEEE Transactions on Biomedical Engineering 44, 60-69.

Gene Expression Data Based Cancer Status Prediction for Individual Patients by an Optimized Fuzzy Cluster and Rule Based System

Michael Pfaff, Reinhard Guthke*, Dirk Woetzel, Dominik Driesch
BioControl Jena GmbH
Wildenbruchstr. 15, D-07745 Jena, Germany
Phone: +49-3641-675511, Fax: +49-3641-675512
email: biocontrol@t-online.de
*Hans Knoell Institute for Natural Product Research
Beutenbergstr. 11, D-07745 Jena, Germany
Phone: +49-3641-656820, Fax: +49-3641-656825
email: rguthke@pmail.hki-jena.de

ABSTRACT: The development of microarray technology to gather gene expression data has made significant progress over the last years. In medicine, it is a major aim to use this technology together with adequate data analysis methods to recognize different health/disease states of patients. Various data analysis methods, ranging from statistics to data mining, have been applied to discover the most informative genes and to predict unknown health/disease states of patients. This paper describes a novel fuzzy cluster and fuzzy rule based approach to cancer status prediction based on gene expression monitoring by DNA microarrays and a data set from 72 leukemia patients taken from the literature. The data set originated from 25 patients with acute lymphoblastic leukemia (ALL) and 47 patients with acute myeloid leukemia (AML). The supervised learning approach applied is entirely based on fuzzy methods and includes: data based clustering, cluster based rule extraction with subsequent rulebase construction and finally cluster and rule based prediction of the cancer status of patients. Clustering was carried out using the fuzzy c-means algorithm and rule extraction using the normalized empirical fuzzy rule rating measure. To optimize the performance of the prediction system two quantitative criteria were introduced and applied: First, the "dummy gene" to identify rules (genes) that are informative at all with respect to the intended prediction and second, the "informative imbalance" to optimize the composition of the rules in the rulebase - solely based on the training data - with respect to a minimum prediction error for the test data, thereby avoiding overfitting to the training data and using as much as possible global information in the training data to better fit the system to unknown data. The optimum rulebase designed contains just 13 genes (5 predicting ALL and 8 predicting AML) out of the 7,129 original ones covered by the microarray gene expression data set and predicts all 38 training patient samples and, more importantly, 33 of the 34 independent test patient samples correctly. The results obtained underline the feasibility of this novel approach since it provided at least the same or even better results compared to those in the literature. The data mining approach applied allows for the development of application-specific minimized assays or arrays for the detection of specific health/disease states of patients based on a prior overall analysis of microarray gene expression data that covers a multitude of genes.

KEYWORDS: Gene Expression Monitoring, Leukemia, Cancer Status Prediction, Informative Genes, Fuzzy Clustering, Fuzzy Rule Extraction, Rulebase Optimization

GENERAL

The paper presents a novel fuzzy data mining approach towards the prediction of health/disease states of patients based on gene expression data obtained by microarray technology.

The study was based on data originally published in [1] and taken from the respective website [2]. The data set is described in detail there. This data was obtained by DNA microarray gene expression monitoring from 72 leukemia patients of whom 47 had acute myeloid leukemia (AML) and 25 had acute lymphoblastic leukemia (ALL).

Methods used and results obtained by the approach applied here were compared to those described in the original paper [1] and a subsequent paper [3] which also used the same data for supervised learning. In both these papers, primarily statistical methods were used to predict known disease states from unknown samples.

Intelligent e-Health Applications in Medicine
Special Session Proc. EUNITE 2002, Albufeira, Portugal, Sept. 19-21, 2002
Published by the University of the Aegean

In the present paper the focus is on fuzzy data mining methods and more precisely on the use of fuzzy clusters together with fuzzy IF-THEN-rules to predict the leukemia cancer status of patients. The use of rules supports the interpretation of the underlying problem.

Also, in the present paper the particular focus is on avoiding overfitting of the prediction system to the training data set. In the centre of attention was therefore an algorithm to generate a prediction system, in this case a clusterbase/rulebase combination, that is more globally informative, i.e. that reflects not only information characteristic of the training data set but also of potential unknown data sets (e.g. the test data set). Therefore, this approach aims to extract as much as possible global information and as little as possible local information from the training data set.

Since the composition of rules in the rulebase is crucial to the robustness of the predictive capability upon change from known (i.e. training) to unknown (i.e. test and new) data, a special algorithm was developed in order to construct an adequate rulebase out of ranked rules.

In the literature cited above [1] [3], which served as a benchmark in this study, the term molecular classification is used. Since, however, the approach applied here is not only based on clustering/classification but in particular on rule extraction and rulebase construction, the term prediction is used throughout to describe the entire approach. More specifically, the term refers to fuzzy cluster and fuzzy rule based prediction as applied here.

Although the work carried out is based on the analysis of microarray spot expression signals as features that represent gene fragments or whole genes and one gene may be represented by more than one spot, in compliance with the literature these features are referred to as genes throughout the paper.

APPROACH AND METHODS

The fuzzy cluster and fuzzy rule based approach to cancer status prediction applied here consists of four major steps. These are: data based clustering to establish a clusterbase, cluster based rule extraction to establish rule sets for the subsequent rulebase construction and finally cluster and rule based prediction. The first step is preceded by data pre-processing.

This approach is illustrated in Figure 1 and was described in greater detail in [4] in relation to a different application in medicine. This figure also shows the specific adaptations made with respect to the current application, namely the introduction of the "dummy gene" and the "informative imbalance" which will be explained later. Also, in the current application, the approach is entirely based on fuzzy methods.

Preceding the first step of the approach, the data was split into a training data set (consisting of 38 patient samples) and a test data set (consisting of 34 patient samples) and pre-processed, i.e. normalized, all according to [1].

In the first step of the approach (see Figure 1), the patient samples for each individual gene were fuzzy clustered into two clusters in order to establish a clusterbase for the problem investigated. This was carried out using the fuzzy c-means algorithm [5].

In the second step of the approach (see Figure 1), uniconditional rules, i.e. rules that have just one IF-part (condition) apart from one THEN-part (conclusion), based on the fuzzy clusters obtained were generated, rated by the normalized empirical rule rating measure [6] and ranked according to this rating measure in order to obtain rule sets with rankings of the individual rules.

The structure of these cluster based rules reads as follows (example):

IF Expression of Gene X is in Cluster x1 THEN Leukemia Status Cluster is AML

The IF-parts (conditions) of the rules contain the expression clusters of the genes, while the THEN-parts (conclusions) of the rules contain the cancer status clusters (either ALL or AML). There are two rule sets, one containing all rules with the cluster ALL in their conclusion and one containing all rules with the cluster AML in their conclusion.

In order to identify genes in the rule sets that are informative at all with respect to the intended application, an additional "dummy gene" was introduced that represents the non-informative "gene". This "dummy gene" is characterized by having the same expression level, i.e. the same cluster assignment, for all patients. Using this "gene" in the IF-parts of rules concludes that all patients belong to the same cancer status cluster. In this sense, it is not discriminating between the cases and equals a guess without information. It can therefore be used to mark the separation between informative and non-informative genes in the ranked rule sets. All rules ranked higher than the ones containing the "dummy gene" are candidates for transfer to the rulebase for prediction.

In the third step of the approach that is merging with its fourth one (see Figure 1), rulebase construction for prediction, top ranked rules well above the "dummy gene" rules from both rule sets (ALL conclusion cluster rule set and AML conclusion cluster rule set) were transferred to the rulebase.

Rulebase construction for prediction was based on the following steps and criteria (see Figure 2):

First, the rulebase should not contain more rules than the number of training samples to avoid overfitting to the training data, i.e. in the case of uniconditional rules investigated here, the number of rules (genes) contained in the rulebase should not exceed the number of training patient samples (here 38).

Second, the number of rules in the rulebase (so far equally balanced between the ALL and AML conclusion cluster rules) should be minimized as much as possible starting from the upper limit (here 38) until the prediction error for the training data set worsens substantially (coarse tuning).

Third, since, for instance, the rules (and the gene expressions behind them) out of the ALL rule set that contribute to the rulebase can be more discriminating with respect to the intended prediction than the ones out of the AML rule set, an informative imbalance can appear in a constructed rulebase so that a quantitative criterion has to be formulated to balance the information content that is contributed by either of the two rule sets (fine tuning).

While coarse tuning reduces the number of rules that contribute to the rulebase, fine tuning is likely to increase this number again, yet below the upper limit, until an optimum is reached.

For fine tuning, the following criterion ("informative imbalance" – *IIB*) was developed:

$$v_1 = \sum_i h_i, \quad v_2 = \sum_j h_j$$

$$IIB = \left| \left(\frac{\overline{|v_1 - v_2|}}{\overline{|v_1|}} \right)_{ALL} - \left(\frac{\overline{|v_1 - v_2|}}{\overline{|v_2|}} \right)_{AML} \right| \longrightarrow Min \qquad (1)$$

IIB: informative imbalance of the rulebase
h: hit rate for one rule and an individual patient
ALL: over all patients in the training data set who have ALL
AML: over all patients in the training data set who have AML
i: index for rules that "fire" for conclusion ALL
j: index for rules that "fire" for conclusion AML

In order to optimize the composition of the rules in the rulebase by fine tuning, different partial rule sets of the two rules sets (with either ALL or AML in the conclusion cluster) were generated and evaluated by the above criterion (1), e.g. rulebase A: 7 top ranked rules of the ALL rule set and 5 top ranked rules of the AML rule set, rulebase B: 5 top ALL rules and 5 top AML rules, rulebase C: 5 top ALL rules and 8 top AML rules, etc. This optimization process is shown in a schematic representation in Figure 3.

The optimized rulebase with the minimal *IIB* value according to (1) together with the underlying clusterbase was then used for prediction (see Figure 1).

Figure 4 shows an example of the prediction voting of the individual uniconditional rules contained in the rulebase when gene expression data from a patient with unknown status is used to predict his status by the designed system. The voting of the individual rules takes place using the fuzzy classification results of the new sample, i.e. its values of fuzzy assignment to the previously established clusters for the individual genes, in the conditions of the individual rules. The respective conclusion of each individual rule then gets a vote that equals the fuzzy assignment value of its condition. All votes of the rules in the ALL conclusion cluster rule set and all in the AML conclusion cluster rule set are then being totalled separately and compared. The conclusion cluster with the higher total vote finally predicts the patient's status.

The prediction system design described above was followed by the validation of the clusterbase's and rulebase's predictive capability using the independent *test* data set (see Figures 1 and 2).

Figure 5 summarizes this approach. The patient samples (gene expression data) in the test data set were fuzzy classified using the clusterbase (derived from the training data set only) for gene expression recognition of the rule conditions before subjecting them to the rulebase (derived also from the training data set only) for the test patients' cancer status prediction by the rule conclusions. In the same way as the test patient samples, new (unknown) patient samples (gene expression data) are processed during the operation stage of the prediction system as designed in the training stage and validated in the test stage to predict the cancer status of patients hitherto unknown to the system.

To prove that the "informative imbalance" criterion is capable of optimizing the composition of the rules in the rulebase with respect to the correct prediction of the *test* patients' cancer status, although the design is solely based on the *training* data, several rulebases (with the established underlying clusterbase) were used to investigate the prediction error for the *test* patients' cancer status versus the *IIB* value, thereby validating the entire approach (as shown in Figures 1, 2, 3, 4, 5).

RESULTS

The original gene expression data set [2] consisting of 513,288 single data and describing 6,817 genes represented by 7,129 spot expression signals for each of the 72 leukemia patient samples was split into a data set of 38 patient samples used for training (270,902 single data) and another one of 34 patient samples used for testing (242,386 single data). Of the 38 training samples 27 originated from patients with ALL and 11 from patients with AML, while of the 34 test samples 20 originated from patients with ALL and 14 from patients with AML. The original data set was split exactly the same way as in [1] and [3].

The data was then normalized and all spot expression signals with zero standard deviation of all the 38 training patient samples were omitted because they can not be clustered into different clusters and not be used for discrimination. All the data pre-processing steps were carried out following the procedure in [1].

The training patient samples for all 6,646 remaining spot expression signals (in the following referred to as genes) were now fuzzy clustered into two clusters using the fuzzy c-means algorithm [5]. The training patient samples for the "dummy gene" were *a priori* set, i.e. crisp clustered, so that all of them belong to just one cluster.

Based on these clusters, altogether 26,588 uniconditional rules, 13,294 for each of the two cancer status clusters (conclusions), ALL and AML, respectively, were then generated, rated by the normalized empirical rule rating measure [6] and ranked accordingly.

Out of the rules of each of the two rule sets obtained those rules were further processed that were ranked higher than the "dummy gene", since they represent the genes that are informative at all as described above. 419 rules for the ALL conclusion cluster rule set and 3,707 rules for the AML conclusion cluster rule set were obtained above the highest of the two rules in each of the two rule sets that stand for the "dummy gene".

Out of these informative genes obtained from the training patient samples, the rulebase was constructed by the coarse tuning described above (see Figure 2) beginning from the 19 top ranked rules and resulting in the 5 top ranked rules for each of the two rule sets. This rulebase is already capable of predicting all 38 *training* patient samples correctly.

The final aim, however, is to predict the unknown, i.e. the 34 *test* patient samples correctly. The rulebase designed solely based on the *training* patient samples therefore has to be validated using the independent *test* patient samples. This was done in the next two steps (see also Figure 2).

After the coarse tuning, the fine tuning as explained above was applied using the *IIB* criterion (1) and the training patient samples to subsequently change the number and the composition of the rules, i.e. the partial ALL and AML conclusion cluster rule sets in the rulebase, until the criterion reached its minimum. This is shown in Figure 6 where the *IIB* criterion (1) is drawn versus the number of top ranked ALL rules and the number of top ranked AML rules contained in the rulebase. It can be clearly seen in this figure for the rulebase compositions investigated that in order to reach minimum values for the informative imbalance of the rulebase, the number of ALL rules has to be smaller (5 or 6) than the number of AML rules (7 to 10). The opposite rulebase composition leads to much higher *IIB* values. Also, in terms of sensitivity, to reach low values of the *IIB* criterion, it is more important that the number of ALL rules is smaller (4 to 6) whereas the number of AML rules can cover a wider range (4 to 10). Vice versa, when the number of AML rules is smaller (4 to 5) and the number of ALL rules covers a wider range (4 to 15) than this will yield only much higher *IIB* values. These results clearly point towards a rulebase where the informative imbalance has to be counter-balanced by a higher number of AML rules compared to the number of ALL rules.

The minimum *IIB* value of 0.0222 was obtained for the optimum rulebase composition consisting of 5 ALL conclusion cluster rules and 8 AML conclusion cluster rules. This resulted in a correct prediction for 33 of the 34 *test* patient samples.

These results are shown in Figure 7 (prediction error for *test* patient samples versus the *IIB* value for the *training* patient samples and different rulebase compositions). This figure also summarizes the trace of the rulebase optimization process using the *training* data set only and its validation using the independent *test* data set. It demonstrates that the informative imbalance is compensated by a counterbalance of the ratio of the ALL conclusion cluster rules and the AML conclusion cluster rules in the different compositions of rulebases investigated. It shows that for *IIB* values above about 0.5 there is an ALL rule majority in the rulebase, while for *IIB* values between about 0.25 and 0.3 there is a rough balance between ALL and AML rules. For *IIB* values below about 0.25, however, there is a majority of AML rules, obviously counterbalancing the informative imbalance of the rules contained in the rulebase. This imbalance originates from a different degree of separation of the clusters of those genes that are contained in the rulebase. In the range below about 0.25 there is obviously a good convergence of the *IIB* value towards its minimum leading to the optimum composition of the rulebase.

Figure 7 also proves that the criterion (1) devised represents the aim of the rulebase optimization based on the preceding steps very well, i.e. to minimize the error of prediction for the *test* patient samples although only *training* patient samples are used during the design of the prediction system (compare Figures 1 to 6).

The approach applied can therefore be considered validated for the problem (disease) investigated here. Although it may be assumed that for similar problems, the approach works equally well, this has to be shown in future studies for other diseases based on different gene expression data.

Results obtained by this fuzzy data mining approach using the optimum rulebase together with the underlying clusterbase for prediction are shown in Figure 8. This figure displays all 4,126 informative genes (spot expression signals) out of the 7,129 original ones (those above the "dummy gene" and not excluded before during pre-processing) in the ranked rule sets, i.e. the 419 genes in the uniconditional rules for the ALL conclusion cluster rule set (small circles) and the 3,707 genes for the AML conclusion cluster rule set (small triangles). They are displayed as the logarithmic mean spot signal values for AML patient samples versus the logarithmic mean spot signal values for ALL patient samples over all training patients samples. The figure also shows all the 13 genes that are contained in the optimum rulebase, i.e. the 5 genes in the ALL conclusion cluster rules (large circles) and the 8 genes in the AML conclusion cluster rules (large triangles). These 5 and 8 genes have been numbered separately in Figure 8 according to their ranking in the rulebase and refer to the ranking numbers in Table 1 where the respective gene description with the gene accession number is given. This table also shows whether these genes are up-regulated or down-regulated.

Genes shown in Figure 8 that lie on the 45° line are not differentially expressed in the ALL and the AML patient samples. Genes that lie above this line are down-regulated in the ALL patient samples compared to the AML patient samples and genes that lie below this line are up-regulated in the ALL patient samples compared to the AML patient samples.

Figure 8 shows that all 419 informative ALL genes identified (small circles) are down-regulated, while the 3,707 informative AML genes identified (small triangles) are either up-regulated or down-regulated.

This figure also shows that the applied approach allows to identify genes that predict one cancer status (here AML) although they are not extremely differentially expressed compared to the other cancer status (here ALL), i.e. to identify discriminating genes close to the 45° line. This can be seen in Figure 8 for the 5 up-regulated genes out of the 8 genes in the rulebase predicting the AML cancer status (large triangles, No. 1, 3, 4, 5, 6). Moreover, the approach can identify discriminating genes predicting one cancer status (here AML) and having similar differential expression levels then the ones found to predict the opposite cancer status (here ALL). This can be seen in Figure 8 just above the 45° line where discriminating genes predicting opposite cancer states are very close to one another (e.g. large circle No. 5 close to large triangle No. 6).

The 13 genes listed in Table 1 as contained in the optimum prediction rulebase and identified as the most relevant ones with respect to the described cancer status prediction are currently being investigated in an experimental study with independent leukemia patient samples in order to validate the results of the approach by experimental means too.

The approach designed can eventually lead to the development of problem-specific minimized assays or microarrays with a limited number of the most relevant genes and their application in medicine and biomedical research in order to detect specific health/disease states after a prior overall analysis of problem related gene expression data that covers a multitude of genes.

DISCUSSION

It was a major aim of this study to provide improved methods for the gene expression data based prediction of health/ disease states, as in this case cancer states, and to benchmark these results with those of other methods.

The prediction results obtained by the optimum rulebase with the underlying clusterbase were therefore compared to results presented in the literature, in particular in [1] and [3], where exactly the same data and the same pre-processing steps were used.

A summary of this comparison is given in Table 2. Apart from the overall results, this table lists the methods applied. Also, since the approaches finally aim to make predictions for individual patients, the results of the three approaches for all 72 individual patient predictions are shown in Table 3.1 for the 38 training samples and in Table 3.2 for the 34 test samples, both together with the actual cancer status of the individual patients.

The predictor described in [1] is based on 50 genes. It predicts only 36 of the 38 training patient samples and just 29 of the 34 test patient samples correctly. This is probably due to an overfitting of the predictor to the training data, which to avoid was the declared aim of the approach presented in this paper.

As described in [3], a 25 gene model based on statistics predicted 32 out of the 34 test patients samples correctly, while a 5 gene model even predicted 33 out of the 34 correctly. Both models also predicted all 38 training patient samples correctly This coincides with the results obtained here, where the same number was predicted correctly by the optimum rulebase containing 13 genes. Most interestingly, the only incorrectly predicted patient sample was the same as in [3] (No. 66). This sample was also incorrectly predicted by the approach described in [1]. This might be viewed as an indication for a wrongly marked sample or the like, since this result was obtained by three independent methods.

Also, only 2 genes overlapped between the 5 gene model in [3] and the 13 gene rulebase reported here. These genes have the accession numbers X95735_at and M84526_at (for gene description see Table 1). This rather small overlap despite identical prediction results may be due to the redundancy of the genes that do not overlap. On the other hand, this also seems to indicate that for the particular problem investigated a rather very limited number of relevant genes is important. This is supported by the fact that these two genes are the only ones that were obtained by all four models of the three approaches. The other overlapping genes between the respective methods are listed in Table 2.

The approach in [3] and the one presented here clearly show a better prediction performance than the one in [1]. While the approach in [3] demonstrates that classical statistical methods can equally well be used for prediction, the data mining approach presented here has further potential when extended from uniconditional to multiconditional rules to be even more adequate to the complexity of gene expression. Moreover, the use of rules supports the interpretability of results and the causal understanding of the underlying processes.

Another advantage of the approach applied here is that top ranked rules with a particular condition that predicts a particular disease state as conclusion do not necessarily imply that the opposite condition predicts the opposite conclusion. As can be seen in Table 1, for example, although "X95735_at *down*-regulated" is predicting *ALL*, "X95735_at *up*-regulated" is not predicting *AML*. On the contrary, "X95735_at *up*-regulated" may point towards a number of other health/disease states and not necessarily towards AML, thereby allowing a much more adequate description of the real conditions behind the problem.

SUMMARY AND CONCLUSIONS

The paper presents a fuzzy logic, cluster and rule based data mining approach to the prediction of the two leukemia cancer states ALL and AML based on microarray gene expression data.

With the methods applied a prediction rulebase consisting of uniconditional rules was constructed containing 5 genes that predict the ALL cancer status and 8 genes that predict the AML cancer status. This rulebase proved to be optimal with respect to the informative imbalance criterion introduced in this work. The prediction using this rulebase yielded 33 correct predictions for 34 patients, i.e. there was only 1 incorrect prediction for the test patient data, i.e. for the data that was not used for training of the rulebase.

The study aimed at benchmarking the present approach with two other approaches in the literature using exactly the same gene expression data and pre-processing steps. While in [3] the approach yielded similar results as in the present one, the approach in [1] showed a weaker performance. The main reason for this may be seen in the lack of an appropriate algorithm to assemble the genes that are finally used within the predictor. In the present approach the "informative imbalance" criterion was introduced and successfully applied to optimize the assembly of genes in the rulebase for prediction.

Although the approach applied here is multivariate already using a voting of an entire rulebase based on various genes, even more complex relations between genes may be represented by multiconditional rules in the rulebase.

Preliminary rulebases with biconditional rules showed good predictions for the test patients (32 out of 34 were predicted correctly). However, the existing "informative imbalance" criterion for rulebases that consist of uniconditional rules needs to be further modified with respect to multiconditional rules.

This work showed that fuzzy cluster and rule based systems are capable of reliable health/disease status prediction when adequate optimization criteria are provided. It has to be emphasized that not only methods to find informative genes are important but also methods to assemble reliable prediction systems.

ACKNOWLEDGEMENT

This work was supported by the German Federal Ministry for Education and Research (BMBF).

REFERENCES

[1] T.R. Golub, D.K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J.P. Mesirov, H. Coller, M.L. Loh, J.R. Downing, M.A. Caligiuri, C.D. Bloomfield and E.S. Lander, 1999, "Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring", *Science*, vol. 286, pp. 531-537.

[2] http://www.genome.wi.mit.edu/MPR

[3] J. Lu, S. Hardy, W. Tao, S. Muse, B. Weir and S. Spruill, 2002, "Classical Statistical Approaches to Molecular Classification of Cancer from Gene Expression Profiling", in: *Methods of Microarray Data Analysis* , Eds.: S.M. Lin, and K.F. Johnson (Kluwer Academic Publishers), pp. 97-108.

[4] Pfaff, M.; Woetzel, D.; Weller, K.; Schroeder, K.; Pohlmeier, R.; Vienken, J.; Stein, G.; Guthke, R.; 2001, "Cluster and Rule Based Prediction in Hemodialysis – Adaptivity vs. Robustness. EUNITE 2001 European Symposium on Intelligent Technologies, Hybrid Systems and their Implementation in Smart Adaptive Systems, December 13-14, 2001, Puerto de la Cruz, Tenerife, Spain. Final Programme & Proceedings (Abstracts + CD), Published by Verlag Mainz, Wissenschaftsverlag, Aachen, Germany,. Session on Computational Intelligence in Medicine. Abstract: Booklet pp. 65/66; Full Paper: CD-ROM pp. 286-335; Shorter Paper Version: In: Adaptive Systems and Hybrid Computational Intelligence in Medicine (Eds.: G.D. Dounias and D.A. Linkens), Special Session Proceedings of the EUNITE 2001 Symposium, Organized by IBA D: Human, Medical & Health Care, December 13-14, 2001, Tenerife, Spain, Published by the University of the Aegean, Greece, in Collaboration with EUNITE, The European Network of Excellence on Intelligent Technologies for Smart Adaptive Systems, pp. 14-31.

[5] J.C. Bezdek, 1981, "Pattern Recognition with Fuzzy Objective Function Algorithms", *Plenum Press*, New York, USA.

[6] R. Guthke and B. Ludwig, 1994, "Generation of Rules for Expert Systems by Statistical Methods of Fermentation Data Analysis", *Acta Biotechnol.*, 14, pp. 13-26.

