A Multi-Modal Case-Based System for Clinical Diagnosis and Treatment in Stress Management

Mobyen Uddin Ahmed, Shahina Begum, Peter Funk

School Of Innovation, Design and Engineering,
Mälardalen University, PO Box 883 SE-721 23, Västerås, Sweden
{firstname.lastname}@mdh.se

Abstract. A difficult issue in stress management is to use biomedical sensor signal in the diagnosis and treatment of stress. Clinicians often base their diagnosis and decision on manual inspection of signals such as, ECG, heart rate, finger temperature etc. However, the complexity associated with the manual analysis and interpretation of the signals makes it difficult even for experienced clinicians. A computer system, classifying the sensor signals is one valuable property assisting a clinician. This paper presents a case-based system that assist a clinician in diagnosis and treatment of stress. The system uses a finger temperature sensor and the variation in the finger temperature is one of the key features in the system. Several artificial intelligence techniques such as textual information retrieval, rule-based reasoning, and fuzzy logic have been combined together with case-based reasoning to enable more reliable and efficient diagnosis and treatment of stress. The performance has been validated implementing a research prototype and close collaboration with the experts. The experimental results suggest that such a system is valuable both for the less experienced clinicians and for experts where the system may be seen as a second option.

1 Introduction

Medical knowledge is today expanding rapidly to the extent that even experts have difficulties to follow all new results, changes and new treatments. Computers surpass humans in the ability to remember. This property is very valuable for clinician work and computer-aided system enable improvements in both diagnosis and treatment. Different methods have proven to be valuable in different diagnosis and treatment situation. Especially methods and techniques from Artificial Intelligence (AI) such as case-based reasoning, textual case based reasoning and fuzzy logic have drawn much attention and proven to be useful in solving tasks previously difficult to solve with traditional computer-based methods. Recent advances show that by combining more than one AI methods and techniques increase the potential for clinical decision support systems. The multi-faceted and complex nature of the medical domain often leads to designing of multi-modal systems [11] [13].

Diagnosis and treatment of stress is such an example of a complex application domain. It is well known that increased stress level can lead to serious health problems. During stress the sympathetic nervous system of our body activates causing a decrease in peripheral circulation which in turn decreases the skin temperature and reverse effect (i.e. parasympathetic nervous systems activates) occurs during the relaxation. Thus the finger skin temperature responds to stress [17]. Since there are large individual variations when looking at the FT, this is a worthy challenge to find a computational solution to apply it in a computer-based system. Case-based reasoning (CBR) is especially suitable for domains with a weak domain theory, i.e. when the domain is difficult to formalize and is empirical. The advantages of CBR in medical domain have been identified in several research works i.e. in [1, 8, 4, 12, 20]. For some applications the integration of CBR and rule-based reasoning have been explored, e.g. in systems like [6, 21]. Cases comprised textual features or textual cases and introducing ontology into the CBR system, to get the advantages, are also implemented in systems such as, in [14, 19]. The use of fuzzy logic in medical informatics has begun in the early 1970s. In fuzzy CBR, fuzzy sets can be used in similarity measure e.g. [5, 7, and 18].

This paper presents a multi-modal and multipurpose-oriented clinical decision support system for stress management. Our previous work in [10] provides a solution for the diagnosis of stress based only on the finger temperature using CBR and fuzzy similarity matching. The system presented in this paper, for stress management, is not only based on the FT sensor data but also considers contextual information i.e. human perception and feelings in textual format. The system applies CBR as a core technique to facilitate experience reuse. Moreover, an effort has been made in this research work to improve the performance of the stress diagnosis task when there are limited numbers of initial cases introducing a fuzzy rule-based classification scheme. Reliability of the performance for diagnosis and decision making is further enhanced using textual information retrieval with ontology. Finally, a three phase CBR framework is also included into the system to assist in treatment i.e. biofeedback training.

2 CBR system for stress management

The construction of multi-purposed and multi-modal medical systems is also becoming a hot topic in the current applied CBR research. Fig. 1 presents the steps to develop a hybrid multi-purpose CBR system to support in diagnosis and treatment of stress-related disorder.

Step 1: Clinical studies show that FT in general decreases with stress and helps to determine stress-related disorders [17]. Analyzing/interpreting finger temperature and understanding large variations of measurements from diverse patients requires knowledge and experience and, without adequate support, erroneous judgment could be made by a less experienced staff.

Step 2 and 3: The measurement is taken from 31 subjects using a temperature sensor in six steps (i.e. Baseline, Deep-breath, Verbal-stress, Relax with positive thinking, Math-stress and Relax) in the calibration phase [10]. Eight woman and twenty three men within the age range of 24 to 51 are participated in this study. The numbers of individual parameters identified and features extracted from the complex data format are briefly presented in section 2.1.