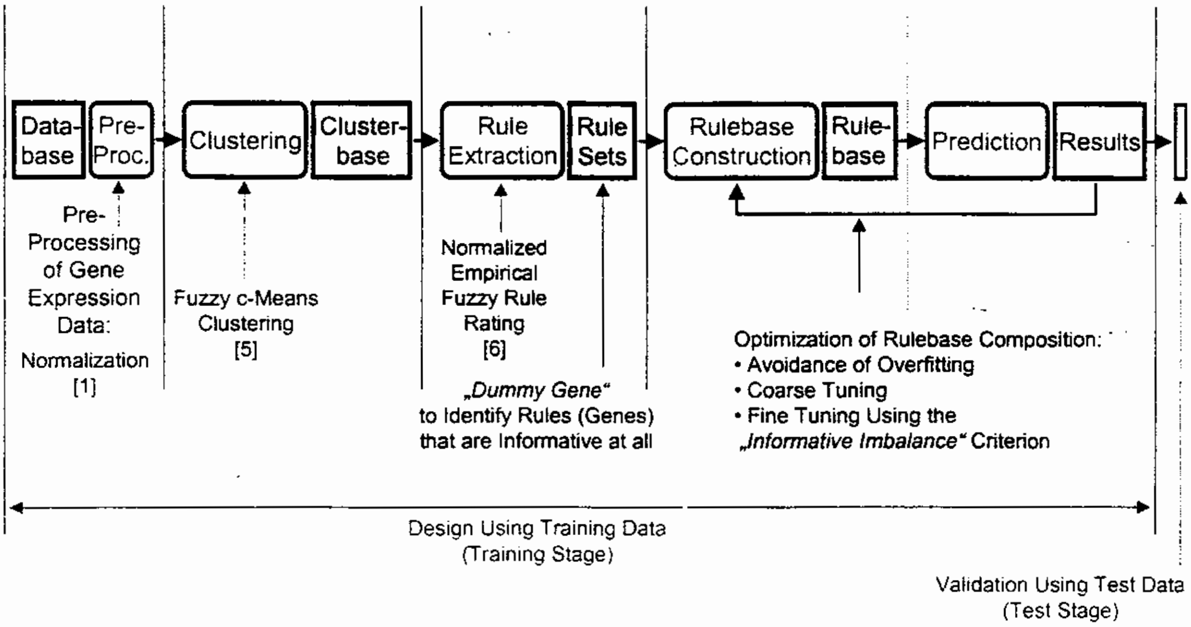


Figure 1: Overview of Approach and Methods

1st AVOIDANCE OF OVERFITTING

Number of rules in the rulebase < Number of patient samples in the training data set
(Rules from top ranked "informative" rules of the rule sets
well above the "dummy gene" obtained from the training data set)

2nd COARSE TUNING

Minimization of the number of rules in the rulebase
(until the prediction error for the training data set worsens substantially)

3rd FINE TUNING

Minimization of the "informative imbalance" criterion based on the training data set
by changing the composition of the rules in the rulebase
(Number of rules might increase to some extent below the above limit)

VALIDATION

Proof that the "informative imbalance" criterion
actually minimizes the prediction error also for the test data set,
although it is only calculated based on the training data set

Figure 2: Steps and Criteria for Rulebase Construction and Optimization

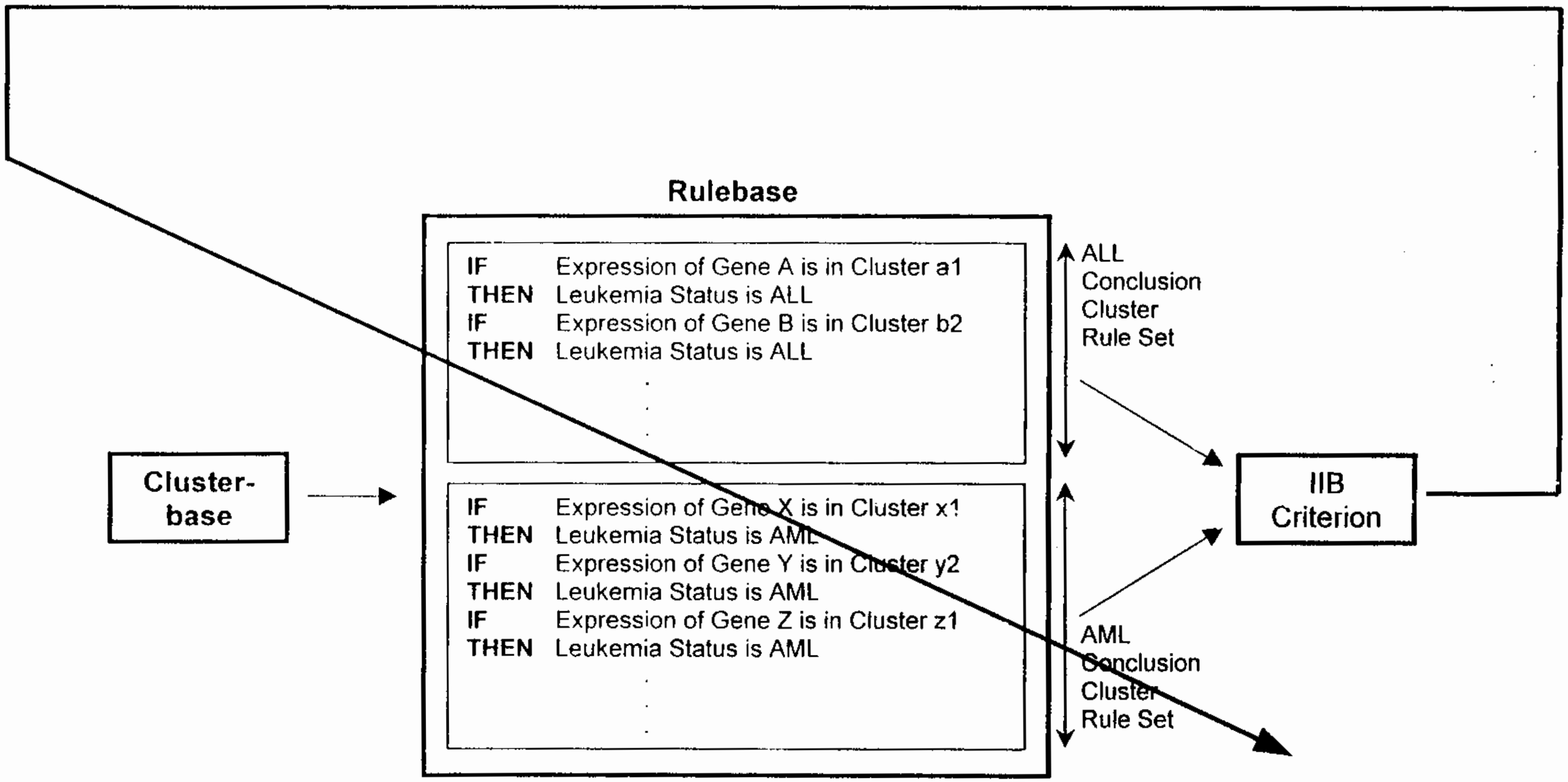


Figure 3: Rulebase Optimization Process

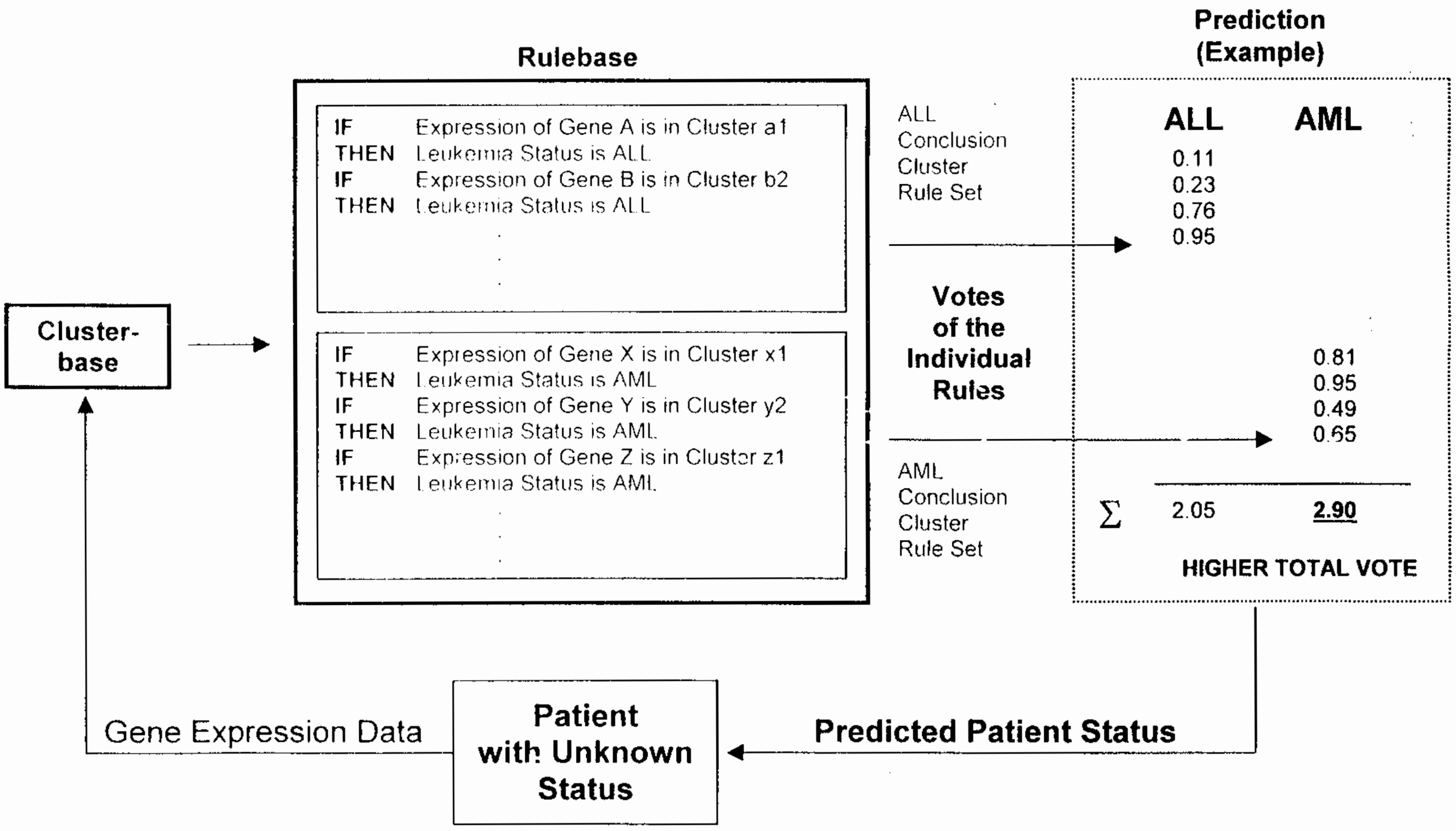


Figure 4: Prediction Voting of the Rulebase

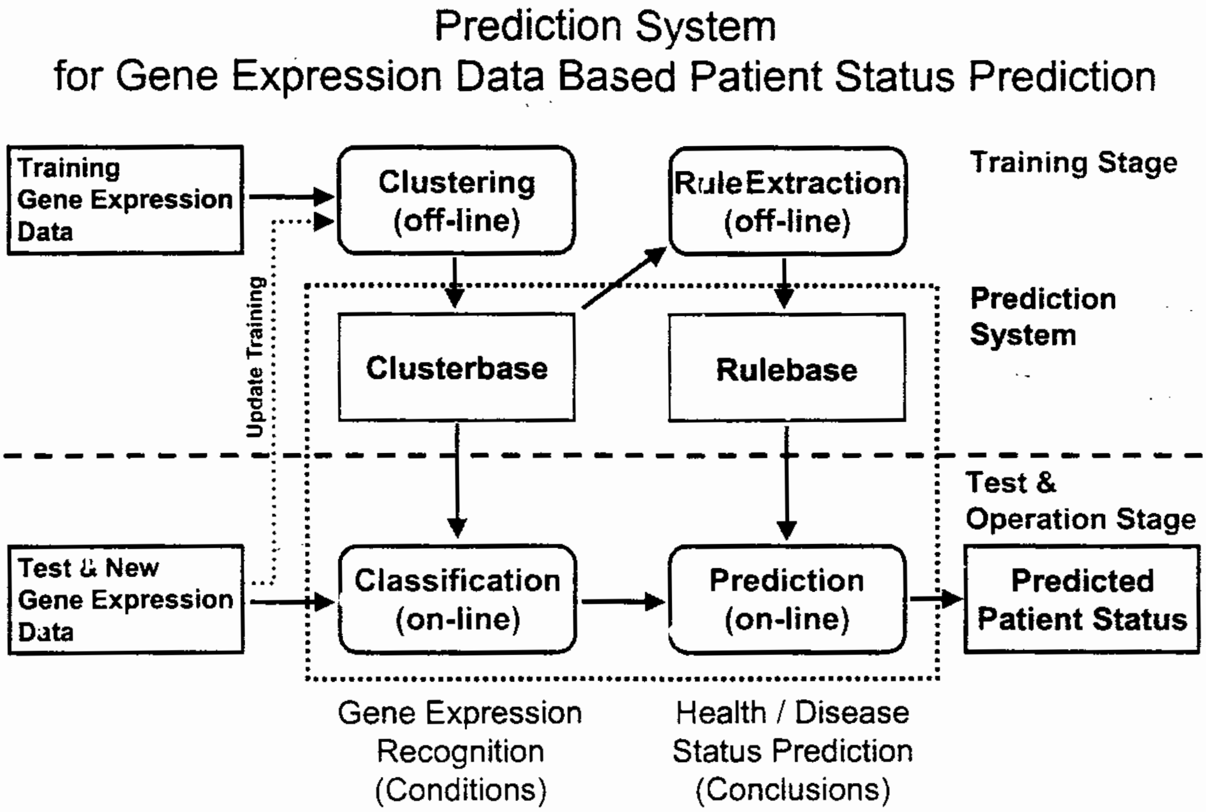


Figure 5: Prediction System for Cluster and Rule Based Patient Status Prediction

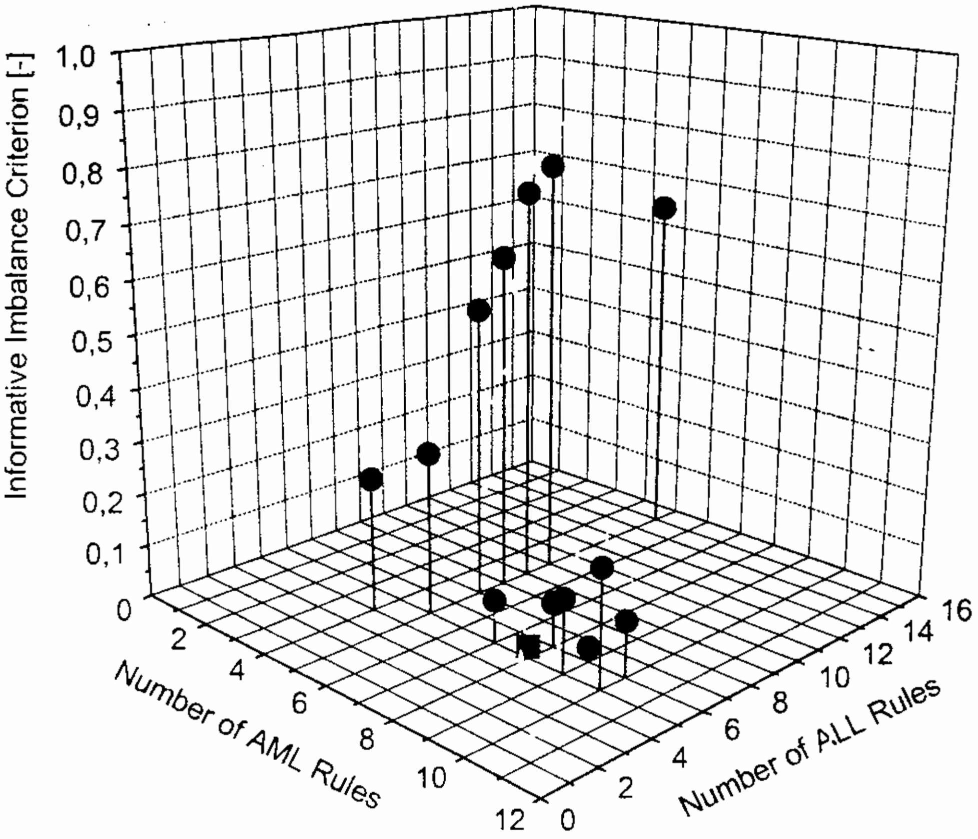


Figure 6: Informative Imbalance Criterion (I) versus the number of top ranked ALL rules and the number of top ranked AML rules contained in the rulebase

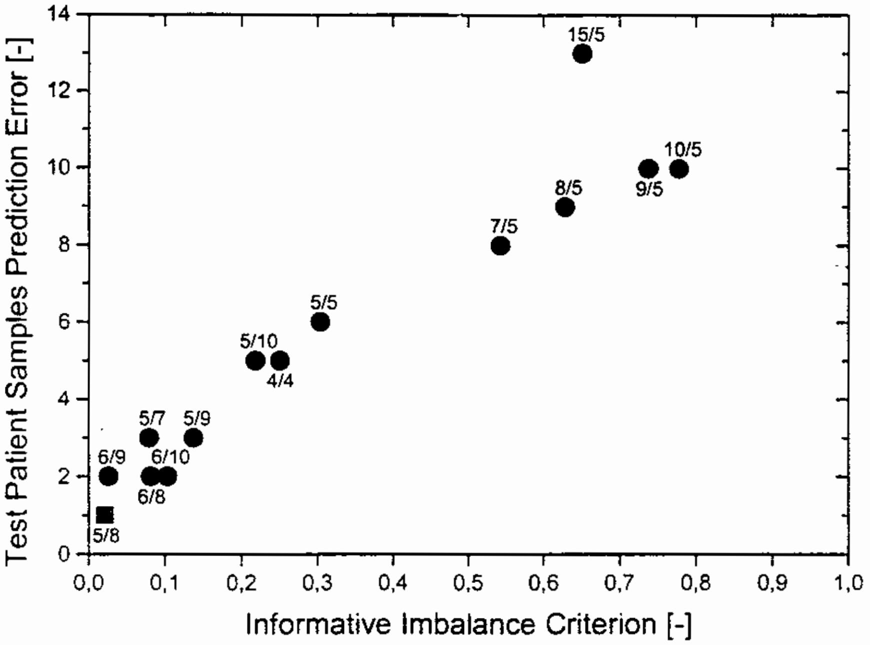


Figure 7: Prediction Error for the Test Patient Samples versus the Informative Imbalance Criterion for Different Rulebase Compositions (Numbers next to the symbols denote the number of ALL conclusion cluster rules (left) and AML conclusion cluster rules (right) contained in the respective rulebase.)

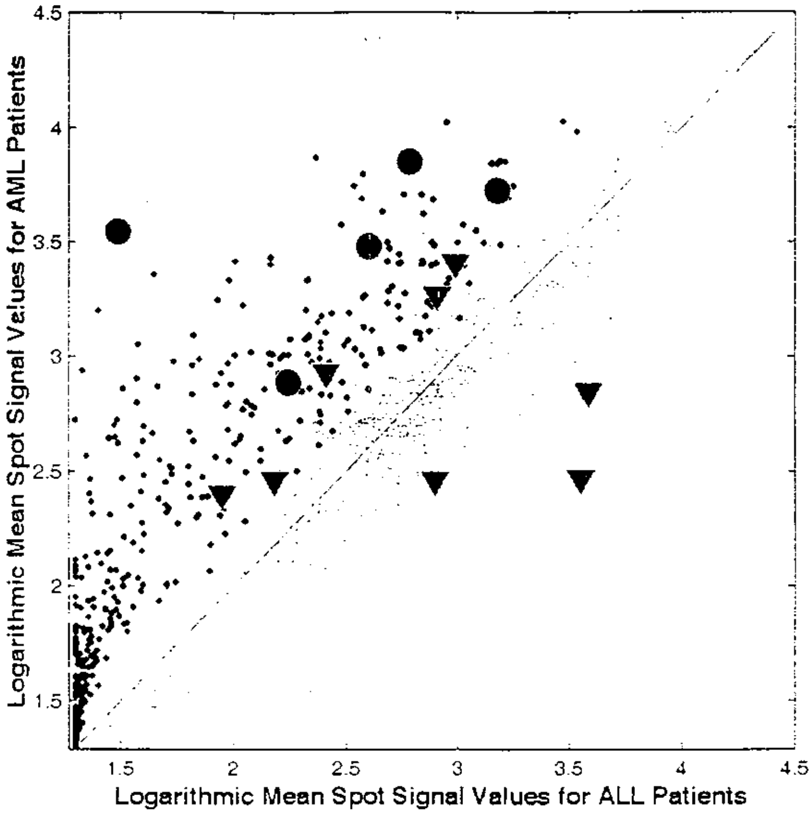


Figure 8: Logarithmic mean spot signal values for AML training patient samples versus logarithmic mean spot signal values for ALL training patient samples. Shown are all 4126 informative genes (small symbols) and the 13 genes contained in the optimum prediction rulebase (large symbols). Numbers refer to those in Table 1.

Ranking	Gene Description	Gene Accession Number	down/up regulated
Predicting ALL			
1	Zyxin	X95735_at	down
2	Phosphotyrosine independent ligand p62 for the Lck SH2 domain mRNA	U46751_at	down
3	DF D component of complement (adipsin)	M84526_at	down
4	Azurocidin gene	M96326_rna1_at	down
5	CD33 antigen (differentiation antigen)	M23197_at	down
Predicting AML			
1	FAH Fumarylacetoacetate	M55150_at	up
2	C-myb gene extracted from Human (c-myb) gene, complete primary cds, and five complete alternatively spliced cds	U22376_cds2_s_at	down
3	Thrombospondin-p50 gene extracted from Human thrombospondin-1 gene, partial cds	U12471_cds1_at	up
4	CDC25A Cell division cycle 25A	M81933_at	up
5	Leukotriene C4 synthase (LTC4S) gene	U50136_rna1_at	up
6	GB DEF = Homeodomain protein HoxA9 mRNA	U82759_at	up
7	MB-1 gene	U05259_rna1_at	down
8	T-COMPLEX PROTEIN 1, EPSILON SUBUNIT	D43950_at	down

Table 1: Ranking, Gene Description, Gene Accession Number and Down-/Up-Regulation of the 13 Genes Contained in the Optimum Prediction Rulebase

	Golub et al. 1999 [1]	*Lu* et al. 2002 [3]	Present Publication 2002
Methods (general)	Statistics	Statistics	Data Mining
Methods (specific)	For Class Prediction: "Neighbourhood Analysis" using a Class Distinction ("idealized expression pattern", i.e. uniformly exppressed gene) and correlation analysis using "weighted vote" and "Prediction Strength" Validation: Cross-Validation using Training Data; Validation using test data	For Prediction: t-Test (Ranking of Genes) stepwise discriminant analysis Validation using test data	Clustering: Fuzzy c-means Rule Extraction: Empirical Fuzzy Rule Rating of uniconditional rules Rulebase Construction for Prediction: "informative imbalance" Validation using test data
Informative Genes	about 1100 (higher correlated with "Class Distinction" than by chance)	500 (based on t-Test Ranking, 250 genes with most positive and 250 with most negative t-values; number arbitrary)	4.126 419 (for AML) 3.707 (for ALL) (based on the use of the "dummy gene" in the rule sets)
Number of Genes Used for Prediction	50 (arbitrary, no specific method)	25 after 1st discriminant analysis 5 after 2nd discriminant analysis	13 (using the "informative imbalance" criterion starting from 38 genes)
Number of Patients predicted correctly out of the 38 training patient samples	36	38	38
Number of Patients predicted correctly out of the 34 test patient samples	29 (using 50 genes)	32 (using 25 genes) 33 (using 5 genes)	33 (using 13 genes)
Prediction Genes Obtained by all Methods	2 [X95735_at; M84526_at]		
Prediction Genes Obtained by any Two methods	3 (between the 5 and 50 gene model) [M27891_at; M84526_at; X95735_at] 6 (between the 25 and 50 gene model) [M27891_at; M84526_at; X95735_at; M13792_at; M62762_at; Y00787_s_at]		
		2 (between the 5 and 13 gene model) [X95735_at; M84526_at] 2 (between the 25 and 13 gene model) [X95735_at; M84526_at]	
	[X95735_at; U46751_at; M84526_at; M23197_at; M96326_rna1_at	10 (between the 50 and the 13 ← gene model) →	M55150_at; U82759_at; U22376_cds2_s_at; U05259_rna1_at; U50136_rna1_at]

Table 2: Summary of Methods and Results of the 3 Approaches Compared

Training Patient Sample No.	Actual Cancer Status	Predicted Cancer Status			
		Golub et al.1999 [1]	*Lu* et al. 2002 [3]		Present Publication
		50 Gene Model	25 Gene Model	5 Gene Model	13 Gene Model
01	ALL	ALL	ALL	ALL	ALL
02	ALL	ALL	ALL	ALL	ALL
03	ALL	ALL	ALL	ALL	ALL
04	ALL	ALL	ALL	ALL	ALL
05	ALL	ALL	ALL	ALL	ALL
06	ALL	ALL	ALL	ALL	ALL
07	ALL	ALL	ALL	ALL	ALL
08	ALL	ALL	ALL	ALL	ALL
09	ALL	ALL	ALL	ALL	ALL
10	ALL	ALL	ALL	ALL	ALL
11	ALL	ALL	ALL	ALL	ALL
12	ALL	***AML***	ALL	ALL	ALL
13	ALL	ALL	ALL	ALL	ALL
14	ALL	ALL	ALL	ALL	ALL
15	ALL	ALL	ALL	ALL	ALL
16	ALL	ALL	ALL	ALL	ALL
17	ALL	ALL	ALL	ALL	ALL
18	ALL	ALL	ALL	ALL	ALL
19	ALL	ALL	ALL	ALL	ALL
20	ALL	ALL	ALL	ALL	ALL
21	ALL	ALL	ALL	ALL	ALL
22	ALL	ALL	ALL	ALL	ALL
23	ALL	ALL	ALL	ALL	ALL
24	ALL	ALL	ALL	ALL	ALL
25	ALL	ALL	ALL	ALL	ALL
26	ALL	ALL	ALL	ALL	ALL
27	ALL	ALL	ALL	ALL	ALL
28	AML	AML	AML	AML	AML
29	AML	AML	AML	AML	AML
30	AML	AML	AML	AML	AML
31	AML	AML	AML	AML	AML
32	AML	AML	AML	AML	AML
33	AML	AML	AML	AML	AML
34	AML	AML	AML	AML	AML
35	AML	***ALL***	AML	AML	AML
36	AML	AML	AML	AML	AML
37	AML	AML	AML	AML	AML
38	AML	AML	AML	AML	AML

Table 3.1: Actual and Predicted Cancer Status for the Individual 38 Training Patient Samples. Prediction Results for the 4 Different Prediction Models Compared.

Test Patient Sample No.	Actual Cancer Status	Predicted Cancer Status			
		Golub et al.1999 [1]	*Lu* et al. 2002 [3]		Present Publication
		50 Gene Model	25 Gene Model	5 Gene Model	13 Gene Model
39	ALL	ALL	ALL	ALL	ALL
40	ALL	ALL	ALL	ALL	ALL
41	ALL	ALL	ALL	ALL	ALL
42	ALL	ALL	ALL	ALL	ALL
43	ALL	ALL	ALL	ALL	ALL
44	ALL	ALL	ALL	ALL	ALL
45	ALL	ALL	ALL	ALL	ALL
46	ALL	ALL	ALL	ALL	ALL
47	ALL	ALL	ALL	ALL	ALL
48	ALL	ALL	ALL	ALL	ALL
49	ALL	ALL	ALL	ALL	ALL
50	AML	AML	AML	AML	AML
51	AML	AML	AML	AML	AML
52	AML	AML	AML	AML	AML
53	AML	AML	AML	AML	AML
54	AML	***ALL***	AML	AML	AML
55	ALL	ALL	ALL	ALL	ALL
56	ALL	ALL	ALL	ALL	ALL
57	AML	***ALL***	AML	AML	AML
58	AML	AML	AML	AML	AML
59	ALL	ALL	ALL	ALL	ALL
60	AML	***ALL***	AML	AML	AML
61	AML	AML	***ALL***	AML	AML
62	AML	AML	AML	AML	AML
63	AML	AML	AML	AML	AML
64	AML	AML	AML	AML	AML
65	AML	AML	AML	AML	AML
66	AML	***ALL***	***ALL***	***ALL***	***ALL***
67	ALL	***AML***	ALL	ALL	ALL
68	ALL	ALL	ALL	ALL	ALL
69	ALL	ALL	ALL	ALL	ALL
70	ALL	ALL	ALL	ALL	ALL
71	ALL	ALL	ALL	ALL	ALL
72	ALL	ALL	ALL	ALL	ALL

Table 3.2: Actual and Predicted Cancer Status for the Individual 34 Test Patient Samples. Prediction Results for the 4 Different Prediction Models Compared.

A Note on Core Regions of Membership Functions

Jürgen Paetz
J.W. Goethe-Universität Frankfurt am Main,
Fachbereich Biologie und Informatik, Institut für Informatik
Robert-Mayer-Straße 11-15, D-60054 Frankfurt am Main, Germany
Phone: 0049-069-798-25119, Fax: 0049-069-798-28353
email: paetz@informatik.uni-frankfurt.de

ABSTRACT: In neuro-fuzzy approaches different membership functions are used for modeling the system's rule set. Two wellknown membership function types are triangle functions and trapezoid functions. In our contribution we demonstrate that trapezoid functions with larger core regions are the more appropriate functions for calculating the membership degrees within neuro-fuzzy systems. If regions of the data of different classes are highly overlapping or if the data is noisy, the values of the membership degrees could be misleading with respect to rule confidence if the core region is modeled too small. In fact, we show that data regions with a high membership degree need not to be the regions with a high rule confidence. This effect that we call membership unrobustness is discussed. We give preliminary benchmark examples and show how this effect influenced our recent work of analysing septic shock patient data.

KEYWORDS: neuro-fuzzy system, membership function, core region, rule confidence, robustness, medical data

INTRODUCTION

An up to date neuro-fuzzy system is the so-called Fuzzy-RecBF-DDA [1], [2], that we recently used in an improved implementation for rule generation and classification in the medical domain [3] (analysis of septic shock patient data). The system is based on the dynamic, geometric adaptation of trapezoids. Another wellknown system is NEFCLASS [4] that is based on a fuzzy backpropagation adaptation process. Fuzzy-RecBF-DDA uses trapezoid membership functions. NEFCLASS uses triangle membership functions. We will not discuss the main differences of the two systems and we will not decide which system is the better one. The interested reader should refer to the literature if he/she is interested in the details of the different systems.

Our approach is a general comparison of the main two types of membership functions: the trapezoid and the triangle membership function, mainly their core regions. Usually for adapting neuro-fuzzy systems the data is divided into training and test data. The training data is used to train the system, the test data is used to calculate the performance of the system. If the data is noisy or if the data of different classes is highly overlapping or if there are not much samples for a performant generalization available - an usual szenario in real world applications - then the statistical properties of the training set may differ from those of the test set. We argue that in these situations trapezoid membership functions with larger core regions should be preferred.

In the next section we define the membership functions and explain the training and test procedure. Then, by some examples we show that it is more reasonable to use trapezoid membership functions with larger core regions instead of triangle membership functions. Particularly, we discuss our more fundamental considerations by septic shock patient data that was recorded in intensive care units from 1993 to 1997 at the Klinikum der J.W. Goethe-Universität Frankfurt am Main [5].

PRELIMINARIES

After presenting the definition of some important fuzzy theory terms we define triangle and trapezoid membership functions. Then, we describe how we will divide the data into training and test data for the adaptive phase of a neuro-fuzzy learning.

Intelligent e-Health Applications in Medicine
Special Session Proc. EUNITE 2002, Albufeira, Portugal, Sept. 19-21, 2002
Published by the University of the Aegean

MEMBERSHIP FUNCTIONS

In Def. 1 we explain the important terms *membership function* and *membership degree*. We define the terms α*-cut*, α*-region* and α*-rule*.

Definition 1:
Let X be a set. A is called a *fuzzy set* if there exist a corresponding *membership function* $\mu : X \to [0,1]$ that is defined everywhere on X. The value $\mu(x)$ is called *membership degree* of $x \in X$. The region in the data space where $\mu(x) = \alpha$ we call α*-cut* if we consider the membership function. The corresponding geometric region we call α*-region*. If additionally the whole rule with the conclusion is considered we speak about α*-rules*. In the special case of $\alpha = 1$, i.e. when considering the 1-cut of a membership function, we speak about *core regions* resp. *core rules*.

In the following example we give the definitions of triangle and trapezoid membership functions in the one-dimensional case, cf. Figure 1.

Example 1:
a) Triangle membership function (1) with $a_1, a_2, b \in I\!R$:

$$\mu_{triangle}(x) := \begin{cases} 1 & , \quad x = b \\ \dfrac{x - a_1}{b - a_1} & , \quad x \in [a_1, b) \\ \dfrac{a_2 - x}{a_2 - b} & , \quad x \in (b, a_2] \\ 0 & , \quad otherwise \end{cases} \tag{1}$$

b) Trapezoid membership function (2) with $a_1, a_2, b_1 b_2 \in I\!R$:

$$\mu_{trapezoid}(x) := \begin{cases} 1 & , \quad x \in [b_1, b_2] \\ \dfrac{x - a_1}{b_1 - a_1} & , \quad x \in [a_1, b_1) \\ \dfrac{a_2 - x}{a_2 - b_2} & , \quad x \in (b_2, a_2] \\ 0 & , \quad otherwise \end{cases} \tag{2}$$

In the m-dimensional case we obtain the membership degree of $x = (x_1, \ldots, x_m)$ by calculating the minimum of the one-dimensional membership degrees of $x_1, \ldots, x_m$ considering the one-dimensional projections of the n-dimensional membership function.

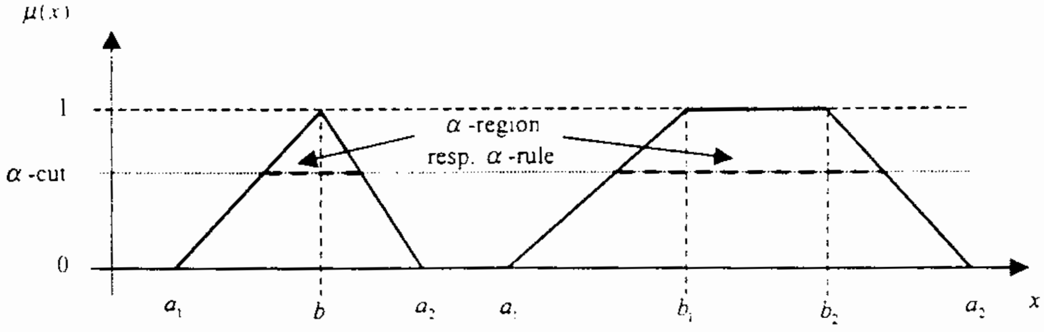


Figure 1: (Left) triangle membership function (Right) trapezoid membership function. The dotted line represents the α-cut. The dashed line in the height of the α-cut represents the α-region resp. the α-rule, the latter if considered with a class label for conclusion.

With d as the dimension of the space the triangle membership functions have $2d$ free parameters. The trapezoid membership functions have $4d$ free parameters. With respect to free parameters the triangle membership functions are the more simpler functions. In the case of $b_1 = b_2$ the trapezoid function is identical to the triangle function.

TRAINING AND TEST

Training a system has to be done carefully. To avoid overfitting the data is usually divided into *training* and *test data*. If we would take all the data for training the system, it would be very well adapted to the training data, but it would not be able to generalize well, i.e. having a good performance on unknown data. Calculating the system's performance on test data is a good way to evaluate a system's generalization ability. Sometimes even more sophisticated data divisions are used (additional validation data, k-fold cross validation, see e.g.[6]).

The statistical properties of the training data differ from those of the test data. If the data environment is noisy or if there are not many samples available the difference is larger. This is the reason why the generalization performance should be calculated only on the test data. If the system is able to tolerate noise and other, different samples then it is well trained and generalizes well.

To calculate the rule performance of an α-rule, i.e. hyper rectangular data region with class label (cf. Def. 1), we define the *rule confidence* as the number of samples of the correct class that lie in the hyper rectangle divided by the number of all the samples that lie in the hyper rectangle. For example, consider the two dimensional unit square. Let there be 4 samples of the class c lying in this square and one of another class $d \neq c$, then the confidence of this α-rule for the class c is 0,8 resp. 80 %.

MEMBERSHIP DEGREES OF TRIANGLE AND TRAPEZOID MEMBERSHIP FUNCTIONS

In this section we present examples which show that it is possible to obtain unexpected results with regard to the rule confidence by using too small core regions together with training and test data.

EXAMPLES FOR "MEMBERSHIP UNROBUSTNESS"

For the first example an artificial dataset is used, motivating the problem of membership unrobustness. For the following examples we use real world datasets to demonstrate the relevance of this problem in practice.
In Example 4 we discuss the problem area by our actual septic shock patient data.

Example 2:
Let us have a more detailed look on the triangle membership function, defined in (1). For easier calculation by hand it is symmetric in our example, i.e. $b - a_1 = a_2 - b$. The upper angle is set to 90°, the lower left and right each to 45°. Thus, the sides of the triangle have a gradient of 1 resp. -1. Let us consider $n_1 = 101$ one-dimensional uniformly distributed samples of the same class c in the interval $[-1,1]$. Add $n_2 = 101$ identical samples with the exception that one sample (the one in the middle, i.e. the one placed at $b = 0$) is a sample of another class d, e.g. an outlier (noise). Let us assume that we have chosen the first $n_1 = 101$ samples for training and the $n_2 = 101$ samples for testing, so that we have obtained the above desribed symmetric triangle membership functions with $a_1 = -1, a_2 = 1, b = 0$.

Now, we calculate the rule confidence for all rectangular regions, considering the membership degrees 0.00;0.02;0.04;...;0.96;0.98;1.00. Because of the outlier at $b = 0$ the rule confidence for class c for the region with membership degree α (α-rule) is $\dfrac{100(1-\alpha)}{101-100\alpha}$ %.

In Figure 2 the rule confidence in relation to α is shown. The higher the membership degree is in Figure 2, the lower is the rule confidence that was calculated on the test data. The minor difference in training and test data (one outlier) leads to a result that is surprising. What we expected is an ascent of rule confidence and not a descent. In our example this unwanted effect appears because the core region of the triangle is a single point only. Thus, only one single point in the test set needs to differ and the result is not reliable.

If we use the trapezoid membership function with a core region of a size covering nine test data samples (including the outlier), the rule confidence at membership degree 1 is 88.89 %, the lowest value for the rule confidence considering all membership degrees in this case, but a much better value than in the triangle case (0 %).

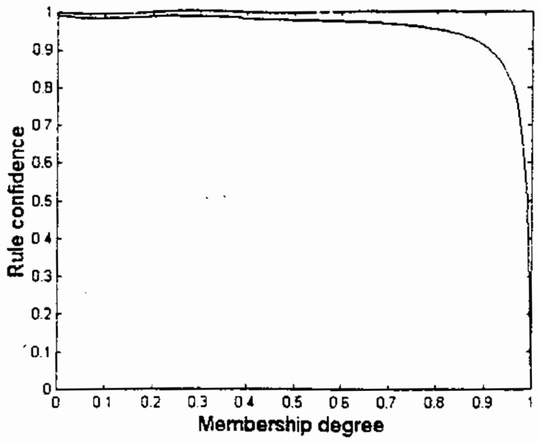


Figure 2: Rule confidence in relation to membership degree. In this example the rule confidence is monotone *descending*.

Now, one could argue that this is a very artificial dataset, but unfortunately the situation remains the same on real world data, see next section. We call this effect *membership unrobustness.*

Example 3:
To demonstrate the effect of membership unrobustness we calculated - in a similar way as in Example 2 - the rule confidence for asymmetric trapezoid membership functions obtained by training the system [3]. As benchmark data we used datasets from the Proben archive [7]. Here, we present the results using the dataset Diabetes1 with overlapping class regions. Approx. 25 % of the data could not be classified correctly. In Figure 3 the rule confidence (mean of all the rules) in relation to the membership degree is shown for both classes of Diabetes1. A higher rule confidence for class one is obtained at lower membership degrees (0.0 to 0.4). The best rule confidence for class two is obtained at the membership degree 0.6 . Thus, the size of the core rules should not be smaller than the size of the rules with these corresponding membership degrees.

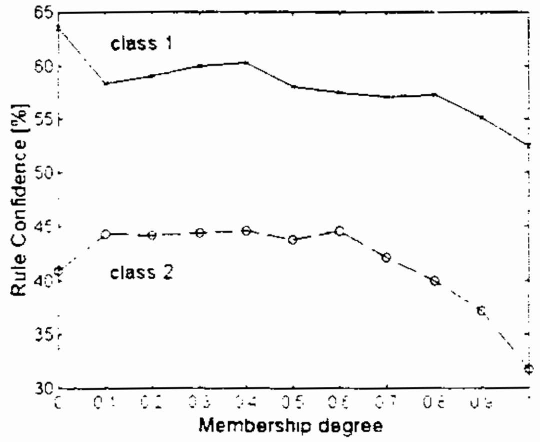


Figure 3 Mean rule confidence (in %) in relation to the membership degree for the two classes (Diabetes1 dataset).

Using too small core regions, the effect of membership unrobustness is very hazardous when cutting α-rules from the trapezoids at higher membership degrees. These α-rules have in fact a lower rule confidence than α-rules of lower membership degrees. Of course, if we would use the triangle membership functions instead of trapezoid membership functions the effect of membership unrobustness would even be stronger since one single point (the top of the triangle) is the smallest not empty core region that is possible.

Example 4:

Let us consider another, more important real world example. In abdominal intensive care medicine patients are in a very critical condition. Often patients develop a septic shock that is associated with a high lethality of about 50 %. It is always related to measurements leaving the normal range (e.g. blood pressure, temperature, respiratory frequency, number of leukocytes), and it is often related to multiorgan failure. We consider 70 patients who developed a septic shock during their stay at the intensive care unit of J.W. Goethe University Clinic. The data was collected from 1993 to 1997 [5]. 38.6 % of the 70 septic shock patients deceased.

To demonstrate the effect of membership unrobustness let us consider the dataset F_{12}, containing the measurements of the 12 most frequently measured variables: creatinin [mg/dl], calcium [mmol/l], arterial pCO_2 [mmHg], pH [-], haematocrit [%], sodium [mmol/l], leukocytes [1000/µl], haemoglobin [g/dl], central venous pressure (CVP) [cmH_2O], temperature [°C], heart frequency [1/min], systolic blood pressure [mmHg]. To limit the influence of missing values, we demanded the existence of a minimum of 10 out of 12 variables for each sample, so that 1698 samples remained out of 2068 (1177 survived, 521 deceased).

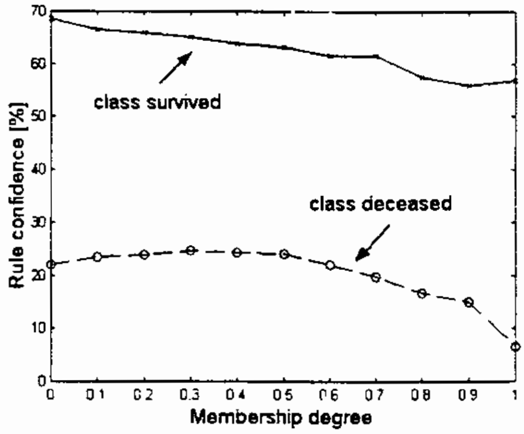


Figure 4: Mean rule confidence (in %) in relation to the membership degree for the two classes (dataset F_{12}).

In Figure 4 we see that the highest rule confidence is obtained at membership degree 0 (0 -rules) for class survived and at membership degree 0.3 (0.3 -rules) for class deceased. This is another example for membership unrobustness using too small core regions. Since the data probability of class deceased is lower than for class survived, the mean values of rule confidence for class survived are naturally higher than for class deceased.

To get an impression of the results, we present one α -rule for the class survived at membership degree $\mu = 0.3$ and one at membership degree $\mu = 1.0$. Rule 1 has a higher confidence on the test data than rule 1*. Thus, rule 1 instead of rule 1* should be presented to a physician. Another benefit of rule 1 is the higher frequency on test data. Due to the lower (statistically too low!) frequency of rule 1* on test data, also the confidence on test data is not as high as for rule 1. In rule 1 resp. 1* the variables systolic blood pressure, arterial pCO_2, haemoglobin, creatinin and sodium are not relevant. Technical note: not all the values of the variables for different membership degrees need to differ (e.g. heart frequency). This is due to algorithm [3] where the trapezoid shape may degenerate to a rectangular shape for some dimensions in the positive and/or negative direction.

Rule 1 ($\mu = 0.3$):
if var heart frequency ≤ 110.16 **and**
var CVP ≥ 7.00 **and**
var pH ≥ 7.37 **and**
var temperature in (34.45,37.42) **and**
var leukocytes ≤ 21.50 **and**
var haematocrit ≥ 26.04 **and**
var calcium ≤ 2.50
then class survived **with** frequency 0.169
and confidence 0.669

Rule 1* ($\mu = 1.0$):
if var heart frequency ≤ 110.16 **and**
var CVP ≥ 7.00 **and**
var pH ≥ 7.38 **and**
var temperature in (37.58,36.78) **and**
var leukocytes ≤ 19.45 **and**
var haematocrit ≥ 28.44 **and**
var calcium ≤ 2.50
then class survived **with** frequency 0.077
and confidence 0.627

Another rule example (rule 2 resp. 2*) for class deceased show the case that no test data sample is implied by a core rule with membership degree $\mu = 1.0$, i.e. frequency $= 0$. This happens every time the rule defining hyper rectangle is statistically too small. Core rule 2* demonstrates that it is not reliable to model too small regions of higher membership degrees. In this case only zero confidence could be attributed for rule 2*.

Rule 2 ($\mu = 0.3$):
if var systolic blood pressure ≥ 110.80 **and**
var CVP ≥ 4.50 **and**
var leukocytes ≥ 16.72 **and**
var sodium ≤ 146.28 **and**
var calcium ≤ 2.19
then class deceased **with** frequency 0.047
and confidence 0.512

Rule 2* ($\mu = 1.0$):
if var systolic blood pressure ≥ 139.20 **and**
var CVP ≥ 15.00 **and**
var leukocytes ≥ 23.56 **and**
var sodium ≤ 131.08 **and**
var calcium ≤ 1.89
then class deceased **with** frequency 0.000
and confidence 0.000

It is possible to generate fuzzy antecedents out of the trapezoids, but this topic we will not discuss here. With the help of such classification rules we have recently built an internet based alarm system to warn physicians whenever a septic shock patient becomes very critical [8].

RECOMMENDATION

As we demonstrated triangle membership functions and trapezoid membership functions with too small core regions are not appropriate for modeling regions (resp. α-rules) where the rule confidence should be very high both for training and test data. Of course, the solution for this dilemma is not to allow the system to generate too small core regions. Since the core region is a single point when using the triangle membership function, it represents the most confident data region in the worst possible way. The region with the highest membership degree represents usually not the region with the highest rule confidence on real world data. Larger regions are statistically more appropriate to model regions of a high membership degree. Therefore, we recommend the use of trapezoid membership functions with larger core regions rather than trapezoid membership functions with small core regions or even triangle membership functions where the core region is a single point only.

CONCLUSION

We discussed the usage of triangle and trapezoid membership functions. Both types of functions are reliable in practice. The main disadvantage using membership functions with too small core regions (e.g. triangle functions) considering applications in practice (real world data) is the less rule confidence of the α-regions resp. α-rules of higher membership degrees α. In fact, for membership degree 1 the corresponding rule confidence could be the worst. Thus, the high membership degree is misleading with respect to rule confidence. By different examples we showed that it is more reliable to use trapezoid functions, avoiding too small core regions. For every dataset and every rule exists an individual best, not too small size of the core region. Finding the optimal size of core regions of membership functions by calculation or adaptation is still an open problem.

Acknowledgement: My work was partially done within the project MEDAN [8], supported by the DFG.

REFERENCES

[1] Berthold, Michael; Huber, Klaus-Peter, 1995, "From Radial to Rectangular Basis Functions: A New Approach for Rule Learning from Large Datasets", Internal Report 15-95, Univ. Karlsruhe.

[2] Huber, Klaus-Peter; Berthold, Michael, 1995, "Building Precise Classifiers with Automatic Rule Extraction", Proc. of the IEEE Int. Conf. on Neural Networks (ICNN), Perth, Australia, vol. 3, pp. 1263 - 1268.

[3] Paetz, Jürgen, 2001, "Metric Rule Generation with Septic Shock Patient Data", Proc. of the 1st IEEE Int. Conf. on Data Mining (ICDM), San Jose/CA, USA, pp. 637 - 638.

[4] Nauck, Detlef; Kruse, Rudolf, 1997, "A Neuro-Fuzzy Method to Learn Fuzzy Classification Rules from Data", Fuzzy Sets and Systems, vol. 89, pp. 277 - 288.

[5] Wade, Sabine; Büssow, Michael; Hanisch Ernst, 1998, "Epidemiology of SIRS, Sepsis and Septic Shock in Surgical Intensive Care Patients", Chirurg, vol. 69, pp. 648 - 655.

[6] Haykin, Simon, 1999, "Neural Networks: A Comprehensive Foundation", Prentice Hall, 2nd ed.

[7] Proben data archive ftp://ftp.informatik.uni-freiburg.de/documents/papers/neuro/proben/proben1/ .

[8] Hanisch, Ernst; Brause, Rüdiger, Project homepage MEDAN (Medical Data Analysis with Neural Networks) www.medan.de .

Classification of Eyetracking Signals with Vector-Based Neural Networks

Martin Golz, David Sommer
University of Applied Sciences Schmalkalden
Department of Computer Science
D-98574 Schmalkalden, Germany
{golz, sommer}@informatik.fh-schmalkalden.de

ABSTRACT: The eye gaze point and the pupil size of five subjects were recorded during an overnight driving simulation task. By scoring the recorded videos clear microsleep events (MSE) and clear non-microsleep events were picked out and the measured signals in the preceding eight seconds were analyzed. The spectral power densities of these segments were classified using Learning Vector Quantization, Self-Organizing Map and Growing Cell Structures. For the latter two networks the supervised and the unsupervised version were applied. Best results were obtained with a modified LVQ3 network.

KEYWORDS: Learning Vector Quantization, Self-Organizing Map, Growing Cell Structures, Eyetracking

INTRODUCTION

Many authors suggested the measurement of the pupil size and the eye movements to estimate a subject's alertness level [5, 9, 10, 11, 12]. The first three groups used electrooculography (EOG); the later two groups used infrared corneal reflection as measurement principle.

The present study employed an eyetracking systems based on the analysis of the combined corneal and foveal reflection [1]. Our intention was not to estimate the alertness level at a time scale of some minutes, but to explore characteristics of the eyetracking signals immediately before the onset of a microsleep.

Five individuals with an age between 19 and 28 years participated in a monotonically overnight driving simulator study. Every hour from 1 a.m. until 7 a.m. one driving session of 25 min length was carried out. The portrait of the driver and the right eye was video recorded. The eyetracker was working in the near infrared with an accuracy of 0.65 deg and measured the pupil diameter (D), and the horizontal (X) and the vertical (Y) component of the eye gaze point in the plane of the driving simulator screen with a sampling rate of 30 Hz.

Microsleep events (MSE) were visually scored off-line by an expert using the video recordings and simultaneously EEG recordings. Clear non-microsleep events (NMSE) were scored in the same manner.

The X- and Y-signals had a series of missing values during eye blinks. They were substituted with Beziér spline interpolation. Additionally outlier elimination was necessary, especially for the Y-signal immediately after an eye blink.

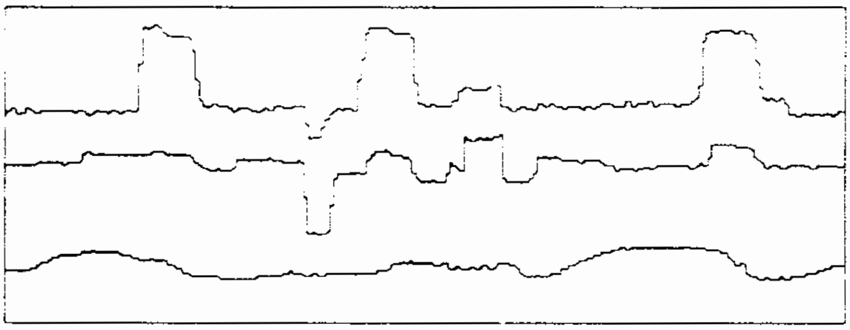


Figure 1: An 8 sec segment of the pupil diameter D (lower graph) and of the eytracker signals X and Y (upper graph).

Intelligent e-Health Applications in Medicine
Special Session Proc. EUNITE 2002 Albufeira Portugal Sept 19-21, 2002
Published by the University of the Aegean

603 segments of the length of 8 sec were taken out of all three signals (X,Y,D) immediately before a MSE and before a NMSE (Fig. 1). Afterwards any linear trend was eliminated and a we.ch window was applied to improve the results of the following Discrete Fourier Transform. The reduction of the total power density due to the windowing was corrected using the Parceval Theorem.

For each signal we got 80 spectral power density samples in the range from 0 to 9.9 Hz with a resolution of 0.125 Hz. All spectral samples were used as an input vector for the neural networks.

To build up a classifier for input vectors of the MSE- and of the NMSE-class we applied three types of vector-based neural networks: the Learning Vector Quantization network (LVQ) [7], the Self-Organizing Maps (SOM) [6] and the Growing Cell Structures (GCS) [3].

The LVQ networks are trained supervised; here the binary information MSE / NMSE was used as teaching input. Kohonen suggested three modifications LVQ1, LVQ2 (LVQ2.1) and LVQ3. The first modification uses an adapted step size, whereas LVQ2 leads to an adaptation of neurons in interclass regions. LVQ3 additionally allows a slight adaptation of weight vectors in intraclass regions [8].

The SOM and the GCS networks are trained unsupervised. After training, both network types were calibrated with the binary MSE / NMSE information. SOM tries to minimize the error of vector quantization and to some extend to find a discrete approximation of the probability density function of the input vectors. GCS are incremental neural networks and are with some restrictions capable to approximate the probability density function of the input vectors. The topological structure is a k-simplex. We chose k = 1 and k = 2 to be able to visualize. Two-dimensional rectangular and one-dimensional topologies were applied to the SOM networks.

RESULTS

Each network was trained with several parameter settings and with several initializations of the weight vectors. Before each training the learning set was randomly partitioned in training set (80%) and in test set (20%). After training had finished, the reclassification rate was estimated by the ratio of correct classified to all applied input vectors of the training set. The classification rate was estimated in the same way with input vectors taken from the test set.

We calculated the classification and reclassification rates in 1.7 10^6 different network simulations with different parameter settings, like number of neurons, learning rate factor and parameters of the neighborhood function and different variables selections for the input vectors and different learning set partitions. Fig. 2 shows an example.

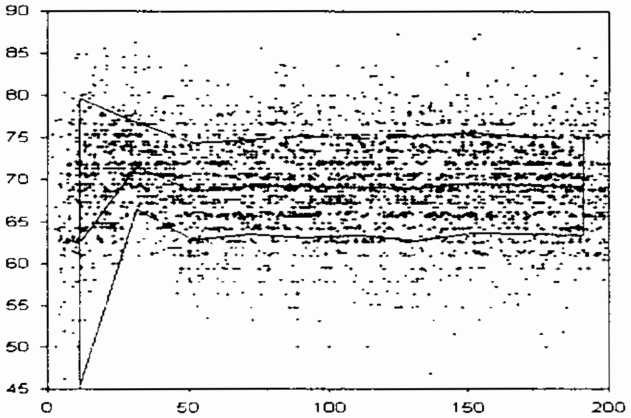


Figure 2: Test-set classification rate (in percent) versus number of neurons for an LVQ3 network. The input vectors contain spectral power densities of the pupil diameter D only. The lines indicate the mean ± standard deviation range.

The optimal number of neurons ranged between 8 and 20. With an increasing number of neurons the LVQ network shows a better adaptation to the training set, the reclassification rate is mostly above 90%; but it shows a decreasing ability to generalize indicated by decreasing classification rates.

The average maximum classification rate was obtained by searching the maximum of the mean + standard deviation (upper curve in Fig. 2) for all different settings of the LVQ networks (Tab. 1). Initialization with median means that we assigned to each component of the weight vectors the median value of this component over all input vectors. During data driven initialization each weight vector was assigned to a randomly selected input vector. Furthermore, in the first

30% of all training iterations the network was trained disregarding the class membership of an input vector to diminish the variance of the classification rate [4].

Network	Initialization	Scaling	D	Y	X	DX	DXY
LVQ1	median	none	77	69	70	71	72
LVQ1	data driven	none	77	68	71	70	71
LVQ1	data driven	square root	76		71	72	
LVQ1	data driven	normalized	75		72	75	
LVQ2	median	none	77		75	74	
LVQ2	data driven	none	79		75	74	
LVQ3	median	none	80		75	74	
LVQ3	data driven	none	80		75	75	
LVQ3	data driven	square root	77		73	74	
LVQ3	data driven	Normalized	73		75	79	

Table 1:. Average maximum test-set classification rate (in percent) with different LVQ networks, different initializations and different scaling applied on different feature sets. (For details see text)

We tried a number of different scaling, but we want to report only the results of no scaling, the square root of each input vector component and the normalization of each component with respect to the sum of all components (relative value). In the columns 'D', 'Y', 'X' the input vectors consisted only of the spectral power densities of the D, Y and X signal respectively. In column 'DX' all spectral values of the D and X signal, and in 'DXY' all spectral values of the D, X and Y signal were used. The best results were obtained with the set of input vectors obtained from the D signal only. Apparently, if we add further components to the input vectors as in the columns 'XD' and 'XYD', the results are not improvable. On the one hand we presented supplementary and independent information to the neural networks, but on the other hand the number of dimensions of the input space was obviously too much.

Typically SOMs calibrated with the MSE and NMSE information are shown in Fig. 3 and Fig. 4 using the U-matrix [13]. The U-matrix represents distances of topological neighbored weight vectors in the input space and is visualized as gray shades. Larger distances of neighbored weight vectors are visualized by darker gray shades. Weight vectors in the NMSE regions show larger distances than weight vectors in MSE regions. Under the assumption that the SOM has found a correct approximation of the probability density function of the input vectors this indicates that the MSE class has a higher probability density and is more compact. The Self-Organizing Map of Fig. 3 also shows that the NMSE input vectors are mapped preferably to the left part and to the lower part of the map, whereas the MSE input vectors are mapped to the right upper part of the map. Between both regions there is a region of overlapping classes.

The differentiated U-matrix (Fig. 4) roughly shows the region of overlapping classes with light shades. The two classes are distributed in only two more or less compact and overlapping regions in the input space. This could explain the decreasing ability of generalization with increasing number of neurons and the onset of this effect at small numbers of neurons already.

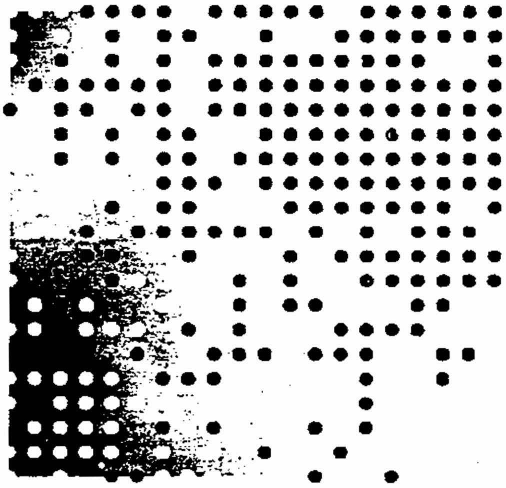


Figure 3: Typical calibrated SOM. Gray shades indicating the U-matrix. Microsleep events (dark nodes) and non-microsleep events (light nodes) are separable with some limitations. Vacancies indicating dead neurons.

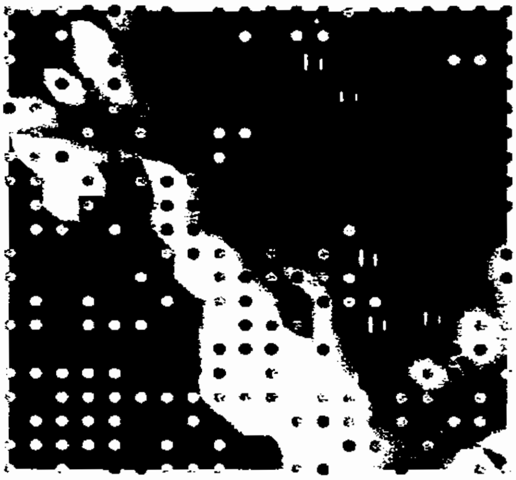


Figure 4: The same SOM with a differentiated U-matrix. Gray shades indicating the U-matrix. Microsleep events (dark nodes) and non-microsleep events (light nodes) are separable with some limitations. Vacancies indicating dead neurons.

The GCS networks were trained and tested with the same method as the SOM. Additionally, there is a fast learning by inserting and deleting neurons depending from a local criterion. Two criteria were proposed [2]: the mean vector quantization error (vqe) and the local probability density function (pdf). For the calculation of the pdf the volume of the n-dimensional voronoi cell was approximated with the volume of the n-dimensional hypercube, generated with the mean local weight vector distance [2].

Network	No. of neur.	dim.	Criter.	D	Y	X	DX	DXY
SOM	20 x 1	1		76		72	70	
SOM	20 x 10	2		74		68	68	
SOM	20 x 20	2		72		66	67	
sv SOM	20 x 1	1		70		67	65	
sv SOM	20 x 20	2		69		62	62	
GCS	300	1	pdf	74		74	70	
GCS	300	1	vqe	75		69	69	
GCS	300	2	pdf	74		69	69	
GCS	300	2	vqe	74		68	68	
sv GCS	300	1	pdf	70		69	70	
sv GCS	300	1	vqe	68		66	65	

Table II:. Average maximum test-set classification rate (in percent) with SOM and GCS networks, different number of neurons applied on different feature sets. Supervised modifications are marked with 'sv'. (For details see text)

Both networks, the SOM and the GCS, came to lower average maximum classification rates compared to LVQ (Tab. II). This is not surprising because their training is unsupervised.

Following the suggestions in [8] a supervised training for SOM was applied. The input vectors are concatenated during training with a binary unit vector of encoded class number. Every class is assigned to one component of the unit vector. In the recall phase the input vectors are applied without the unit vector. In two- and three-dimensional examples one can see, that this modification leads to a higher density of prototype vectors in interclass regions. This idea was applied to the GCS network in the same manner.

Surprisingly, the supervised versions showed poorer classification rates, approximately 5% below the results of the unsupervised version. Compared to the unsupervised version the ability to generalize was poorer and the adaptation (training set classification rate) was improving.

With SOM and with GCS the best results were obtained processing D data only and mapping on one-dimensional topology. In this case it is not considerable to choose vqe or pdf as fast learning criterion function. If pdf and one-dimensional topology was chosen the results were about equal for signal D and for X.

The visualization of the topology yielded no results. Between one and three separate topological nets grew during training. No net contained a large majority of input vectors of the MSE class.

REFERENCES

[1] Cleveland, D; Cleveland, N (1992). Eyegaze Eyetracking System. Imagina – 11th Monte-Carlo International Forum on New Images; Monte-Carlo, USA.

[2] Fritzke, B (1992). Wachsende Zellstrukturen – ein slebstorganisierendes neuronales Netzwerkmodell. University of Erlangen, PhD thesis (german).

[3] Fritzke, B (1994). Growing Cell Structures - a self-organizing network for unsuper-vised and supervised learning. Neural Networks, 7, 1441-1460.

[4] Golz, M; Sommer, D; Lembcke, T; Kurella, B (1998). Classification of the pre-stimulus-EEG of k-complexes using competetive neural networks. Proc. Of the 6th European Congress on Intelligent Techniques and Soft Computing. EUFIT'98. Verlag Mainz, Aachen, Germany, Vol. 3, 1767-71

[5] Hyoki, K; Shigeta, M.;Tsuno, M; Kawamuro, Y; Kinoshita, T (1998). Quantitative electro-oculography and electroencephalography as indices of alertness. Electroenc. Clin. Neurophysiol., 106, 3, 213-219.

[6] Kohonen, T (1982). Self-organized formation of topologically correct feature maps. Biol Cybern, 43, 59-69.

[7] Kohonen, T (1988). Learning Vector Quantization. Neural Networks, 1 (Suppl.1), 303

[8] Kohonen, T (2000). Self-Organizing Maps. 3rd Ed. Springer Series in Information Sciences, Vol. 30, Springer, Berlin, Heidelberg, New York.

[9] McPartland, RJ; Kupfer, D (1978). Computerized measures of electro-oculographic activation during sleep. Int J Biomed Comput, 9, 409-419.

[10] Ogilvie, RD; McDonagh, DM; Stone, SN; Wilkinson, RT (1988). Eye movements and the detection of sleep onset. Psychophysiology, 25, 81-91.

[11] Saito, S (1992). Does fatigue exist in a quantitative measurement of eye movements? Ergonomics, 35,5/6,607-615.

[12] Schmidt, D; Abel, FA; Dell'Osso, LF; Daroff, RB (1979). Saccadic velocity characteristics: intrinsic variability and fatigue. Aviation, Space and Environmental Medicine, 50, 1979, 393-395

[13] Ultsch, A; Siemon, HP (1989). Exploratory data analysis: using Kohonen networks on transputers. Univ. of Dortmund, Technical Report 329, Dortmund, Dec. 1989.

Multiple Training of Vector-Based Neural Networks to detect Density Centers in Input Space

David Sommer, Martin Golz
University of Applied Sciences Schmalkalden
Department of Computer Science, Postfach 182
D-98574 Schmalkalden, Germany
{sommer, golz}@informatik.fh-schmalkalden.de

ABSTRACT: In this paper we will propose a method for enhanced evaluation of classified input vectors. For an arbitrary noisy and labeled data set, which can have several maxima of the probability density function (density centers), an automatic distinction of input vectors into "good classifiable" and "bad classifiable" is done. In a first training and testing phase (filter phase) the goodness of each input vector is calculated by means of vector based artificial neural networks, typically Learning Vector Quantization. In a second phase (evaluation phase) only "good classifiable" input vectors are used to train a Self-Organizing Map (SOM). For an artificial and a real world data set with compact but overlapping probability density functions it is shown that the complexity of the SOM output space is reduced. In some applications it might be possible to enhance the generalization properties of the SOM classifyer by using only "good classifiable" input vectors during training.

KEYWORDS: Learning Vector Quantization, Self-Organizing Map

INTRODUCTION

Two basic assumptions of pattern recognition are compactness and separability of regions in the feature space. But many real world classification tasks suffer more or less from violations of these assumptions. One reason may be uncertainty of class membership. Other reasons one may give for are noisy processes. Automatically learning algorithms, like neural networks, are often overtaxed. A large number of neurons garantees high adaptivity to the training data set, but not necessarily to a high adaptivity to the test data set (generalization ability). For example, numerical simulations of a two-class problem in high dimensional input space, where data samples came from two Gaussian processes with identical parameters, resulted in very good reclassification rates (correct training set classifications) of 95 % and more. But classification rates (correct test set classifications) remained poor (50 %, the rate of arbitrarily classification). This has to be expected, because there is nothing to generalize in the data set.
The assumption of compact and separable regions is related to the concept of voronoi cells of nearest neighbour type classifiers, like vector based neural networks. After training a classifier should have voronoi cells containing input vectors of one class only. Otherwise, if this is not achievable, each Voronoi cell should have highest possible purity, that means the cell contains a large relative amount of input vectors with an uniform label.

METHOD

One way to improve the purity of Voronoi cells could be obtained by learning over multiple training. Before each training run the classifier is initialized randomly. After training the purity of each voronoi cell is calculated and assigned as weighting factors to their input vectors. Input vectors of cells with high purity are assigned to a high weight and input vectors of cells with low purity are assigned to a low weight. The weights for each input vector are stored for later visualization and further processing.
LVQ1 is a version of LVQ and belongs to the supervised learning; vector based neural networks [3]. With the principle of competition learning an adaptation of the prototype vectors to the distribution of the input vectors will be aimed. This paradigm is known as vector quantization. As similarity measure between input vector **x** and prototype vector **w** the euclidian distance was used. During training the prototype vector $\mathbf{w}_c$, which is closest to **x**, is updated at iteration index t by:

$$\Delta \mathbf{w}_c(t) = \pm \eta(t) \left[\mathbf{x}(t) - \mathbf{w}_c(t) \right] \qquad (1)$$

Intelligent e-Health Applications in Medicine
Special Session Proc EUNITE 2002, Albufeira Portugal Sept 19-21, 2002
Published by the University of the Aegean

If input vector **x** and winner neuron $\mathbf{w}_c$ belongs to the same class, than using the positive sign of Eq. (1); i.e. $\mathbf{w}_c$ decreases the distance to **x**. Otherwise $\mathbf{w}_c$ increases the distance to **x**. With increasing iteration index t-the step size $\eta(t)$ becomes smaller, until finishing the training by a criterion [5].

In order to avoid dead neurons and a large variance of the classification rate due to random initialization a modified version of LVQ1 [1] was applied.

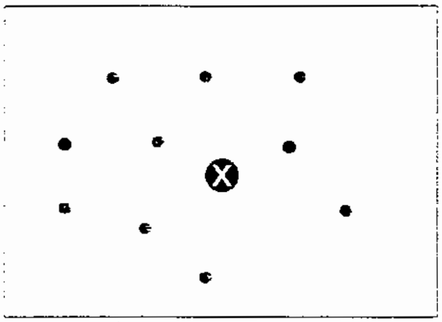


Figure 1: voronoi cell with prototype vector(cross) and 10 input vectors (small circles)

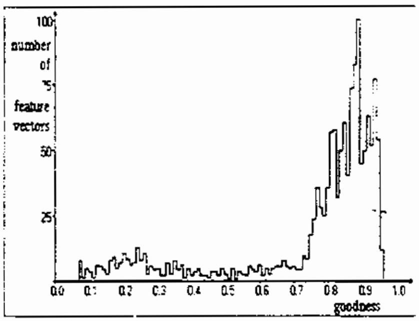


Figure2: histogram of input vectors weights from Fig. 3

After training the purity of each voronoi cells has to be calculated. In the case of two classes the purity g_i was estimated by:

$$g_i = n_i / (n_i + k_i) \qquad (2)$$

where n_i is the number of major input vectors of one class and k_i is the number of input vectors of the other class. The denominator of Eq. (2) is the size of the voronoi set. For the example (Fig. 1) one obtains: $n_i = 8$, $k_i = 2$, $g_i = 0.8$.

The g_i is assigned to each input vector of the major class in the voronoi cell of the prototype vector $\mathbf{w}_i$. The input vectors of the other class are assigned to $(1-g_i)$.

The training is repeated many times with new random initializations. The weights g_i, stored for every input vector, is averaged from training to training. Input vectors lying in compact regions of one class have a probability to reach high weights.

TWO-DIMENSIONAL EXAMPLE

Two overlapping classes, each distributed in two regions, were generated in two-dimensional space (Fig. 3). All 2,000 input vectors were applied to multiple training of LVQ1 networks with 20 neurons. The number of training runs was T = 100

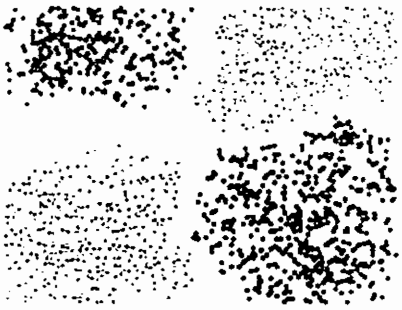


Figure3: Gaussian mixture data of 2 classes (class1: grey, class2: black)

Figure 4: Filtered data from Fig. 3

The calculated histogram (Fig. 2) shows high input vector weights for the most input vectors. The simulation shows that singular input vectors of one class in a region of the other class (e.g. black dots in the upper right and the low left light gray regions in Fig. 3) were assigned to very low weights.

Introducing an arbitrary threshold leads to Fig. 4, where about 20 percent of all input vectors having weights lower than 0.65 were blanked. Without the blanked input vectors the voronoi cells of the LVQ network reach fully purity.

MULTIDIMENSIONAL EXAMPLE

The following example comes from a motoric ability test. 21 normal subjects aged between 18 and 32 years (17males, 4 females) took part in a test to hold the balance on the left leg, while standing on a double movable base plate. The resulting two-dimensional swinging movement x(t) and y(t) was measured and digitized with a sampling rate of 90 Hz. The test was done without and under the influence of alcohol [4].

Segments of the length of 5 sec were extracted and the spectral power densities were estimated using discrete fourier transform. Further features were estimated by the empirical covariance matrix of all samples and by calculation of the entropy leading to learning set of 810 classified 36-dimensional input vectors.

The multidimensional learning set was processed in the same way by multiple LVQ1 training, but visualization is out of reach. Therefore the Self-Organizing Map (SOM) as a dimensionality reducing neural networks was applied [2]. For the calibration of the SOM the class information of input vectors were used.

The separability of two classes on the 20 x 30 map (Fig. 5) is complicate. There might be an interclass region with input vectors of both classes.

After applying the multiple LVQ1 training procedure with 100 trainings and 20 neurons the vectors were weighted and setting the threshold to 0.5, leading to blanking of 160 input vectors. The resulting map shows well separable regions in the output space (Fig. 6).

By choosing the number of neurons for the LVQ1 training one can control the number of voronoi cells and therefore the complexity of the interclass border. A high number of voronoi cells lead to a high purity and high weights of the input vectors. But the extension of the interclass region is also controlled by the weights and by the threshold.

Figure 5: 20x30 SOM for the non-filtered feature vectors of BIOSWING data

Fig. 1. 20 x 30 SOM for the filtered feature vectors of BIOSWING data

dark gray: 'with alcohol class'; light gray: 'without alcohol class'

The proposed method as a combination of filtering and visualization could be used for removing of problematic feature vectors. Unproblematic feature vectors are assigned to high weights and are unchanged for following analyses. The method can simplify classifiers to represent fewer compact regions, whose validity has to estimate separately.

REFERENCES

[1] Golz, M; Sommer, D; Lembcke, T; Kurella, B; Classification of the Prestimulus-EEG of K-complexes using Competitive Learning Networks; Proc. Eufit'98, Aachen, 1998, 1767-1771.

[2] Kohonen, T.; Self-organized formation of topologically correct feature maps; Biol. Cybernet, 43 (1982), 59-69.

[3] Kohonen, T.; The Self-Organizing Map; Proc. of the IEEE, Vol. 78, No. 9,1990, 1464-1480.

[4] Recknagel, S.; Applikation neuroinformatischer Methoden für die Klassifikation von Daten aus einem motorischen Leistungstest; University of Applied Sciences, Schmalkalden, 1999 (in german).

[5] Zell, A.; Simulation Neuronaler Netze; Addison-Wesley, 1994 (german).

NEURAL NETWORK BASED MULTISENSOR FUSION APPLIED IN REMOTE SENSING IN ECOLOGY

Dimitrios Moshou, Cedric Bravo, Stijn Wahlen, Dimitri Lemaire*, Henri Maraite*, Herman Ramon
Laboratory for Biomechatronics and Processing
Katholieke Universiteit Leuven
Kasteelpark Arenberg 30, 3001 Heverlee, Belgium
Phone: +32-16321922, Fax: +32-16328590
email: {dimitrios.moshou, herman.ramon}@agr.kuleuven.ac.be
*Unité de Phytopathologie
Faculté d'ingénierie biologique, agronomique et environmentale
Université catholique de Louvain, UCL
Place Croix-du-Sud, 2 Boîte 3
B-1348 Louvain-la-Neuve, Belgium

ABSTRACT: Excessive use of pesticides for plant disease treatment raises the danger of toxic residues on agricultural products reaching consumers. Because pesticides belong to the highest cost component in the production costs of field crops and have been identified as a major contributor to ground water contamination, their use must be reduced dramatically. This can be obtained by targeting of pesticides or those places where they are needed. One solution is to use remote sensing. The difference in spectral reflectance and fluorescence response between healthy and diseased wheat plants was investigated. Visualisation of data properties has been investigated with the use of Self-Organizing Maps. Disease detection algorithms have been developed based on the combination of SOMs with sensor fusion. Through the use of SOM, classification performance increased to more than 99%. These results are very encouraging for the development of a cost-effective optical device for recognising different kind of stresses automatically.

KEYWORDS: Neural Networks, Self-Organizing Systems, Classification, Agriculture, Plant diseases, data mining

INTRODUCTION

Excessive use of pesticides for plant disease treatment increases costs and raises the danger of toxic residue levels on agricultural products reaching consumers. Because pesticides belong to the highest cost component in the production costs of field crops and have been identified as a major contributor to ground water contamination, their use must be reduced dramatically. This can be obtained by a more precise targeting of pesticides on those places in the field where they are needed

The use of remote sensing through optical instruments provides a solution to the problem of automatic disease and stress detection. The difference in spectral reflectance between healthy and diseased wheat plants was investigated in order to use it as a means for automatic detection of plant diseases. Further, fluorescence kinetics measurements were used to gain extra information about the stress situation of the plant.

The spectral reflectance was measured by means of an imaging spectrograph [1] in order to detect different stresses as early as possible. Fluorescence Kinetics measurements were made in order to enhance stress discrimination through sensor fusion with the reflectance measurements. The fluorescence was recorded with the PEA (Plant Efficiency Analyser) fluorimeter.

Multisensor fusion was used together with different classifiers like QDA and SOM. The advantage of using fusion compared with each sensor alone is proven both in the case of the QDA and the SOM. The SOM gives superior results vs. QDA. The combination of fusion and the SOM neural network give superior results and provide an effective solution to the automated plant disease and stress detection thus facilitating environmental control of unwanted situations or water stress problems.

REMOTE SENSING - SPECTRAL INSTRUMENTATION

The equipment used during the experiments consisted of a Digiteyes visual monochromatic camera on which a V9 Specim spectrograph was mounted. The proposed weed sensor produces a reflectance spectrum of each point on a narrow linear stripe on the target surface [1]. It consists of 3 parts: an objective lens, an imaging spectrograph, and a camera. The working principle of an imaging spectrograph is shown in Fig. 2.

The objective lens (not shown in figure 2) projects the image of a field patch on the slit aperture of the spectrograph. This slit extracts a small stripe from the patch on the ground. The front doublet in the spectrograph collimates the light coming from the slit. The light is then split into its spectral components by diffraction. The second doublet forms an image of the diffracted light on a (monochrome) camera. In this way, the camera has one spatial axis and one orthogonal spectral axis. The system used was designed and made commercially available by Specim [1].

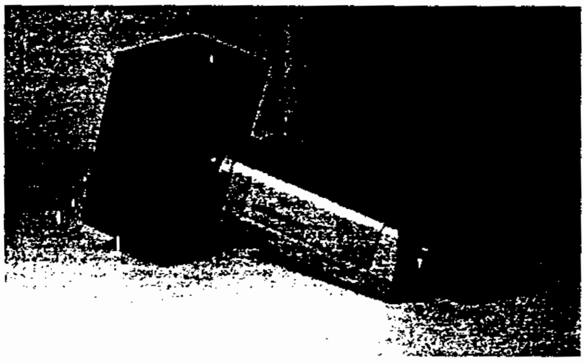


Figure 1: the Specim V9 Spectrograph mounted on a Digiteyes monochromatic camera.

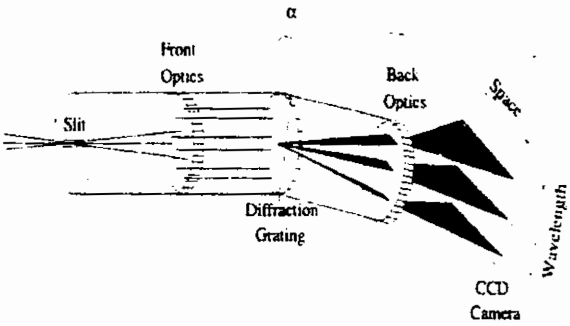


Figure 2: Working principle of spectrograph.

REMOTE SENSING - FLUORESCENCE KINETICS INSTRUMENTATION

The fluorescence was recorded with the PEA (Plant Efficiency Analyser) fluorometer of Hansatech. The fluorescence is excited by ultra-bright LEDs with a peak wavelength of 650 nm. Chlorophyll fluorescence signals are detected using a PIN photocell after passing through a long pass filter (50% transmission at 720 nm). The recording time during the experiments was 1 sec with a resolution of 10 microsec during the first 2 millisec and after that with a resolution of 1 millisec. A leaf clip was used to avoid the entrance of ambient light into the photocell.

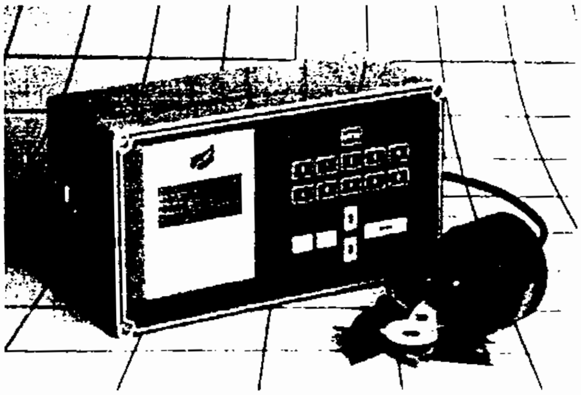


Figure 3: Illustration of the PEA Hansatech fluorometer

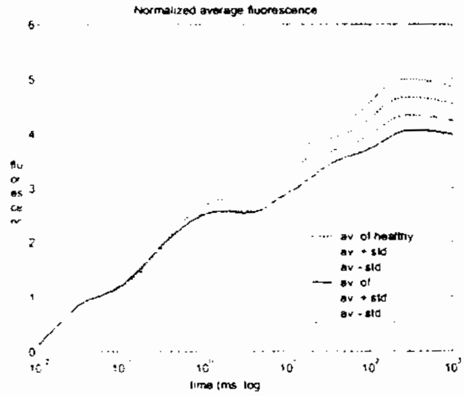


Figure 4: The fluorometer records the transient response of the chloroplasts to light excitation thus giving a quick health check

MULTISENSOR FUSION

Multi-sensor data fusion systems combine data from multiple sensors to perform inferences that may not be possible from a single sensor alone [2], [3]. Applications span military problems (target identification, threat assessment), remote sensing problems (location of mineral resources), medical diagnosis, control of complex machinery, autonomous robots [3] and automated manufacturing. Data fusion is analogous to the ongoing cognitive process used by humans to integrate data continually from their senses to make inferences about the external world.

Development issues for data fusion systems depend on the phenomena that are observed, the type of sensors utilised, and the inferences sought. These inferences in turn determine the type of techniques required. Generally applications aimed at higher-level inferences require techniques from the artificial intelligence domain such as expert systems, template matching, neural networks and fuzzy logic.

A fundamental issue is the architecture selection [2]. The issue revolves about the question of where to fuse the data in the processing flow of two or more sensors. There exist three basic multisensor fusion architectures:

- **Centralised fusion architecture**: fuses raw data from multiple sensors to determine the identity of the observed object. It is useful in the case that the sensors observe the same physical manifestation of an object (e.g. infrared and visual image).

Advantage: the minimum information loss (no feature extraction) leads to high classification accuracy.
Disadvantage: high computational requirements on the final decision calculation since feature extraction is not used.

- **Centralised fusion with feature vectors**: preprocessing is applied to each sensor reading in order to extract a feature vector. These feature vectors are concatenated in a single vector and are used as input to a classification algorithm.

Disadvantage: the feature extraction leads to information loss.
Advantage: quick computation of the final decision calculation since feature extraction is used, reduced communications requirements between sensors and fusion process. The use of feature vectors allows data from noncommensurate sensors to be fused.

- **Autonomous fusion**: In this approach the output from each sensor is a decision (i.e. declaration of identity). Smart sensors can be integrated in this kind of fusion.

Disadvantage: this approach results in a major information loss because the information processing for each sensor may result in a local optimisation rather than a globally optimal solution.
Advantage: very fast computation of the final decision calculation since partial decisions are available.

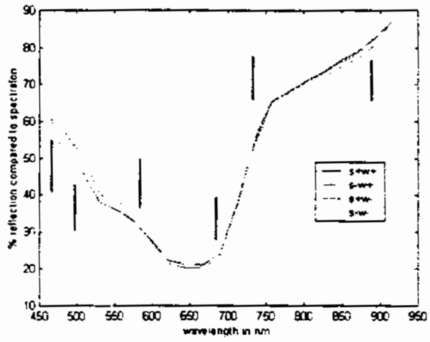
Taking into account the above considerations, the optimal architecture for disease and stress detection would be one that results into a minimum information loss but can be fast enough to be implemented in real-time. Centralised fusion with feature vectors seems the most appropriate because it satisfies both requirements of accuracy and speed.

FUSION OF REMOTE SENSING DATA

For the fusion of different sensors, the spectral results were added to fluorescent kinetic measurements. Four different treatments were produced in order to induce the different stress situations that had to be detected. The four treatments were coded and performed according to the following list.

1. s-w-: control treatment: healthy and well watered
2. s+w-: inoculated treatment with Septoria tritici, well watered
3. s-w+: healthy treatment, deficient water supply
4. s+w+: inoculated treatment, deficient water supply

In a first step fluorescent data and spectral data are investigated separately through quadratic analysis. The spectral data consisted of 6 variables (6 wavebands) found in the 'best' waveband combination suggested in the spectral report. The fluorescent data consisted out of two variables
Overview of the used variables:

Spectral reflectance data	Fluorescence data:
	
6 wavebands of 20nm wide: 463nm +/- 10nm 508nm +/- 10nm 576nm +/- 10nm 689nm +/- 10nm 735nm +/- 10nm 916nm +/- 10nm	F_0 : fluorescence at 0.05ms F_V/F_M : efficiency of the primary photochemistry

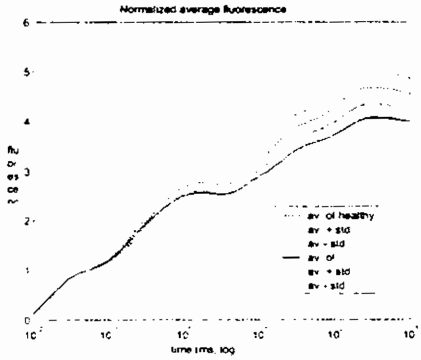


Figure 5: Position of different wavebands which separate the different stress types over the entire monitoring period during the experiment on Septoria and Water Stress.

Figure 6: average normalised fluorescence kinetics curves for healthy and diseased plants. Both treatments grouped water stressed and well watered plants.

By choosing a 'best' combination of wavebands, results are similar to ones found by using an F-test. The 'best' combination consists of two NIR, two red, one green and one blue wavebands. This combination uses most of the variation of the entire spectrum, which explains the ability of separating diseases in both early as latter disease development stages.

Further, the spectral features were put together with the fluorescence features in order to try a fusion of the two sensors.

SELF-ORGANIZING MAP

The Self-Organizing Map (SOM) [4] is a neural network (NN) that maps signals (x) from a high-dimensional space to a one- or two-dimensional discrete lattice of neuron units (s). Each neuron stores a weight (w_s). The map preserves topological relationships between inputs in a way that neighbouring inputs in the input space are mapped to neighbouring neurons in the map space. A graphical representation illustrating the functioning of the SOM is shown in figure 7.

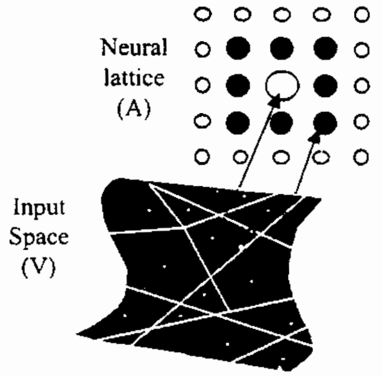


Figure 7: Graphical representation of the SOM

A way of using the SOM to find correlations between the data is to label the neurons of the SOM using a different set than the training set and finding the best-matching-units (BMUs) for every example in the testing set or labelling set. Some of the neurons that are selected most frequently by examples of one class are labelled based on a voting procedure. These neurons are then able to estimate the class of a new example presented to the SOM by calculating the Euclidean distance of the example vector to the codebook vector of each neuron and finding the BMU. The label of the BMU is then the estimated class of the new example vector. The neurons that are not labelled constitute the borders of the classes and show the degree at which example observations from one class can be misclassified as observations belonging to another class. After training , each node acts as a representation of a particular class of the data vectors that were input into the network. Nodes which are adjacent to one another represent similar, but not identical vector classes.

VISUALISATION OF DATA PROPERTIES USING SOM

The labelled SOM shows the different classified samples from the testing dataset. The labels of the different units correspond to a majority voting procedure, which gives the label according to the class attribute of the proportion of the hits. Such a labelled SOM is shown in figure 8.

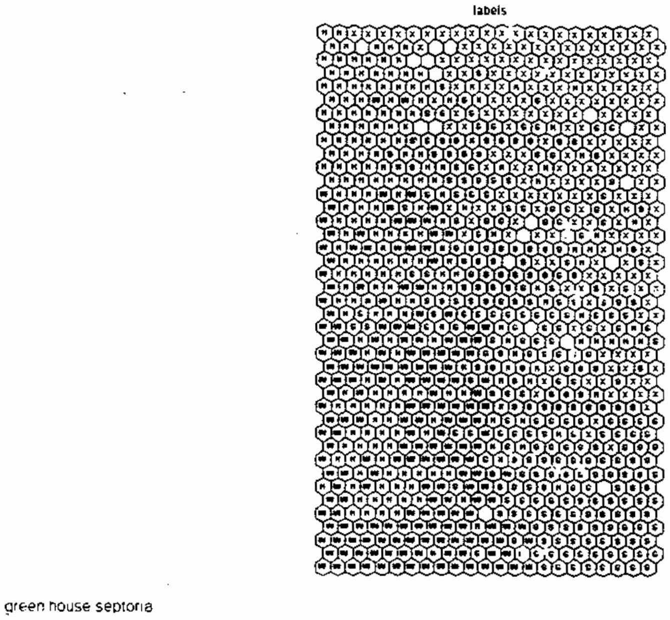


Figure 8: Labelled SOM showing class topology. Labels are showing healthy (=H), diseased (=S), water-stressed (=W) and both water deficient and diseased (=X) units related to early disease detection.

The vectors that have been stored as codebooks represent the median vectors for this neuronal unit. Each unit carries one such codebook. The way these codebooks vary shows their topological relationship and in effect the relationship between the different components of each codebook. Such a SOM with codebooks presented for each unit is shown in figure 9. The relation between the different components can be inferred from the shape of each codebook.

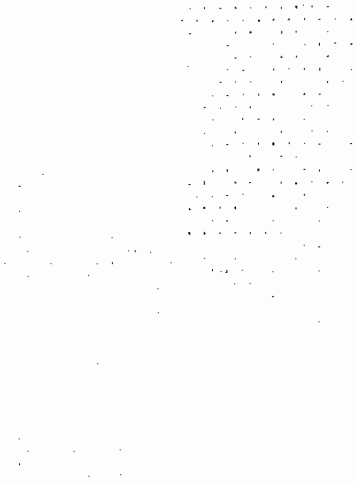


Figure 9: Feature relation for the 4 spectral reflectance features for early disease detection

Certain codebook vectors of the SOM have a much larger distance from the rest of the codebook vectors. This distortion of the SOM appears because certain of the input vectors have exhibited abnormally large values and have therefore created a codebook vector that can detect changes in the input signal when these occurr. Such behaviour can be actually visualised through the utilisation of a U-matrix. The U-matrix [5] is a matrix of distances between neighboring map units. Distance matrices essentially show the density of prototype vectors in different parts of the map. Color-coding can be used to indicate distortion in case of prototype vectors that have large distance from their neighbors thus revealing the structure of the data. Variable 3, which represents the 543 nm band carries the most novel information. The classes of stress are shown simultaneously with the U-matrix for the component plane related to the spectral feature corresponding to the 463 nm (+/- 10 nm) band. The colormap in figure 10 shows the distortion gradient to the direction of diseased plants. The structure of the distortion of the SOM gives a visualisation of the class borders determined by each SOM component plane that corresponds to a spectral feature. It can be deduced that the blue wavaband 463 nm (+/- 10 nm) provides most of the discrimination capability in separating the disease related spectra from the other combinations (water stress, healthy or water stressed and diseased simultaneously).

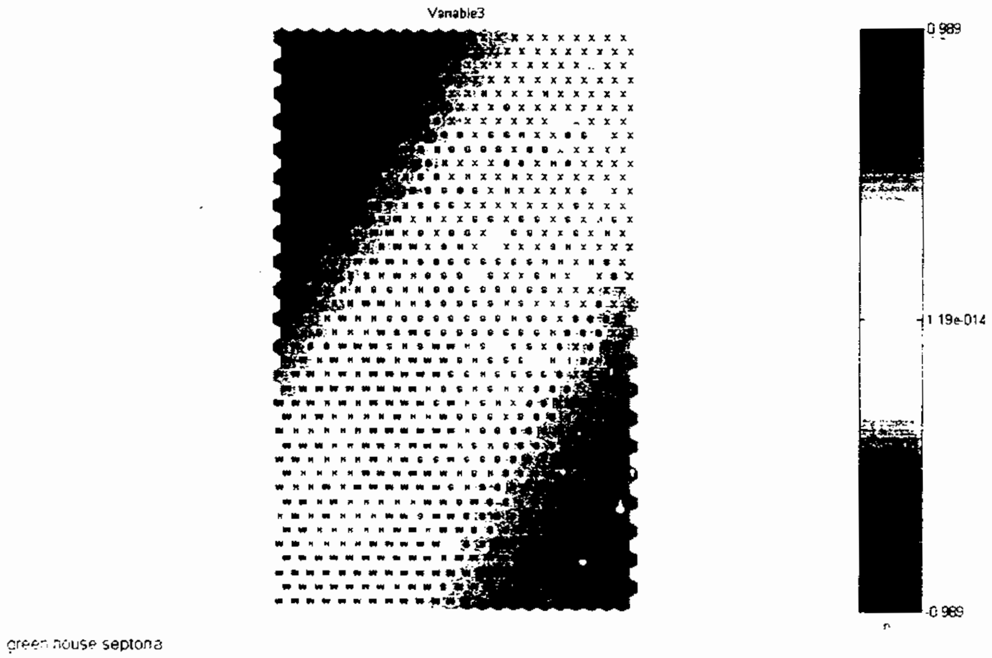


Figure 10: Labelled SOM Clusters superimposed on the variation colormap of the most discriminating spectral component (576 nm). Variable 3 is the most discriminating component because the fluorescence features were put as variables 1 and 2. Labels are showing healthy (=H), diseased (=S), water-stressed (=W) and both water deficient and diseased (=X) units related to early disease detection.

RESULTS AND ASSESMENT OF THEIR SIGNIFICANCE

First a QDA (Quadratic Discriminant Analysis) classifier has been used. A high error rate is noted when using spectral reflectance features (table 1). In contrary fluorescence features gave better results in disease detection (table 2). Adding spectral data to the fluorescence data a dataset of 8 variables is created. A SOM with 20x20 neurons was applied on the combined dataset of spectral and fluorescence measurements. Total misclassification rates dropped to 1.7% for disease discrimination (table 4).

Table 1. Performance of reflectance based QDA classifier

	Detected %		
Re		S+	S-

	s+	76.3	23.7
	s-	13.4	86.6

Table 2. Performance of fluorescence based QDA classifier

		Detected %	
Real 100		s+	s-
	s+	88.4	11.6
	s-	10.4	89.6

Table 3. Performance of fusion based QDA classifier

		Detected %	
Real 100		s+	s-
	s+	92.1	7.9
	s-	6.3	93.7

Table 4. Performance of fusion based SOM classifier (s: septoria, w: water stress, +: present, -: absent). S-w- denotes the healthy control plants.

		Detected %			
Real 100		s+w+	s-w+	s+w-	s-w-
	s+w+	99.3	0	0.7	0
	s-w+	0	97.9	1.4	0.7
	s+w-	0	2.5	96.1	1.4
	s-w-	0	0.7	1.1	98.2

CONCLUSIONS

Disease and stress infestation on plants was successfully detected through the use of remote sensing. The remote sensors were a spectrograph and a fluorometer. The developed technique used a hybrid classification scheme consisting of a multisensor fusion system and a Self-Organising Map. The SOM provides a variety of tools for visualising the correlations and significance of certain spectral variables that affect most the detection capability of the proposed intelligent disease recognition system.

Multisensor fusion was used together with different classifiers like QDA and SOM. The advantage of using fusion compared with each sensor alone is proven both in the case of the QDA and the SOM. A SOM using sensor fusion has been proven to be able of optimally recognizing different stress situations on winter wheat under ambient lighting conditions. The resulting system is able to recognise the simultaneous onset of disease and water stress from the other stress factors. The SOM gives superior results vs. QDA. The combination of fusion and the SOM neural network give very promising results and provide an effective solution to the automated detection of plant pathological sistuations thus facilitating environmental control of plant diseases or water stress problems through remote sensing and intelligent hybrid stress identification algorithms.

ACKNOWLEDGEMENT

This research has been performed in the EU OPTIDIS project. The OPTIDIS project has been funded by the EC under the Quality of Life Programme – Framework V.

REFERENCES

[1] E. Herrala, J. Okkonen, T. Hyvarinen, M. Aikio and J. Lammasniemi, 1994, "Imaging spectrometer for process industry applications", SPIE, 2248, pp. 33-40.

[2] Hall. D. L., 1992, "Mathematical Techniques in Multisensor Data Fusion", Artech House, Boston-London.

[3] Dam, J. van, Krösse B., Groen, F., 1996, "Neural network applications in sensor fusion for an autonomous mobile robot", In: Dorst, L.., Lambalgen, M. van, Voorbaak, F., editors, Reasoning with uncertainty in Robotics, Springer, pp 263-277.

[4] Kohonen. T., 1982, "Self-Organized formation of topographically correct feature maps", Biological Cybernetics 43, pp. 59-69.

[5] Ultsch. A., Siemon, H.P., 1990, "Kohonen's Self-Organizing Feature Maps for Exploratory Data Analysis". In Proceedings of International Neural Network Conference (INNC'90), Dordrecht, Netherlands, Kluwer, pp. 305-308.

Additional Method for the Evaluation of the Training Capturing: Intelligent System for the Estimation of the Human Supervisory Operator Stress Level under the Normal and Safety Critical Circumstances

Ladislava Janku*, Lenka Lhotska and Vaclav Kremen
Czech Technical University in Prague, Faculty of Electrical Engineering
The Dept. of Cybernetics,
Gerstner Laboratory for Intelligent Control and Decision Making
Technicka 2,167 29, Praha 6, Czech Republic
Phone: +420-2-2435-7306, Fax: +420-2-2435-7224
e-mail: {lhotska, kremen}@labe.felk.cvut.cz
*CertiCon,a.s, Center of Applied Cybernetics[1]
Vaclavska 13, Praha 2, Czech Republic
Phone: +420-2-214-60-237
e-mail: jankul@lab.felk.cvut.cz

ABSTRACT: This paper deals with the design and implementation of the intelligent system for the estimation of the human supervisory operator stress level from the real-time acquired physiological parameters, both under the normal and safety critical conditions, and the human supervisory operators training improvement through this system application. The system for the synchronised parallel physiological parameters acquisition and monitoring was presented in detail. This system consisted of the following components – measurement devices, client computers, server, and SQL server. Seven client modules were implemented, from which four ones are more sophisticated, namely the ECG client module, EOG client module, eye-movement module, and general signal pre-processing module. The other client modules were implemented as simple data-collection clients performing no more sophisticated signal analysis and parameterisation. The stress curves reconstruction is discussed. The pilot experiments focused on the stress level classification through the feed-forward neural network and fuzzy rule system were described.

KEYWORDS: Artificial Intelligence in Stress Classification, Human Supervisory Operator Training, Safety Assessment, Synchronised Measurement, Physiological Data Acquisition, Server-Client Architecture, Signal Processing, Fuzzy Expert System

INTRODUCTION

From the viewpoint of the tremendous growth in the area of a control of large technological processes, the method enabling a real-time monitoring of a human supervisory operator's psychical stress level appears as a suitable approach promising control process fault minimization through the elimination of faults caused by the human factor. Particular attention should be given to the application of the stress/fatigue level estimation techniques, applied both to the supervisory operators training process and to the real supervising of plant and control processes.

During the recent years, a wide spectrum of simulations of human behaviour and cognition, covering the whole area from artificial intelligence and fuzzy logic to control theoretic techniques, appeared on the scene [1][8][9][10]. However, there is still great deal the can be done in the area of a real-time monitoring of a human operator's psychical stress and/or fatigue levels and integration of such estimators in simulation systems. An experimental Yoshikawa and Takahashi's study [2] [3] focused on a research in the line of human's cognitive features, appears as a fundamental work in this field. A concept of the two-way adaptive human machine interface was presented. Basic concept was later integrated in Mutual Adaptive Interface [4][5][6].

In recent years, several research projects were realized in the Gerstner Laboratory, focused on the evaluation of dependencies between the supervisory human operator psychical state (stress/fatigue) and the non-invasive measured physiological parameters, and on the application of AI to the problem of the stress/fatigue levels estimation and/or their dynamics description and prediction. Our previous research activities in this area are described in [11][12][13].

---

[1] This work was supported by the Ministry of Education of the Czech Republic under Project LN00B096.

Intelligent e-Health Applications in Medicine
Special Session Proc. EUNITE 2002, Albufeira, Portugal Sept 19-21, 2002
Published by the University of the Aegean

This paper deals with the semi-distributed system for the real-time monitoring of the physiological parameters and their evaluation through the selected AI methods. The application of this system enables in essential in the area of training to compare and evaluate the physiological capabilities of the human supervisory operators, evaluate their strengths and weaknesses. It opens the door to the area of self-improvement and enables to discover situations in which the supervisory operators reach their boundaries and they are not able to explore their knowledge in an optimal way because of stress. The presented semi-distributed system architecture takes in account estimated stress and/or fatigue levels of real power plant human operators. A vast amount of physiological data collected during the simulated experiments gives a satisfactory baseline for estimation procedures. Intelligent techniques have been applied to perform collected data analysis. All these techniques contain knowledge on relationships among physiological parameters (heart frequency, respiratory frequency, skin galvanic reaction, systolic and diastolic blood pressures, muscular tonus, etc.) and levels of stress or fatigue. Because of impossibility to connect all the measurement equipment to one computer or to synchronise collected data sample per sample, a semi-distributed system for the data collection, based on the client-server paradigm, has been developed.

MEASUREMENT OF PHYSIOLOGICAL PARAMETERS – HUMAN SUPERVISORY OPERATORS TRAINING SUPPORT

The design, management and evolution of the educational plans are carried out by the specialised certified training centers in accordance to requirements set by the particular power plants. The basic human supervisory operator training consists of lectures (theory) and practical seminars. Involvement of the human supervisory operators in the training on power plant simulators has two main effects, namely they acquire practical skills necessary for their future work on a real power plant and the training enables to discover such of them, who are not able to explore knowledge acquired on the theoretical lectures in a praxis.

The training on the power plant simulators covers both the normal and safety critical condition examples. Each example is repeated several times to improve acquired experience. In an ideal case, the trained example should be repeated as long as it is clear than the human supervisory operator is fully trained for it. Normally, it is hard to decide if the human supervisory operator is fully trained or not, because of many factors. Particularly, many problems reflect basic psychological principles that the natural capabilities of the human supervisory operators differ in many ways. Each of the human operators needs different number of repeating of the trained example to be fully trained for it. Another problem is that some examples can be very stressful for the particular supervisory operators and there is no mechanism how to discover them.

Real-time measurement of the physiological parameters followed by the real-time stress level estimation can support the training process in two ways:

- *To set a stress curve* for each human supervisory operator for particular examples. This stress curve enables to discover more stressful parts of the trained example, to identify them and modify the training process to eliminate these weaknesses.
- *To define a capturing of the trained example* for each human supervisory operator and for each trained example. This capturing of the trained example consisted of three factors – correctness, time and stress level. Human supervisory operators have to find a correct solution of the trained example in some admitted time (usually seconds). The last factor characterizes a stress level set to the human supervisory operator, while he solves the trained example. If the stress level is high, the particular example must be trained again. The *capturing of the trained example* enables to find an optimal number of repetitions of the training process of each particular example for each particular human supervisory operator.

Involvement of both these stress characteristic into an educational process of the human supervisory operators conduces for three main effects. Firstly, they improve their system knowledge and acquire practical skills in the scope of their individual capabilities and abilities. Secondly, this method enables to determine the most stressful circumstances and move the whole educational process in the direction of stress suppression; it can be cut down through the repetition. Finally, this approach is the best from the financial point of view, namely the training costs are reduced through the training individualization.

SYSTEM FOR THE SYNCHRONISED PARALLEL PHYSIOLOGICAL PARAMETERS ACQUISITION AND MONITORING

The system architecture was designed on assumption that the measurement of the physiological parameters and their pre-processing can be executed independently, e.g. there is no dependency among them which would required a capturing of them into one module. On this assumption, the system for the synchronised parallel monitoring of the physiological parameters could be designed on the basis of the client-server architecture. The basic system architecture is shown on the Figure 1.

BASIC SYSTEM ARCHITECTURE

The system for the synchronised parallel monitoring of the physiological parameters consisted of the following components – measurement devices, client computers, server, and SQL server (database machine). The clients - data collection modules - collect data obtained from different measurement equipment. After the signal segmentation and time-stamps assignment to these segments, the simple or more sophisticated real-time pre-processing techniques are applied. Characteristic features are computed for each segment and, accompanied by the appropriate time-stamp, are sent to server. Simultaneously, the collected signal segments with time stamps are stored in the local databases. The way in which the data are processed at the level of clients - data collection modules - depends at any given time on the operations supported by the procedures implemented in these modules; they can be renovated in a very simple way. This semi-distributed approach to data collection and processing enables to make the best account of computational resources; it also reduces an amount of data transferred to the server. Server is also responsible for starting and stopping clients at any time and for the client synchronization.

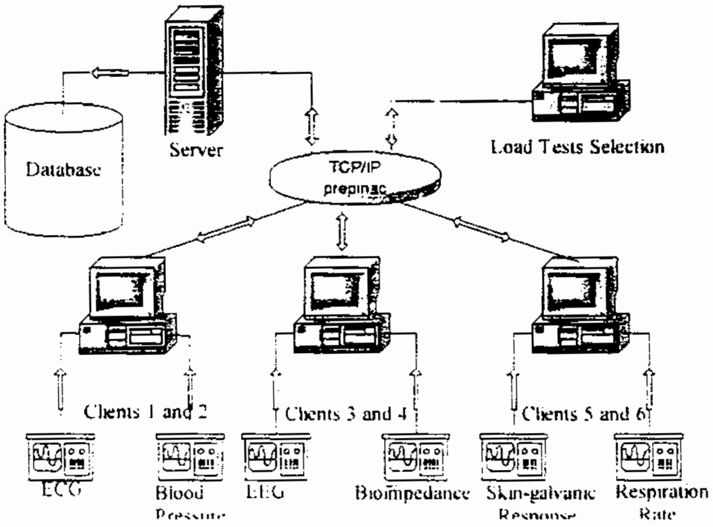


Figure 1: Semi-distributed system architecture for data collection and analysis

GENERAL FUNCTIONALITY

This part refers to the two basic system components – server and client. Server is controlled through the user interface designed with the aim to enable user to control whole measurement process, e.g. server and client modules. Both the application server and application client features are listed bellow.

Application server features:

- Communication and data collection through the TCP/IP network
- A capability to evaluate received parameters and select appropriate action in a time interval shorter than 0.5 s
- A capability to display results of evaluation in a required form
- Biofeedback
- A simple way on modification/ addition of procedures for parameters evaluation
- Extensibility
- Standard communication interface
- A capability to fully explore the mathematical and statistical libraries

- Utilization of MS Windows or special handlers for control of PC peripheries
- Utilization of Hardware Abstraction Level, OS services, COM and ActiveX technologies
- Connection to Database

Application client features:

- Real-time communication with the measurement equipment/ measurement software
- A capability for real-time pre-processing of the measured data
- A capability to send the results of the measured data pre-processing to the server
- Standard communication interface
- A capability to fully explore signal processing libraries
- Utilization of MS Windows or special handlers for control of PC peripheries

Each client module consisted of three layers – communication interface, a layer of measurement device/software control and communication, and signal/data processing layer.

CLIENT MODULES

Until now, seven client modules were implemented, from which four ones are more sophisticated, namely the ECG client module, EOG client module, eye-movement module, and general signal pre-processing module. The other client modules, namely diastolic and systolic blood pressures client module, skin galvanic response (SGR) client module, and respiration rate module, were implemented as simple data-collection clients performing no more sophisticated signal analysis and parameterisation. In the future, adding EEG and Bioimpedance client modules will extend the family of the system components. The configuration of client modules is made in accordance to the configuration file saved in the storage at the server at the beginning of the measurement. Table shows the basic functionality of the client modules.

Table I:. Client Modules Features Overview

Client module ID	Common features							Special features
	FIR filters	IIR filters	FFT	Cepstral Analysis	Wavelet Analysis	Parameterisation	Energy	
ECG	x	x	x		x	x	x	QRS complex detection RR interval detection QT interval detection Band spectral analysis Band time frequency analysis
EOG	x	x	x				x	Events detection HMM
Eye-movement	x	x			x			Image processing Iris detection Iris movement detection
Common	x	x	x	x		x	x	-
Blood pressures	x	x					x	-
SGR	x	x						-
Respiration rate	x	x						-

FUNCTIONALITY AND TIME COMPLEXITY VERIFICATION

This paragraph refers to the part of the development process devoted to the system functionality and time complexity verification. Two important verification tests were selected for demonstration in this paper

- *Communication functions and procedures time complexity.*
- *A speed of reading server variables by the clients.* The requirement of client modules synchronisation can be regarded as the one of the most important requirements set to the system for synchronised parallel monitoring of the physiological parameters. These verification procedures test the application functions and procedures responsible for the client synchronisation with aim to identify the largest time delay between pairs of the client modules. This largest time interval defines synchronisation accuracy.

Verification test I: Communication functions and procedures time complexity

This verification test mapped the effectiveness of the application functions and procedures. The MATLAB Profiler – a standard part of the MATLAB software - was utilised for the functions and procedures time complexity verification. This tool disposes of capabilities to map time complexity of all in MATLAB written functions and explore them at all levels. The results of this verification test were continuously used for the code improvement.

Verification test II: A Speed of reading server variables by the clients

In the framework of the operating system (OS) from the family of MS Windows OS, it is impossible to measure the run-time of a part of a program code. This feature is determined through the Windows OS architecture and cannot be eliminated. While the application runs, there is any Windows OS service, which has a higher priority than the other services and applications. However, only the mean value of the run-time of any function/ apart of program code can be determined. In the system for synchronised parallel monitoring of the physiological parameters, the speed of reading server variables by the clients is derived from the run-time of the function *getval*. The verification test was performed under the following hardware and software configuration.

- 1 to 3 notebooks HP Omnibook 4150, Intel Pentium II/450 MHz, 128 MB RAM
- OS: MS Windows 2000 Professional
- Notebooks were connected together in the frame of one segment of an Ethernet network
- HP ProCurve Switch (10 Mb/s)
- Notebooks network cards configuration: half duplex mode, speed 10 Mb/s

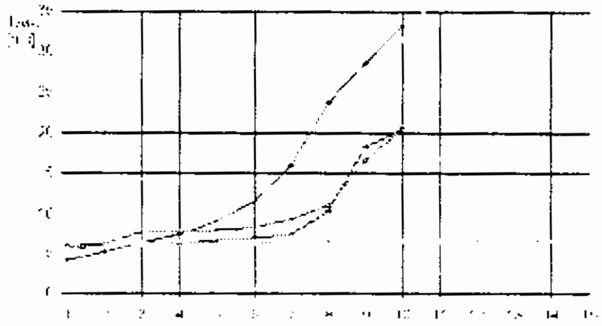


Figure 2: Mean value of the run-time of the getval function. Legend: y-axis: mean value of the run-time, x-axis: the number of the running clients. More information can be found in the text.

During the verification test, the dependency of the run-time of getval function was observed with regards to the number of running client modules. Figure 2 shows results of the verification test. The distribution of the server and clients to the 3 computers was the following: red color – server and all clients were run on one PC, violet – server was run on the first PC and clients were run on the second PC, green color – server and several clients were run on the first PC and the other clients on the second PC, light blue color – server was run on the first PC, clients were run on the second and third PC's.

IMPLEMENTATION

This part refers to some remarks and technical details of the implementation of the system for the synchronised parallel physiological data acquisition and monitoring. The system architecture was designed for the operating systems of the MS Windows family in accordance to the norms 457 291 483 306 Microsoft. Following OS are supported: MS Windows 2000 professional, MS Windows XP professional Edition, MS Windows 98 and MS Windows 95. The core of the system was implemented in MATLAB 5.3. The client modules were implemented in MATLAB 5.3 extended with the Signal and Statistic Toolboxes. The core of the system, and server and client functions explore the capabilities of the OS services, COM and ActiveX techniques.

REAL-TIME ESTIMATION OF STRESS CURVES AND STRESS LEVEL

From the theoretical viewpoint, the techniques of real time estimation of stress curves and stress levels can be considered as pattern reconstruction/ pattern classification methods. All these techniques contain knowledge on relationships among physiological parameters and levels of stress. The stress curve reconstruction method is based on mapping between the n-dimensional space of the physiological parameters and the space of values of the stress intensity function. In both cases, the independent variable is time correlated with the phases of a trained example, if we turn back to the problems of human supervisory operator training. In the other hand, the stress level estimation is a pattern classification method, which requires the existence of the stress intensity scale and the existence of mapping from the n-dimensional space of the physiological parameters (a space of patterns) to the space of classes. Next text describes both approaches, but it is necessary to say, that just only the pilot studies and systems were done.

STRESS CURVES RECONSTRUCTION

Stress Curve is a graphical representation of a Stress Intensity Function S(t): t → R. *Stress Intensity Function S(t)* is a mapping from t to R, which characterizes a level of stress set to the human operator in time. Let the $M \subseteq R^n$ be a set of input vectors of the measured physiological parameters and $H_i \subseteq R$ a set of all possible values of the stress intensity function $S_i(t)$. Then, the mapping from φ: $M \rightarrow H$ assigns a value h to each vector m of input pattern set M.
Generally speaking, this mapping is not time-independent, but for the computational reasons, we assume that it is unique and time-independent for each human supervisory operator in a short-time period. This mapping is influenced by many factors, e.g. medicament taking, alcohol drinking, age, fatigue, and all these factors must be taken into account. In fact, this mapping φ is approximated through the mapping φ' which is estimated through the soft-computing based system. We decided to utilize a fuzzy rule system, which enables formulate known relationships among physiological parameters, special factors and stress intensity function values explicitly. This knowledge was obtained partially from the consultation with experts, and partially from the training data acquired during the preliminary tests. More rigorous approach would be based on invasive measurements – blood analysis.

STRESS LEVEL CLASSIFICATION

While the task of the stress curve reconstruction was focused on the approximation of the mapping φ: $M \rightarrow H$, the stress level classification process classifies each input vector from $M \subseteq R^n$ to one of the three classes. Two computational intelligence approaches were applied to estimate the psychical stress level – fuzzy rule system & feed-forward neural net. Both classification approaches performed classification to the categories described by the same three-value scale – low psychical stress, medium psychical stress, high psychical stress. Fuzzy rule system consisted of 27 fuzzy rules defined for 3 inputs & 3 outputs (classification to 3 classes). Applied neural net classifier consisted in 3-layer feed forward neural net (3 neurons in input layer, 7 neurons in hidden layer & 3 neurons in output layer). Well-known back-propagation algorithm was used for this neural net learning. Two types of neuron output functions were used – linear & sigmoid functions. A set of simple functions describing the relations among the psychical stress levels and simulated physiological parameters were created. These functions were derived from the real physiological data histories. Training data was generated using these functions, and it consisted of a matrix of 27 input vectors (blood pressures, heart rate, respiratory rate). Results obtained by both classifiers were very similar. Obtained results imply competency and correctness of this approach to actual psychical stress level estimation problem.

CONCLUSION

From the viewpoint of safety assessment in the area of supervision of large technological processes, the human-machine systems performing real-time human supervisory operator's psychical state analysis as suitable systems for the safety risks minimization and the training improvement.
The additional method for the evaluation of the training capturing was presented. The objectives of this project cover the design and implementation of the intelligent system for the estimation of the human supervisory operator stress level from the real-time acquired physiological parameters, both under the normal and safety critical conditions. The system for the synchronised parallel physiological parameters acquisition and monitoring was presented in detail and its implementation in MATLAB 5.3 under the family of the MS Windows operating systems was presented. This system consisted of the following components – measurement devices, client computers, server, and SQL server (database machine). Seven client modules were implemented, from which four ones are more sophisticated, namely the ECG client module, EOG client module, eye-movement module, and general signal pre-processing module. The other client

modules, namely diastolic and systolic blood pressures client module, skin galvanic response client module, and respiration rate module, were implemented as simple data-collection clients performing no more sophisticated signal analysis and parameterisation.

Real-time measurement of the physiological parameters followed by the real-time stress level estimation can support the training process in two ways – a stress curve reconstruction for each human supervisory operator for particular examples and a definition of the capturing of the trained example for each human supervisory operator and for each trained example. Also stress level classification can be useful, particularly in the examples, where the accuracy and resolution of the stress curve are not necessary. Stress curve reconstruction and stress level classification processes were described. Pilot experiments in this area were based on application of fuzzy rule system and feed-forward neural network. Both these methods were designed and test on the limited training set of the measured physiological data.

ACKNOWLEDGEMENTS

Research work of Ladislava Janku was supported by the Ministry of Education of Czech Republic under Project LN00B096. Research work of Lenka Lhotska and Vaclav Kremen was funded by the Ministry of Education of Czech Republic under the research program 'Transdisciplinary Biomedical Engineering Research', No. MSM 210000012.

REFERENCES

[1] Cacciabue, P.C.: Modelling and Simulation of Human Behaviour in System Control, Springer Verlag, 1998

[2] Yoshikawa,H., Kitamura,M., Takahashi,M.: Development of Biocybernetic Analysis System (BIOCYS) for MMI - Related Cognitive Factor Analysis, Institute of Atomic Energy, Kyoto University Press, Japan, 1985

[3] Yoshikawa,H., Takahashi,M., Arai,T.: Concept on Mutual Adaptive Interface and Related Experimental Study, Institute of Atomic Energy, Kyoto University Press, Japan, 1994

[4] Takahashi,M., Kitamura, M., Yoshikawa,H.: Development of Real-Time Cognitive State Estimator, Institute of Atomic Energy, Kyoto University Press, Japan, 1995

[5] Takahashi,M., Kubo,O., Yoshikawa,H.: Mutual Adaptive Interface: Basic Concept, Institute of Atomic Energy, Kyoto University Press, Japan, 1993

[6] Kubo,O., Takahashi,M., Yoshikawa,H.: Mutual Adaptive Interface: Laboratory Experiments for Human State Estimation, Institute of Atomic Energy, Kyoto University Press, Japan 1993

[7] Fujita, Y. et al.: Designing a Knowledge Based Operator Support System for Practical Applications, Nuclear Technology 95, 1991

[8] Bar, A., Cohen, P.R., Feigenbaum, E.A: The Handbook of Artificial Intelligence – Volume IV., Addison-Wesley, Rading, MA, U.S.A, 1989

[9] Amalberti, R., Deblon, F.:Cognitive Modelling of fighter Aircraft Process Control; A Steps Towards an Intelligent Onboard Assistance System, International Journal of Man-Machine Studies, 36, 1992

[10] Corker, K.M., Smith, B.: An Architecture and Modelling for Cognitive Engineering Simulation Analysis: Application to Advanced Aviation Analysis, AAIA Conference on Computing in Aerospace, San Diego, CA, U.S.A. 1993

[11] Sorf, M. - Eck, V. - Janku, L. - Lhotska, L: Application of Neural Network for Stress Classification. In: The State of the Art in Computational Intelligence. Heidelberg : Physica-verlag, Germany, 2000

[12] Eck, V., Janku, L., Sorf, M., Novak, D., Fejtova, M., Roknic, J.: Applications of the Fuzzy Rule System and Neural Net to the Human Operator Stress Level Estimation. In: Proceedings. Zittau: Institut für Prozesstechnik und Messtechnik, Germany, 2001

[13] Janku, L., Sorf, M., Lhotska, L., Eck, V.: Five Approaches to Computational Personality Classification Problem. In: Kuncheva (Ed.), Proceedings of CIMA 2001, ICSC/NAISO conference, UK, 2001

Computational Intelligence - Design Tool for Use in Biotechnological Process Monitoring

Klaus-Dietrich Kramer, Steffen Patzwahl, Thomas Nacke*, Dieter Frense*
Hochschule Harz - University of Applied Sciences
Department of Automation and Computer Science
Friedrichstraße 57-59, D-38855 Wernigerode
Phone: ++49(0)3943 / 659-317, Fax: ++49(0)3943 / 659-107
email: kkramer@hs-harz.de, spatzwahl@hs-harz.de
* Institute for Bioprocessing and Analytical Measurement Techniques e.V. (iba)
Rosenhof, D-37308 Heiligenstadt
Phone: ++49(0)3606 / 671-163, Fax: ++49(0)3606 / 671-200
email: thomas.nacke@iba-heiligenstadt.de

ABSTRACT: This paper describes a universal software tool for biotechnological process monitoring. This tool combines classical model building and process control methods with computational intelligence techniques in one shared user interface. The particular role of classificators is taken into account by implementing different classification algorithms. In addition, modelling is based directly on plant online sensor signals. Thus, also data mining, especially in the form of signal conversion and feature extraction, is considered in the tool. The application presented is a fuzzy control model to control substrate supply to a biogas pilot plant developed by means of this software tool.

KEYWORDS: computational intelligence, fuzzy control, neural network, feature extraction, data mining, bioprocess

INTRODUCTION

In the past few years computational intelligence (CI) procedures – besides classical methods or to complement such – have gained increasing significance in modelling applications. Fuzzy control (FC) systems or applications of artificial neural networks are the state-of-the-art. CI methods have become especially indispensable in finding solutions in bioprocess engineering-related process classification, modelling or control, due to grossly non-linear interrelationships of process variables often in connection with no or hardly existing process expertise. Some applications are described in [5],[7],[10],[12],[13],[16] and [17] below.

Prerequisite for the implementation of monitoring and/or control system for a bioprocess requires is a basic process analysis ([2],[20]). This paper assumes that the bioprocess to be described is executed in a plant equipped with online sensory analysis of process variables so that online process data records are available. Process analysis is tasked with evaluating the series of statistical values of process data, if need be, in connection with process expertise. Central task is the development of an appropriate process model. The initially essential aspect is the type of modelling. Depending on the target (model with an analog output quantity or classificator) there are quite a number of model building options. Classical modelling methods use, e.g., differential equations or statistic classificators, while CI-based models employ artificial neural networks and/or fuzzy logic. Frequently it is especially a combination of classical and CI-based models that yields the desired solution for a process model or a control strategy.

The works covered in this paper were aimed at assisting users in the selection of a suitable modelling type. The developed tool „Model Builder" integrates several modelling methods and makes such available to the user as a graphic development interface. Each model can be parameterised and put to a data-based testing as required. Easy and swift model building and simulative verification enable the user to realise very quickly different modelling approach to a given task. In particular the easy comparability of the developed models considerably facilitates finding of an optimal modelling type. The implementation of controller structures is also taken into consideration

A further essential aspect of model building is the proper selection of model input variables. In many cases, also especially in bioprocesses, optimal model input variables can only be identified via a preceding feature formation. This is why also simple methods for feature extraction from time data series have been implemented into this software tool.

Intelligent e-Health Applications in Medicine
Special Session Proc EUNITE 2002 Albufeira Portugal Sept 19-21, 2002
Published by the University of the Aegean

Data mining is completed by the sub-processes „Conversion of sensory voltage signals“ and „Scaling“ as data pre-processing and processing functions.

THE SOFTWARE TOOL Model Builder

BASIC CONCEPT

The Model Builder is a software tool for MS Windows^{TM} operating systems. It provides the user with the possibility to interactive development, parameterisation and data-based testing of system models, including preceding data mining functions. Figure 1 shows the main functional units of the Model Builder. All data mining steps as well as different model-building algorithms are displayed for the user block-oriented via a graphic user interface.

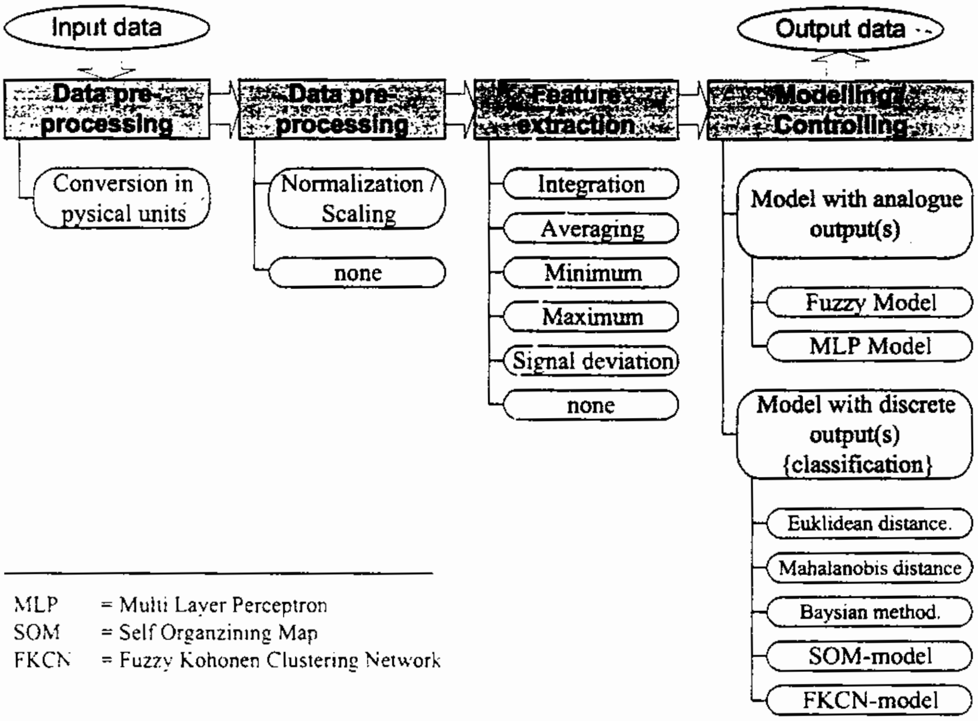


Figure 1: Main functional units (modules) of the Model Builder

DATA MINING

The first data mining step is the preparation of data. Under consideration of its practical applicability the Model Builder concept assumes that direct sensor output signals from the plant are used as input signals for the model to be developed. Typically, such sensor output signals are voltage values representing process signal values. A conversion rule for each input variable can be deposited in the module „data preprocessing“ which enables the calculation of physical units of process variables proceeding from sensor-provided voltage values.
The module „Data processing“ is available as the second data mining step and can be used to scale input signals. Typically, all input signals are adjusted to the 0 - 1 range. This is especially necessary when artificial neural networks are used because the neuron activation function is only defined in this range.
The last step of data mining is provided in the Model Builder by the module „Feature extraction“, where the user can use the functions „Integration“, „Averaging“, „Minimum value determination“, „Maximum value determination“ and „Signal deviation determination“. The feature extraction is aimed at the forming of features related to series of statistical input values. For one series of statistical values with 144 measured values (scanning rate: 10 min, observation period: 24 h), e.g., the arithmetical mean can be input as a feature describing a series of statistical values. The aims of

feature extraction are data reduction and above all filtering of significant signal features to facilitate later modelling [19] and [22].

MODEL FORMING / CONTROLLER DESIGN

The block „model building / control" provides several model-building methods using both classical and CI-based algorithms. To start with the individual model-building methods are distinguished in terms of model output variables into system models with analog model output and system models with discrete model output (classificators (confer Figure 1)). Implementation of numerous classificatory-based model building algorithms facilitates easy process monitoring, e.g. by classifying certain process sensor signals into feature groups („process in underload condition", „process in optimum condition", „process in overload condition", „error"). The model-building methods, in particular their implementation and parameterisation, are explained below in greater detail.

Fuzzy Model

A fuzzy-model ([8],[21]) can be used for model building or process control in cases where the process to be modelled cannot be described mathematically by means of differential equations, but where experts can provide a qualitative or verbal description of the relations between process influencing variables. The Model Builder provides a complete development environment for fuzzy models and fuzzy controllers.

MLP Model

If a mathematical process description is not possible and in the absence of process expertise, artificial neural networks are one opportunity to still obtain a process model. However, precondition for the application of artificial neural networks is the existence of process-describing data records. The Model Builder includes a Multi-Layer Perceptron (MLP) with a backprop learning algorithm ([15], [22]) which enables it to "learn" relations between inputs and outputs of a training data record file and apply such to new, not learned input data in the recall phase. The MLP can be employed as a process controller [10].

Statistical classificators (Euclid, Mahalanobis, Bayes)

Statistical classificators sort certain value pairs of model inputs (objects) on the basis of distance and/or truth dimensions into one or more predefined classes (model output) ([3],[8],[11],[14]). This classification is one form of model building. The classification results make process monitoring or automatic process intervention possible.
The common feature of all statistical classificators is that they operate according to the following algorithm:

1. Determination of similarities / distances / affiliations / truth values of the kth situation to all K classes
2. Determination of similarity ranks of the classes
3. Selection of the class with maximum similarity or minimum distance

$$c = \arg\min_{k=1}^{K} d_k(\mathbf{x}) \qquad (1)$$

The Model Builder includes distance and similarity dimensions after Euclid (Formula 2), after Mahalanobis (Formula 3) and after Bayes (Formula 4) ([3],[14],[11],[22])

$$d_k(\mathbf{x}) = \|\mathbf{x}-\boldsymbol{\mu}_k\| = \sqrt{\sum_{i=1}^{n}(x_i-\mu_{ki})^2} \qquad (2)$$

$$d_k(\mathbf{x}) = (\mathbf{x}-\boldsymbol{\mu}_k)^T \cdot \mathbf{C}_k^{-1} \cdot (\mathbf{x}-\boldsymbol{\mu}_k) \qquad (3)$$

$$d_k(\mathbf{x}) = -\log|\mathbf{C}_k| - (\mathbf{x}^T-\boldsymbol{\mu}_k^T) \cdot \mathbf{C}_k^{-1}(\mathbf{x}-\boldsymbol{\mu}_k) \qquad (4)$$

As required by the target of the classification method and according to the properties of input data records the user can define different classificators and verify their results with test data records.

SOM Model

Self-organising artificial neural networks (self-organising maps = SOM) are capable of automatically recognising neighbourhood properties in input data patterns. Hence, SOM are ideally suited for data classification. Provided that there is a suitable network algorithm and a matching weight matrix allocation, a certain coherent neighbourhood area for each input data pattern is activated in the output configuration while the other properties remain inactive. Unsupervised learning in the training phase results in self-organisation of weights as a function of frequency distribution of the provided input patterns and their neighbourhood properties ([12],[15]). In order to be usable as a classificator each neuron has to be assigned a meaning in a labelling phase. The SOM algorithm in the Model Builder combines the training and labelling phases into one step and, in a further step, can apply the trained SOM to new data.

FKCN Model

The Fuzzy Kohonen Clustering Network (FKCN) [1] is a neuro-fuzzy model where fuzzy logic and artificial neural networks ceased to exist independently side by side but blend into each other in the form of an overall model [22]. Especially self-organising artificial neural networks [9] and the fuzzy c-means algorithm ([1],[22]) were combined. The main objective of the symbiosis of fuzzy logic and artificial neural networks is to combine the advantages and compensate the shortcomings of the individual methods.

All above-mentioned model-building methods can also be used for process controller development. To this end the process model has to be designed by using an actuating variable as model output.

APPLICATION IN BIOGAS PLANTS (Applicative Example)

BIOGAS GENERATION AND CONTROL ENGINEERING PROBLEMS

Biogas generation from organic residues is a complex process, especially in anaerobic co-fermentation, and proceeds via a number of intermediate products. The substrate composition is not constant and is subjected to substance-related and seasonal variations. Economic use of biogas plants for energy generation requires a gas yield as continually as possible at an optimal carbon degradation. Meeting of requirement makes it necessary to control the quantity of fed substrate as a function of the desired gas yield volume. Conventional control systems meet their technical limits equipment when having to control such anaerobic fermentation processes for biogas generation. The difficulty lies in the non-linear relations between substrate supply and biogas extraction which are linked to large time constants. In addition, those control engineering problems in anaerobic fermentation are aggravated by order-related co-fermentation of different organic wastes. The latter entails differing space loads of the fermentation plant. The anaerobic fermentation process is relatively complex and very difficult to describe in mathematical and model terms due to the non-linearity between substrate feeding, biomass growth and gas yield [13].

FUZZY PROCESS-CONTROL MODEL

A process-description database was generated as a basis to build this model building. The lab-scale biogas plant (fermenter volume 6 l) with extensive online measured data acquisition is shown in Figure 2.

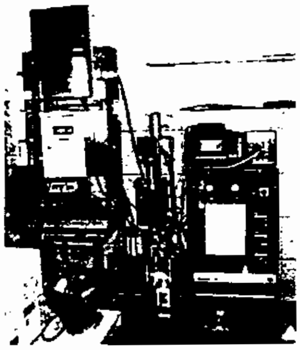


Figure 2: Picture of the lab-scale biogas pilot plant (iba – Heiligenstadt)

Following initial data evaluation with correlation analyses the below measured data were assessed to be process- and model-relevant.

- Biogas volume (BGV) [dm^3]
- CH_4 concentration in the generated gas (CH4) [%]
- CO_2 concentration in the generated gas (CO2) [%]
- O_2 concentration in the generated gas (O2) [%]
- pH (pH) [-]

The fuzzy model [17] developed by means of the Model Builder is shown in Figure 3. This model operates with seven input variables that are generated via the data-mining feature extraction function.

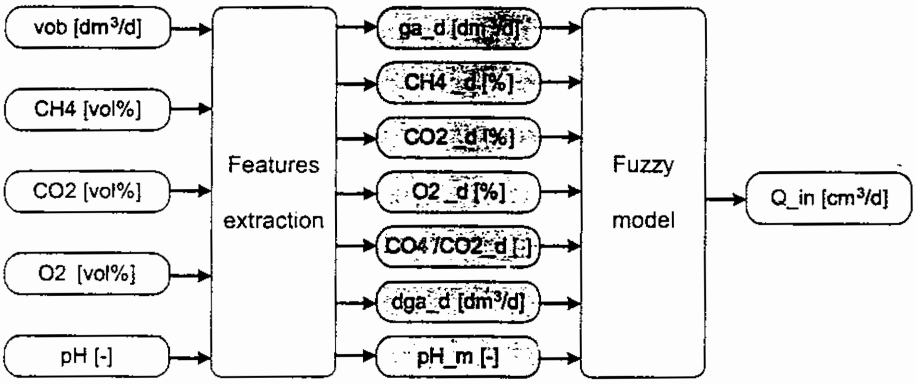


Figure 3: Structure of the fuzzy models with preceding feature extraction [17]

The model input variables were the following process-descriptive features:

- Gas quantity per day (GM) [dm^3/d]
- CO_4 content per day (CH4_d) [%]
- CO_2 content per day (CO2_d) [%]
- O_2 content per day (O2_d) [%]
- Quotient of CH_4 and CO_2 content per day (CH4/CO2_d) [-]
- Variation of gas volume between current and preceding day (dGM) [dm^3/d]
- Arithmetic mean of pH (pH_m) [-]

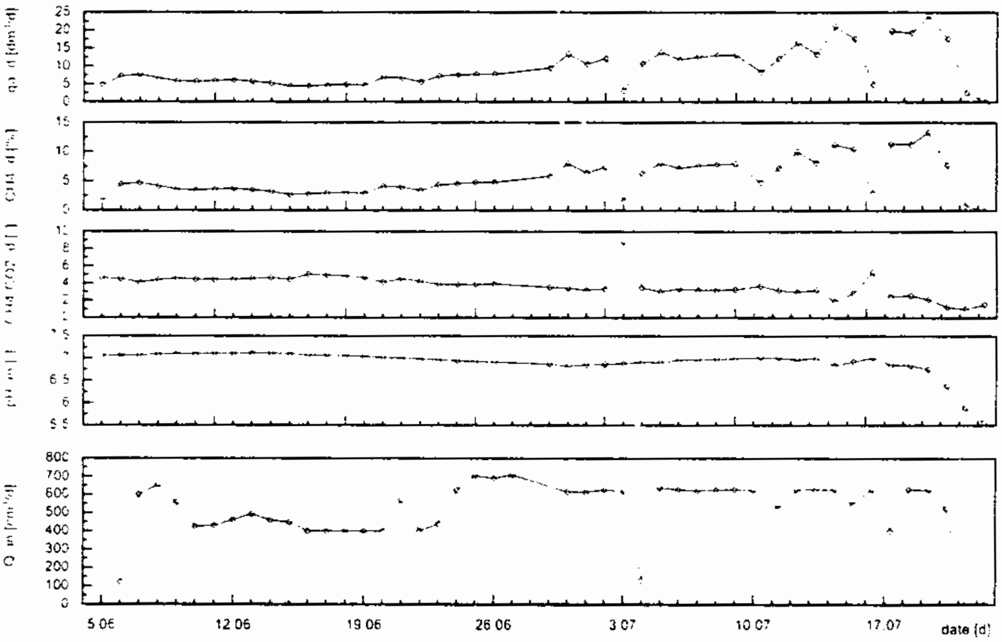


Figure 4: Result of a fuzzy system simulation (progression of features and actuating variables)

The optimal additional dosing rate (Q_{in}) is determined as an output variable and can be directly used for process control as an actuating variable. The control basis (22 rules) and the parameters of the affiliation functions of the linguistic terms were prepared together with process experts (Fuzzy expert system).

The model relates to the substrates of fats and hydrolysis as input materials into the fermentation process. The principal purpose of the fuzzy model is to ensure process safety, i.e. a process failure is prevented. Moreover, the model optimises total gas and CH_4 yields.

The result of a simulation with real process data is shown in Figure 4. This figure depicts the time progression of three out of seven features (= fuzzy system inputs) and the progression of the optimal additional dosing rate (= fuzzy system output variable). After 44 days the retention time was reduced and the organic load increased. This resulted in a risk for the methane-generating process (as recognisable from the severely lowering pH). The system, logically, defines a lower additional dosing rate as output variable in order to re-stabilise the process [13].

SUMMARY

This paper presents a universal software system for monitoring biotechnological processes. The software tool „Model Builder“ combines classical model-building methods and CI methods in a modular graphic user interface. Particular attention was devoted to the classificator model type, which is of great importance in process monitoring, by implementing different classificatory algorithms. Beside model building, also process control algorithms can be developed and put to a data-based testing. The simple model structure and model testing enable the user to very swiftly realise first process models or process controllers. Optimal process models or optimal process controllers can be selected by comparison of simulation results.

The Model Builder is embedded into an overall algorithm which on the basis of process data that are online acquired by plant sensors is capable of developing and testing on the basis of data a process -model and/or a process controller via downstream data mining with data preprocessing, data processing and feature extraction. Current activities focus on a suitable model export to a target hardware system (µC/DSP system or guidance system).

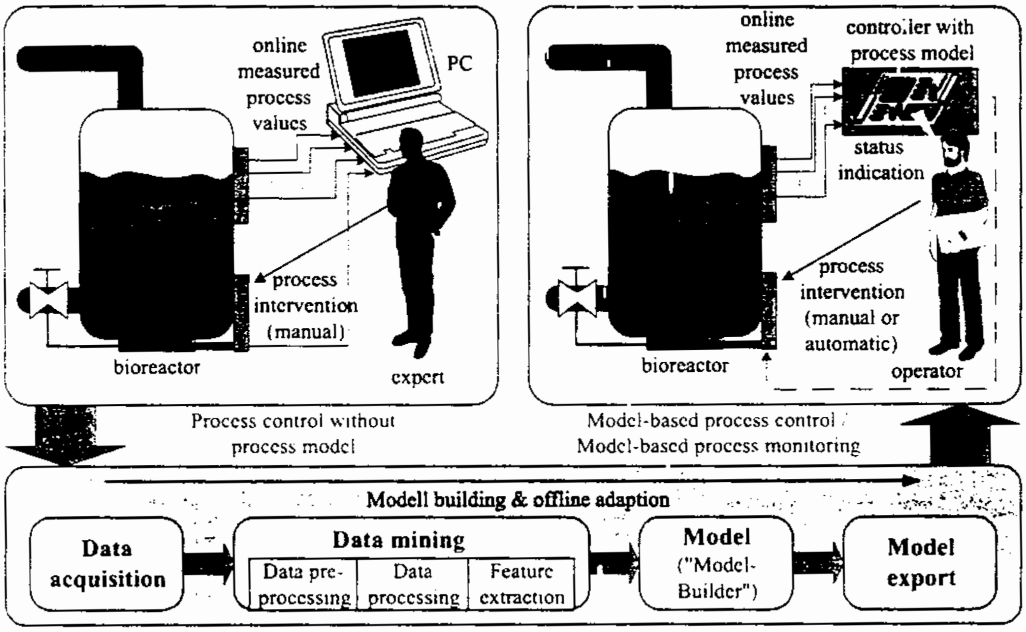


Figure 5: Overall algorithm for data-based development of models / controller structures

ACKNOWLEDGEMENT

Part of this work has been supported by the German Federal Ministry of Education and Research (BMBF).

LITERATURE

[1] **Bezdek, J., Pal, S. K.:** *Fuzzy Models for Pattern Recognition.* New York: IEEE Press, 1992.

[2] **Bocklisch, S. F.:** *Prozessanalyse mit unscharfen Verfahren.* Berlin: Verlag Technik, 1987.

[3] **Canty, M. J.:** *Fernerkennung mit neuronalen Netzen – Theorie – Algorithmen – Programme.* Renningen-Malmsheim: Expert-Verlag, 1999, S. 122-125.

[4] **Günther, M.:** *Kontinuierliche und zeitdiskrete Regelungen. Stuttgart.* B. G. Teubner-Verlag, 1997.

[5] **Estaben, M., Polit, M., Steyer, J. P.:** *Fuzzy Control for an anaerobic digester.* Control Eng. Practice, Vol. 5, No. 98, p.1303-1310, 1997.

[6] **Höppner, F., Klawonn, F., Kruse R.:** *Fuzzy-Clusteranalyse, Verfahren für die Bilderkennung, Klassifizierung und Datenanalyse.* Braunschweig, Wiesbaden: Vieweg-Verlag, 1997.

[7] **Isermann, R.:** *Zur Anwendung der Fuzzy-Logik in der Regelungstechnik.* Automatisierungstechnische Praxis 38, München, Wien: Oldenbourg Verlag, S. 24-26, 1996.

[8] **Koch, M., Kuhn, T., Wernstedt, J.:** *Fuzzy Control – Optimale Nachbildung und Entwurf optimaler Entscheidungen.* München, Wien: Oldenbourg-Verlag, 1996.

[9] **Kohonen, T.:** *Self-organizing and associative memory.* Heidelberg, Berlin, New York: Springer Verlag, 1989.

[10] **Linko, S., Zhu, Y.-H., Linko, P.:** *Applying neural networks as software sensors for enzyme engineering.* TIBTECH, Vol. 17, p. 155-162, 1999.

[11] **Mahalanobis, P. C.:** *On the Generalized Distance on Statistic.* Proc. Nat. Inst. India, No. 2, p. 49-55, 1936.

[12] **Patzwahl, S., Nacke, T., Kramer, K.-D.:** *Selforganizing neural networks to supervise the content of water in biological filtration plants.* 7th European Congress on Intelligent Techniques and Soft Computing, Proceedings of EUFIT '99 Aachen, Germany, Final Programme, Abstracts of the Papers & Proceedings on CD-ROM, Aachen, 1. Edition Aachen: Wissenschaftsverlag, S. 104 ff., 1999.

[13] **Patzwahl, S. et al.:** *Microcontroller-Based Fuzzy System to Optimize the Anaerobic Digestion in Biogas Reactors.* Computational Intelligence. Theory and Applications, International Conference, 7th Fuzzy Days, Dortmund, Germany, October 1-3, 2001. Proceedings, London, Heidelberg, Berlin: Springer-Verlag, p. 2-10, 2001.

[14] **Ripley, B. D.:** *Pattern Recognition and Neural Networks.* Cambridge: Cambridge University Press, 1992.

[15] **Schmitter, E.-D.:** *Neuronale Netze - Einführung, Programmierbeispiele, praktische Anwendungen.* 1. Auflage, Holzkirchen: W. Hofacker Verlag, 1991.

[16] **Souza De, M. B., Ferreira, L. S., Folly, R. O. M.:** *Development of an alcohol fermentation control system based on biosensor measurements interpreted by neural networks.* Sensors and Actuators B 3713, p.1-6, 2001.

[17] **Stenz, R., Kuhn, J.:** *Vergleich: Fuzzy-Automatisierung und konventionelle Automatisierung einer Batch-Destillationskolonne.* Automatisierungstechnische Praxis 35, München, Wien: Oldenbourg Verlag, S. 228-295, 1993.

[18] **Tautz, T.:** *Entwicklung und praktische Realisierung eines mikrocontroller-basierten CI-Regelungssystems für Biogasanlagen.* Diplomarbeit, Hochschule Harz, Wernigerode, 2001.

[19] **Wang, X. Z.:** Data mining and knowledge discovery for process monitoring and control. London, Heidelberg, Berlin: Springer-Verlag, 1999.

[20] **Wernstedt, J.:** *Experimentelle Prozessanalyse.* Berlin: Verlag Technik, 1989.

[21] **Zadeh, L.:** Fuzzy Sets. Information and Control, 8, S.338-353, 1965.

[22] **Zimmermann, H.-J.:** *Datenanalyse – Anwendung von DataEngine mit Fuzzy-Technologien und Neuronalen Netzen.* Düsseldorf: VDI-Verlag, 1995.

Solving clinical problems with artificial neural networks: some case studies

José D. Martín*, Emilio Soria*, Gustavo Camps*,
Antonio J. Serrano*, José R. Sepúlveda* and Víctor Jiménez+
*Digital Signal Processing Group, Electronics Engineering Department.
Faculty of Physics, University of València.
C/ Dr. Moliner, 50, 46100 Burjassot (València). Spain.
Phone: +34-963160198, Fax: +34-963160466
email:{jdmg,soriae,ajserran,gcamps,jsepulve}@uv.es
+Pharmacy Service, University Hospital Dr. Peset, València. Spain.
email: jimenez_vic@gva.es

ABSTRACT: In this paper some medical problems are solved using artificial neural networks. We have identified patients with potential risk of post-Chemotherapy Emesis and potentially intoxicated patients treated with Digoxin. Neural networks have been also used for predicting Cyclosporine A and Erythropoietin blood concentrations. Several neural networks (multilayer perceptron, Elman recurrent neural network, FIR networks, and network ensembles) have been used. Results show that neural networks are very suitable tools for medical classification and prediction tasks, outperforming the mostly used methods in this field (logistic regression and multivariate analysis). Several software applications have been developed in order to improve clinical outcomes and to reduce costs to the Health Care System.

KEYWORDS: artificial neural networks, classification, prediction, software application, Decision Support Systems.

INTRODUCTION

Monitoring patients treated with drugs is an essential task nowadays. Prediction of blood concentrations in order to administer a proper dose and prevention of intoxication continues to be an important aspect in clinical decisions. In fact, the health care team in a hospital decides daily the dose to be administered by assessment of the patients' factors and their clinical conditions. This decision-making process can be aided with mathematical models and usually classical or statistical methods are used. However, reactions that occur in the human body are much more complex than those simulated by theoretic equations, which often do not meet the underlying hypothesis. Despite of this evident fact, a review of the literature shows basically two linear models: logistic regression for classification and multivariate analysis for prediction. These models offer good outcomes and some knowledge gain in the problem since one can see the linear influence of the model inputs on the dependent variable. However, they offer, in general, poor results when relationships between variables are not linear, 'a priori' assumptions about the underlying system are not theoretically and empirically well-founded, when data are sparse, the probability distribution function is unknown, non-uniform sampling in the time series is evident, or highly unbalanced classes are encountered ([1],[2]). On the other side, neural networks are non-linear regression models in which no previous knowledge of the problem is needed and it is not strictly necessary to assume a specific relationship between variables. These features allow the user better versatility and generalization performance than using classical methods.

In this paper we present a four-year project focused on the application of classical and neural approaches for solving pharmaceutical problems. Several software applications have been developed in order to improve clinical outcomes and reduce costs to the Health Care System. The manuscript is organized as follows: in the next section we describe a series of applications in which neural networks has been used to solve some classification and prediction problems; we will end with a brief discussion containing concluding remarks and a proposal for further work.

Intelligent e-Health Applications in Medicine
Special Session Proc EUNITE 2002 Albufeira Portugal Sept 19-21, 2002
Published by the University of the Aegean

SOFTWARE APPLICATIONS

CLASSIFICATIONS OF PATIENTS WITH POTENTIAL RISK OF DIGOXIN INTOXICATION

Digoxin is a drug that does not reduce the mortality rate among patients with congestive heart failure but is considered useful in controlling the symptoms involved in this pathology. Due to its narrow therapeutical range, commonly accepted between 0.8 and 2 ng/ml, more than 10% of the patients treated with it can suffer toxic effects. A patient is usually considered at risk for digoxin poisoning when the blood concentration of the drug is above the therapeutic range, i.e. 2 ng/ml. The data used corresponds to 257 patients monitored in the Pharmacy Unit of the University Hospital Dr. Peset in València, Spain. The best model obtained is a multilayer perceptron (MLP) with sensitivity (Se) and specificity (Sp) of more than 80% in the validation set[1]. It is made up of 14 input variables (including physiological and treatment variables), nine neurons in the hidden layer and one output node that distinguishes patients with risk of been intoxicated (Plasmatic Concentration (PC) ≥ 2ng/mL) [3].

A software for monitoring the clinical evolution of the patients and for predicting and preventing the symptoms of digitalis toxicity was developed and is used currently. This tool was developed in Visual Basic [4] and consists of several forms where the data are collected and introduced to the model. Results are showed in another screen along with the recommendations for the specialist in each case. A module for database management, help files and tips for correct use of the program are also included. In **Fehler! Unbekanntes Schalterargument.** some windows are shown. Great success and confidence was achieved in the preliminary tests and nowadays the tool is working in the Pharmacy service.

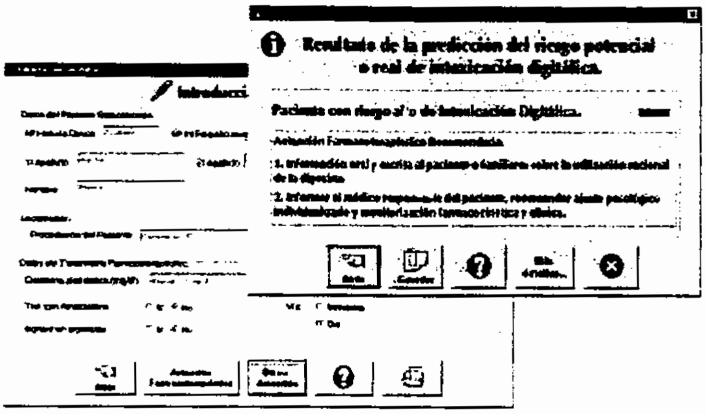


Figure **Fehler! Unbekanntes Schalterargument.**: Two windows of the software PreTox-DGX. Data are entered in a form and, after presenting the inclusion/exclusion criteria, results are presented. User can also retrieve statistics and comparison with logistic regression is also effective.

CLASSIFICATION OF PATIENTS ACCORDING TO THE DEGREE OF POSTCHEMOTHERAPY EMETIC PROTECTION WITH CYSPLATIN

Anti-neoplastic therapy forms part of most cancer treatments, in combination with surgery or radiotherapy. As a result of the toxicity of the drugs administered, oncologic patients have all kinds of negative reactions. Emesis (vomiting and nausea) is, from the patient's point of view, the worst side effect of chemotherapy. Treatment with cysplatin is especially difficult because of this drug's high emetogenic capacity in a great many patients. A questionnaire on vomiting and nausea run on patients treated with cysplatin from April 1996, to March 1998, at the University Hospital Dr. Peset and it was used to develop a neural network for predicting, on the basis of certain characteristics of the patients and the treatment, emesis during the 24h after antineoplastic chemotherapy, which is the highest risk period.

A multilayer perceptron was developed using 212 patterns and then validated using another 107. In the validation set specificity was 73% and sensitivity 85% (79% success rate) [5]. This model was introduced into a web site that also contained a series of elements to help with the model and the problem. In **Fehler! Unbekanntes Schalterargument.** a sample of the web site is shown.

---

[1] Sensitivity is the correct classification percentage on intoxicated patients. Specificity is the correct classification percentage on non-intoxicated patients.

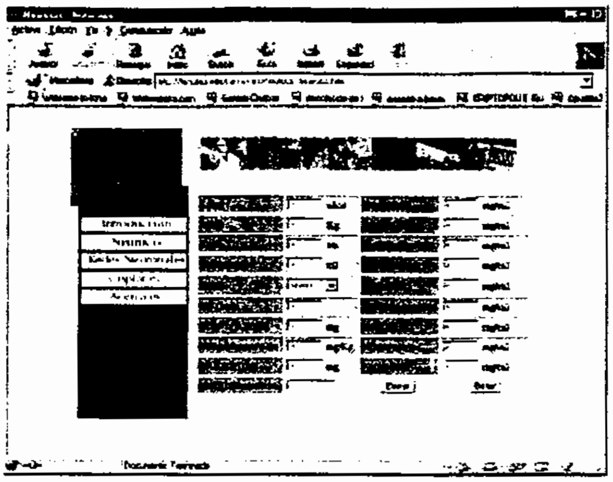


Figure **Fehler! Unbekanntes Schalterargument.**: Web site containing the emesis classification model. The figure shows the window for data introduction.

PREDICTION OF ERYTHROPOIETIN (EPO) DOSE

EPO is administered externally to increase hemoglobin (Hb) or hematocrit (Ht) in anemic patients. In general, the goal is to hold Hb levels at about 12 g/dL and/or Ht at 35% [6] The final goal of our work is to provide a mathematical model for predicting the EPO dose that guarantees the appropriate Hb level in every patient.

Patients used to obtain and validate the model belong to Valencia's Nephroclub Center. The population used to generate the model comprises 77 patients (495 patterns) and another 33 patients (174 patterns) were used to validate the models. In accordance to international directives, results are expressed as a percentage of success rates when the prediction error is below 0.5 g/dL. We used the MLP, the Finite Impulse Response (FIR) network and the Elman recurrent network for predicting the hemoglobin concentration. Results with all of them were very similar (in above the 97 % of the predictions the committed error was less than 0.5 g/dl, which it is a correct prediction according to the experts) [7]. Furthermore, we used a neural networks ensemble formed by these three networks, but since they offered very correlated outputs, the results were not improved.

The best model obtained was implemented in a user-friendly software that consists of a main screen on which the prediction appears after data from the preceding month are filled in, and the possibility of changing the dose of EPO depending on how the dose affects the Hb level. Once the dose has been decided on, it must be stored. In **Fehler! Unbekanntes Schalterargument.** a capture of this tool is shown.

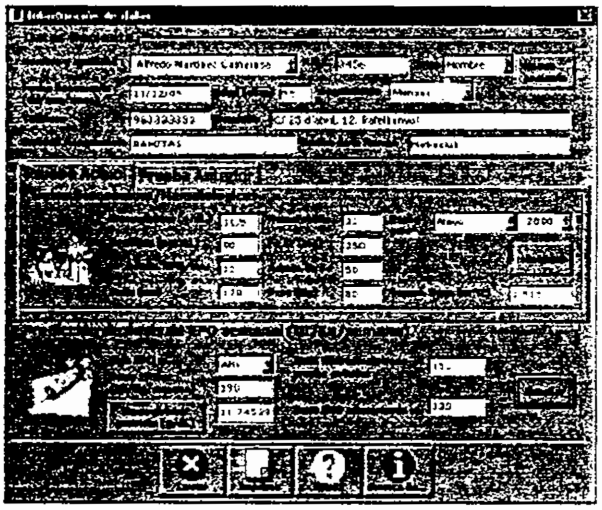


Figure **Fehler! Unbekanntes Schalterargument.**: Main window of the application for predicting EPO dose (fictitious patient and data).

PREDICTION OF CYCLOSPORINE BLOOD CONCENTRATION

Cyclosporine A (CyA) is generally considered a critical dose drug. Its narrow therapeutical range is an important issue in the clinical management of transplant patients, whereas underdosing may result in graft loss and overdosing causes kidney damage, increases opportunistic infections, systolic and diastolic pressure, and cholesterol. Moreover, the pharmacokinetic behavior of CyA presents a substantial inter- and intra-individual variability [8].
Thirty-two renal allograft recipients treated in the Nephrology Service of the University Hospital Dr. Peset were included in this study. Time series for every patient are of different length ranging from 13 to 22 samples. Two thirds of the population was used for training the models and the rest for their validation. Each *pattern* or example was constituted by the present and past values of the variables described.
We have used three neural models for the prediction (Multilayer Perceptron, FIR network and Elman network[2]), and also we have considered combined forecasts. We focused on simple linear combinations such as the equal weights (EW) method, the minimum-variance (MV) method and the Optimal Linear Combination (OLC). All of them are extensively described in [9]. The FIR network shows the best outcomes among the single models, but with neural-network ensembles, outcomes are clearly improved. The best mixture was formed by the OLC [10]. More complex models using different number of each kind of experts have also been attempted but results were similar. A software application implementing the FIR network was developed in order to predict the CyA (see **Fehler! Unbekanntes Schalterargument.**). One can try the parameters of the network, and after the train runs. Results are showed in another window.

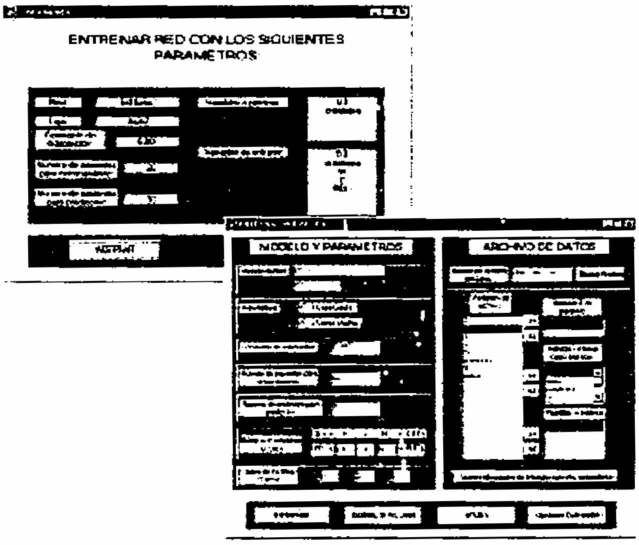


Figure **Fehler! Unbekanntes Schalterargument.**: Windows of the application for predicting the CyA dose.

DISCUSSION

The present report have presented the use of neural networks in medical and pharmacy applications by offering four real applications. Results have always been improved when using these models in comparison to the early stages when linear models were the first, easy though limited approaches attempted.
At present these tools are being tested for clinical-decision aid via software programs and web pages. Initial results indicate that neural networks can be used to reach a better modelization of the body's reactions to medication and, consequently, to improve treatment of patients. We are working at present on more complex models (IIR and Gamma networks, kernel-based methods and neurofuzzy networks) which can provide more sophisticated performance in complex situations. On the other hand, we are researching on rules and knowledge extraction from trained networks. We expect that an understanding of accurate models will provide us with information about which variables are

---

[2] ARMA models were also attempted but results were poor due to the non-uniform sampling, the inherent complexity of the series and their short length.

important and which meaningless, and about treatments and suitable protocols for the posology individualization and the identification and prevention of toxic effects in the patients.

AKNOWLEDGEMENTS

This work has been partially supported by the University of València project "Desarrollo de sistemas neuronales para ser usados como modelos farmacocinéticos" with reference number UV01-15.

REFERENCES

[1] Bishop, C.M., 1996, "Neural networks for pattern recognition". Clarendon-Press, Oxford.

[2] Martin, R.J., 1998, "Autoregression and irregular sampling: Filtering". Signal Processing, no 69, pp. 229-248.

[3] Jiménez, Víctor; Albert, Asunción; Soria, Emilio; Camps, Gustavo; Serrano, Antonio J., 2000, "Prediction of Digoxin Plasma Potentially Toxic Levels by Using a Neural Network Model", ASHP 1999 Midyear Clinical Meeting. Orlando, FL, USA.

[4] Schneider, D.I., 1998, "Essentials of Visual Basic 5.0 Programming". Prentice Hall.

[5] Soria, Emilio; Serrano, Antonio J.; Calpe, Javier; Resurreción Arcas, Iván; Magdalena, Rafael; Jiménez, Victor, 2000, " Determination of the emesis protection level in chemotherapy treatments with Cisplatin using neural networks". Chicago 2000 World Congress on Medical Physics and Biomedical Engineering. Medical Physics, vol. 27, no 6, pp. 1384, June 2000, Chicago, IL, USA.

[6] Balhman, J.; Schoter, K.H.; Scigalla, P; et al, 1984, "Morbidity and mortality in hemodialysis patiens with and without erythropoietin treatment: A controled study". Contrib Nephrol 1984: 76: 250-256.

[7] Martin, José D.; Soria, Emilio; Jiménez, Victor; Camps, Gustavo; Serrano, Antonio J.; Pérez, Juan J., 2000, "Nonlinear prediction of rhEPO by using Neural Networks". Chicago 2000 World Congress on Medical Physics and Biomedical Engineering, Medical Physics, vol. 27, no 6, pp.1378, June 2000, Chicago, IL, USA.

[8] Lindholm, A. 1991, "Factors influencing the pharmacokinetics of cyclosporine in man", Therapeutic Drug Monitoring, vol. 13, no. 6, pp. 465-477, November 1991.

[9] Hasem, S., 1997, "Optimal linear combinations of neural networks", Neural Networks, vol. 10, no. 4, pp. 599-614, 1997.

[10] Camps, Gustavo, Soria, Emilio; Martin, José D.; Serrano, Antonio, J., Pérez, Juan J.; Jiménez, Victor, 2001, "Neural Networks Ensemble for Cyclosporine Concentration Monitoring". International Conference on Artificial Neural Networks ICANN 2001 Lectures Notes in Computer Science 2130, pp. 706-711, August 2001, Vienna, Austria

Adaptive Segmentation and Classification of Non-Stationary and Time-Varying Signals

Lenka Lhotska, Milos Zvolanek
Gerstner Laboratory, Department of Cybernetics,
Czech Technical University in Prague, Faculty of Electrical Engineering,
Technická 2, CZ-166 27 Prague 6, Czech Republic
Phone: +420-2-24353933, Fax: +420-2-24311081
E-mail:lhotska@fel.cvut.cz, milos.zvolanek@worldonline.cz

ABSTRACT: Signal processing specialists are often confronted with the problem of recognizing a sequence of events corresponding to the signature of a particular process state. We tackle the problem of discovering, without the "manual" aid of an expert, implicit relations and temporal constraints from a collection of dated events detected on temporally structured signals. The proposed approach associates tightly signal processing and learning methods. The signal-processing module abstracts the signal data into stationary segments and calculates their basic characteristics. These values are then used as inputs for cluster analysis that finds the best division of all segments into clusters. Resulting clusters are used as training and testing sets for neural network learning. The approach is illustrated on classification of events on EEG records.

KEYWORDS: digital signal processing, cluster analysis, neural network, EEG signal, adaptive segmentation, classification

INTRODUCTION

Signal processing specialists are often confronted with the problem of recognizing a sequence of events corresponding to the signature of a particular process state. This difficulty may be even greater in biomedical domains where behaviour of processes is poorly known and consequently the related scenarios that must be recognized. This is the case in neurology, for instance, where the relation between EEG peaks (transient paroxysmal events associated to the pathology) and epilepsy crises is not well explained. The question is: "if a relation exists between signals and process states, how can it be automatically qualified?"

Another problem that may be successfully solved using automatic processing is analysis of long-term signal recordings. Nowadays it is impossible to check the recordings manually. Usually it is not sufficient to indicate maximum and minimum values but more frequently it is necessary to detect certain shapes in the signal that correspond to significant events. In such cases standard (conventional) temporal and frequency analyses are not satisfactory.

We propose to associate signal processing and machine learning to classify the events. The signal-processing module abstracts the signal data into stationary segments and calculates their basic characteristics. These values are then used as inputs for cluster analysis that finds the best division of all segments into clusters. Resulting clusters are used as training and testing sets for neural network learning.

We illustrate our approach on classification of EEG signal of epileptic patients. The analysis and evaluation of long-term EEG recordings have recently gained increasing importance. One of the problems that are connected with the evaluation of EEG signals is that it necessitates visual checking of such a recording performed by a physician. In case the physician has to check and evaluate long-term EEG recordings computer-aided data analysis might be of great help.

Our case study work deals with the issue of application of artificial intelligence methods to the analysis of EEG signals. The work describes the design and implementation of a system, which performs an automatic analysis of EEG signals.

Several attempts to detect epileptic seizures have already been made. They use formalisms such as conventional temporal and frequency analyses [1], [2], [3], quantitative characterization of underlying non-linear dynamical systems [4], local texture features in conjunction with wavelet transform [5], neural networks [6].

In the next sections we describe used theoretical methods, proposed approach that combines the methods, and give some details about the implemented system. Finally we present experiments and obtained results.

Intelligent e-Health Applications in Medicine
Special Session Proc. EUNITE 2002 Albufeira Portugal Sept. 19-21 2002
Published by the University of the Aegean

MATHEMATICAL METHODS FOR BASIC SIGNAL ANALYSIS

A signal is a description of how one parameter is related to another parameter. Time and frequency are the most common parameters to appear on the horizontal axis of acquired signals; the vertical axis may represent voltage, light intensity, or an infinite number of other parameters. In general, signals can be analysed both in temporal and frequency domains [7]. The basic characteristics in temporal domain are the mean (the average value of a signal) and signal variance (square of standard deviation). Another important parameter is signal-to-noise ratio, which is equal to the mean divided by the standard deviation. In frequency domain, one of the basic transforms used for signal description is Fourier transform. It may be applied both to continuous and discrete signals. Discrete signals are processed by Discrete Fourier transform. Its efficient software implementation is called Fast Fourier Transform (FFT). Its advantage in comparison to analysis using digital filters is that we acquire as results not only amplitude spectrum but also phase spectrum. Then it is possible to calculate further signal characteristics, for example correlation function, power spectra. Correlation contains information on relation of signal value $x(t)$ for time $t = t_0$ and value of this signal $x(t)$ (or another signal $y(t)$) for time $t_1 = t_0 + \tau$. Correlation represents optimal way to detect a known waveform in a signal. Autocorrelation is used for a signal correlated with itself. This function is very useful and frequently used because its Fourier transform is the power spectrum of the original signal. It is used, for example for detection of signal periodicity.

ADAPTIVE SEGMENTATION

In automatic signal analysis, extraction of informative attributes with the greatest possible discriminative ability belongs to important tasks. If we use signal divided to intervals of constant length for acquisition of informative attributes, non-stationariness of the signal may cause distortion of characteristics estimation. Segments defined in this way may contain mixture of waves of different frequencies and shapes. It is preferable to divide signal to segments of different interval length that are stationary. There exist several approaches to adaptive segmentation [8], [9] which divide signals to stationary segments.

ADAPTIVE SEGMENTATION BASED ON AUTOCORRELATION FUNCTION

Let us have two windows, one of them is reference window and is fixed to the beginning of the segment. The second window is sliding (testing) and slides along the signal. Signal characteristics are calculated in the fixed reference window. The same characteristics are calculated in the sliding window. From the differences of signal characteristics in both windows measure of difference (deviation from stationariness) is determined – see eq. (1).

This measure corresponds to difference of signals in both windows. If the measure of difference exceeds defined threshold, the point is marked as segment border. The reference window is fixed to the beginning of a new segment and the procedure is repeated.

Measure of difference is calculated from the difference of autocorrelations of reference and testing windows

$$DIFF = \sum_{n=0}^{ACFL} \frac{\left[R_{W1}(n) - R_{W2}(n)\right]}{R_{W1}(0) R_{W2}(0)} \qquad (1)$$

where R_W is autocorrelation of the window, W_1, W_2 are windows, ACFL is number of coefficients of autocorrelation. The disadvantage of this method is delay of indication of segment border time stamp behind the real border. This delay must be estimated from further signal course.

ADAPTIVE SEGMENTATION BASED ON TWO JOINT WINDOWS

The idea of sliding two joint windows follows the method proposed in [10]. It is based on calculation of differences of two windows. The difference is calculated from spectra of both windows, using FFT. The method is very slow because the difference is calculated using FFT for each window shift. In addition, the method indicates borders even around real segment borders. Therefore in our approach the autocorrelation is used for calculating measure of difference, similarly to [8]. Advantage of this new method is indication of segment borders without delay and possibility of multichannel segmentation.

Following procedure enables to indicate segment borders: Two joint windows slide along the signal. For each window the same signal characteristics are calculated. Measure of difference is determined from the differences of signal characteristics in both windows. This measure corresponds to difference of signals in both windows. If the measure of difference exceeds defined threshold, the point is marked as segment border.

CLUSTER ANALYSIS

Cluster analysis is a widely used technique whose goal is to partition a set of patterns into disjoint and homogeneous clusters [11]. Each pattern is described by features (attributes) and their values. Cluster analysis can be used as classification method as well. It belongs to methods of learning without teacher, which is in some cases necessary, especially if we do not know classified examples in advance. It is often used for a priori classification before using methods requiring training sets with known classification. Cluster analysis enables to adjust classification not only without information about correct classification, but also in extreme case without information about optimal number of classes. It uses for pattern identification similarity measure. It searches for "natural" data structure (if it exists). Patterns belonging to the same cluster have higher value of similarity measure than patterns belonging to different clusters. Various metrics are used for calculation of similarity measures. The most frequently used ones are Euclidean and Mahalanobis metrics. Mahalanobis metrics respects dependency of individual attributes and standardizes the attributes.

DATA STANDARDIZATION

In practical applications, values of different attributes may differ in order. If we use Euclidean metrics for such data, attributes with high values will dominate the cluster analysis. Thus if we consider all attributes having the same importance it is necessary to perform attribute standardization.

Let us have n patterns forming a matrix Z. $z_{i,j} = (z_{i1}, z_{i2}, \ldots, z_{ip})$, $i \in <1, n>$. Each pattern has p attributes. We perform following operations for all patterns:

We calculate the mean $\bar{z}_j$ and standard deviation s_j for $j \in <1,..,p>$

$$\bar{z}_j = \frac{1}{n}\sum_{i=1}^{n} z_{i,j} \tag{2}$$

$$s_j = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(z_{i,j} - \bar{z}_j\right)^2} \tag{3}$$

We transform original attribute values to standardized values according to (4).

$$z'_{i,j} = \frac{z_{i,j} - \bar{z}_j}{s_j} \tag{4}$$

The standardized attributes are dimensionless quantities.

CLUSTERING METHODS

In principal, the clustering methods can be divided into two groups, namely hierarchical and non-hierarchical methods. Let us briefly describe principle of non-hierarchical methods because this approach has been selected for implementation.

The core of non-hierarchical methods is finding optimal breakdown of clusters according to specified criterion of breakdown optimality. If the criterion reaches its extreme, the breakdown is optimal. There are two parts of the problem. The first one solves the task of the breakdown to the most suitable number of clusters that should constitute the optimal breakdown. The second one solves the task of finding optimal breakdown to given number of clusters.

NEURAL NETWORKS

In general, artificial neural network is a structure for parallel data processing that is composed of many simple processing elements. Each of them can accept arbitrary finite number of input data in parallel and send arbitrary finite number of equal information about state of its single output. Most frequently used neural networks are multilayer neural

networks, usually with an input layer, one hidden layer, and an output layer. According to topology, they can be divided into two groups, namely feed-forward networks and feedback networks.
The neural network model used has been a standard three layer feed-forward network, trained with the resilient back propagation algorithm. A sigmoid activation function has been employed throughout. Resilient back propagation algorithm usually converges faster than other algorithms.

DESIGN AND IMPLEMENTATION OF THE SYSTEM

The system for adaptive segmentation and classification of nonstationary and time-varying signals has been designed and implemented as a modular system - see figure 1.

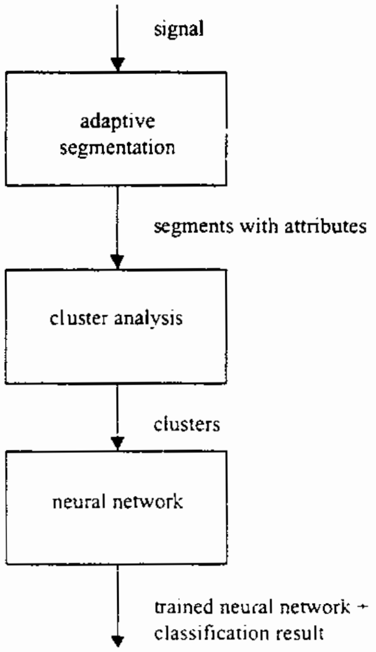


Figure 1. Functional structure of the program

ADAPTIVE SEGMENTATION MODULE

The first module performs adaptive segmentation. Since the input signal may have different course it is possible to set parameters of adaptive segmentation. The most important adjustable parameters are: window type for FFT, window length (= number of samples), number of points for Fourier transform, minimal length of segment, number of samples for filtering, parameters for calculating window difference (e.g. average amplitude, average frequency, variation). When all parameters are set segment borders are identified. If a segment is too short it can be joined to that neighbouring segment, which has more similar value of signal variation. Frequency characteristics of windows corresponding to power spectra are displayed.

Selection and calculation of attributes

After segmentation, attributes that will be used for classification are selected. The following attributes describing signal can be used: average AC amplitude in the segment, variance of AC amplitude in the segment, maximum positive and minimum negative values of amplitude in the segment, maximum value of the first derivation of the signal in the segment, maximum value of the second derivation of the signal in the segment, average value of frequency in the segment, amplitude values in defined frequency bands (e.g. for EEG in alpha, beta, theta, and delta bands). All attribute values are standardized. These values serve as input data for cluster analysis module. All attribute values for all segments and all channels can be exported to a text file as well. It is possible to export values both before and after standardization.

CLUSTER ANALYSIS MODULE

The method used is non-hierarchical clustering of k-means [12], [13]. The user may define number of clusters. Implicit value is 6 clusters. However, the number of clusters can be changed in dependence on character of processed signal. Another possibility is to apply estimation of number of clusters using functional of breakdown quality. Besides number of clusters it is possible to select attributes that will be used for clustering. If a segment is misclassified from the point of view of an expert, its class can be changed manually.

NEURAL NETWORK MODULE

Implemented neural network has one hidden layer at which number of neurons can be changed. Number of input neurons is given by number of attributes set in cluster analysis. Number of output layers is equal to number of classes (clusters). When specifying number of neurons in the hidden layer following facts must be considered. The more neurons are in the hidden layer the worse the ability of the neural network to generalize is. And vice versa – the less neurons are in the hidden layer, the worse the ability of the neural network to learn a sample from training data. Besides number of neurons in the hidden layer, further parameters for training can be changed, namely step of weight incrementation and decrementation, reset of weights, number of learning epochs, maximum MSE error. Trained neural network can be saved to a file and later used for classification.

EXPERIMENTAL DATA AND RESULTS

For experiments, EEG records of epileptic patients have been used. System 10-20 (Montreal convention) has been used for EEG recording, i.e. signals from 21 channels have been acquired. The input file has standard text format. For testing the implemented program we have used two records of one epileptic patient. The signal contains epileptic graphelements of type spike-wave complex of frequency 3 Hz. In both records there are two epileptic seizures. The first signal has been used as training data. The signal has been segmented and segment attributes calculated. Then cluster analysis has partitioned segments into classes. Classified segments have been used for neural network training. The second signal has been segmented as well. These segments have been used for neural network testing.
Table I shows partitioning of segments of the training signal done by cluster analysis.

Class	Number of segments	Description
1	749	Alpha and theta activities
2	185	Artefacts (high waves)
3	1888	Beta and alpha activities
4	88	Epileptic segments
5	29	Epileptic segments

Table I: Classes of segments of the training signal

Two neural networks have been trained. The first one with standardized data, the second one with original data. Classification error on training data has been 3 per cent for the first neural network and 4.5 per cent for the second neural network. Classification of testing data has brought interesting results. Using standardized attributes, some segments belonging to class 1 and having higher amplitude have been misclassified to class 2. Using original attributes has brought improvement of class homogeneity, especially to classes detecting epileptic segments and artefacts.

CONCLUSION

One of the aims of signal classification system development is to ease work (in the described application work of medical doctors). These systems are to help the doctor interpret long-term EEG records correctly and successively to propose the most appropriate treatment. Further applications of these systems can be in education of new doctors. A number of algorithms exist which may be employed to classify unknown EEG signals. Basically, they can be divided into two groups. The first group is based on rules defined by human experts. The rules are biased towards human expertise. The second group utilises various forms of learning, thus avoiding human biasing. Both groups of algorithms have their advantages and drawbacks.

We have verified suitability of application of chain of several methods for signal processing and evaluation, namely adaptive segmentation, cluster analysis, and neural network. Since the system enables to export data processed by each module other methods can be used for their processing as well. We suppose that this system can be used for classification of other signals as well.

Future work can be divided into following main directions:

- utilisation of decision trees for classification;
- application of data mining and knowledge discovery methods in order to find possibly new relations among signal features;
- application of instance-based reasoning (IBR) to the same problem.

ACKNOWLEDGMENTS

The research described in the paper has been supported by the research program No. MSM 210000012 "Transdisciplinary Biomedical Engineering Research" of the CTU Prague, sponsored by the Ministry of Education, Youth and Sports of the Czech Republic and by MIRACLE - Machine Intelligent Research and Application Centre for Learning Excellence, EU project "Centre of Excellence", ICA1-CT-2000-70002.

REFERENCES

[1] Davey B.L.K. et al., 1989, "Expert system approach to detection of epileptiform activity in the EEG", Med. Biol. Eng. Comput. 27(4), pp. 365-370

[2] Gotman J., 1982, "Automatic recognition of epileptic seizure in the EEG", Electroencephalogr. Clin. Neurophysiol. 83, pp. 271-280

[3] Qu H.; Gotman J., 1993, "Improvement in seizure detection performance by automatic adaptation to the EEG of each patient", Electroencephalogr. Clin. Neurophysiol. 86, pp. 79-87

[4] Mayer-Kress G., 1994, "Localized measures for nonstationary time-series of physiological data", Integrative Physiological Behav. Sci. 29 (3), pp. 205-210

[5] Petrosian A.; Homan R.W.; Pemmaraju S.; Mitra S., 1995, "Wavet-based texture analysis of EEG signal for prediction of epileptic seizure", SPIE Proceedings on Wavelet Applications in Signal and Image Processing, Vol. 2569, San Diego, CA, pp. 189-194

[6] Weng W.; Khorasani K., 1996, "An adaptive structure neural networks with application to EEG automatic seizure detection", Neural Networks 9(7), pp. 1223-1240

[7] Smith S.W., 1997, "Digital Signal Processing", California Technical Publishing, San Diego, CA, USA

[8] Bodenstein G.; Praetorius H.M., 1977, "Feature extraction from the electroencephalogram by adaptive segmentation", Proc. IEEE, 65, pp. 642-652

[9] Krajca V.; Petranek S.; Ondruska V.; Doruzkova J.; Pecharova D., 1992, "Multichannel adaptive segmentation – basis for EEG record pre-processing", Lekar a technika, 2, 23, pp. 28-34 (in Czech)

[10] Silin D.Z.A.; Skrzlev K.M., 1997, "Avtomaticeskaja segmentacija EEG" Zurnal vyssej nervovej dejatelnosti, 36, pp. 1152-1155

[11] Jain A.K.; Dubes R.C., 1988, "Algorithms for Clustering Data", Prentice-Hall, Englewoods Cliffs, NJ

[12] Alsabti K.; Ranka S.; Singh V., 1995, "An Efficient K-means Clustering Algorithm", Proceedings of the First Workshop on High Performance Data Mining

[13] Huang Z., 1998, "Extensions to the K-means algorithm for clustering large data sets with categorical values", Data Mining and Knowledge Discovery, 2, pp. 283-304

Author Index

Subject Index