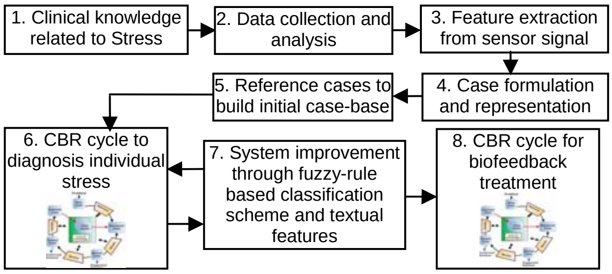


Fig. 1. Schematic diagram of the stress management system in the IPOS

Step 4 and 5: A new problem case is formulated with 19 features in total. The problem description part of a case contains a vector of the features extracted from the FT measurements and the solution part provides a level of stress. The levels of stress are denoted as Very Relaxed, Relaxed, Normal/Stable, Stressed and Very Stressed and the initial case base, with 53 reference cases from 31 subjects, is classified by the domain expert.

Step 6: To diagnose individual stress level, a new FT measurement (formulated as a problem case) is inputted into the CBR cycle. The new problem case is then matched using different matching algorithms including *modified distance function*; *similarity matrix* and *fuzzy similarity matching*. A *modified distance* function uses Euclidean distance to calculate distance between the features of two cases. Hence, all the symbolic features are converted into numeric values before calculating the distance for example, for a feature 'gender' male is converted to one (1) and female is two (2). The function *similarity matrix* is represented as a table where the similarity value between two features is determined by the domain expert. For example, the similarity between same genders is defined as 1 otherwise 0.5. In *fuzzy similarity*, a triangular membership function (mf) replaces a crisp value of the features for new and old cases with a membership grade of 1. In both the cases, the width of the membership function is fuzzified by 50% in each side. Fuzzy intersection is employed between the two fuzzy sets to get a new fuzzy set which represents the overlapping area between them.

$$sim(C_f, S_f) = s_f(m1, m2) = \max(om/m1, om/m2) \quad \textbf{(1)}$$

Similarity between the old case (S_f) and the new case (C_f) is now calculated using equation 1 where *m1*, *m2* and *om* is the area of each fuzzy set. The system can provide matching outcome in a sorted list of best matching cases according to their similarity values in three circumstances: when a new problem case is matched with all the solved cases in a case base (between subject and class), within a class where the class information is provided by the user and also within a subject.

Step 7: A fuzzy rule-based classification scheme [2] and textual features [4] are introduced to provide improved performance in the stress diagnosis task. Detailed information of the system improvement is presented in section 2.2 and 2.3

Step 8: The last step in fig 1 focuses on the CBR system in biofeedback treatment. A three phase CBR framework [3] is deployed to classify a patient, estimate initial parameters and to make recommendations for biofeedback training. A detailed description on the three phases is given in section 2.4.

2.1 Feature mining from the biomedical sensor signal

An experienced clinician often classify FT signal manually without being pointed out intentionally all the features he/she uses in the classification. However, identifying appropriate features is of great importance in developing a computer-based system. To determine important features the system uses 15 minutes measurements (time, temperature) in 1800 samples, together with other numeric (age, room-temperature, hours since meal, etc) and symbolic (gender, food and drink taken, sleep at night, etc) parameters. After analyzing a number of sample measurements it was found that the FT decreases during the *verbal stress condition* and increases in *relax condition*. In our opinion, either mean value or standard deviation of the FT measurement might not be an indicative for stress. For instance, consider two signals one is increasing from 20 to 30, the other decreasing from 30 to 20, and then both have same mean/standard deviation value in the duration, but indicate opposite for stress levels. As an alternative way, we guess that the mean of the slope value might be a feasible feature to convey relation with stress. If the mean slope is sufficiently positive, it will be an indication of relax, otherwise an indication for stress. But if the mean slope is around zero, it shows a situation with high uncertainty for decision or weak decision. According to closer discussion with clinicians, the derivative of each step of FT measurement (from calibration phase) is used to introduce "degree of changes" as an indication of the FT changes. A low angle value, e.g. zero or close to zero indicates no change or stable in finger temperature. A high positive angle value indicates rising FT, while a negative angle, e.g. -20° indicates falling FT.

Total signal, except the baseline, is divided into 12 parts with one minute time interval and 12 features (i.e. *Step2_Part1*, *Step2_Part2*, *Step3_Part1*, …, *Step6_Part1*, *Step6_Part2*) are extracted. The system thereafter formulates a new problem case combining these generated features and other features namely *start temperature, end temperature, minimum temperature*, *maximum temperature* and *difference between ceiling and floor*. Also we consider human defined features such as, sex, hours since last meal etc. This new formulated case is then applied in diagnosing and treatment plan of stress by using the CBR cycle.

2.2 Fuzzy rule-base reasoning for creating artificial cases

The cases stored in the case library should be both representative and comprehensive to cover a wide spectrum of possible situations. The composition of the case library is one of the key factors that decide the ultimate performance of a CBR system. Initially, this CBR system has a limited number of available cases which reduces the performance of the system. Therefore, a fuzzy rule-based classification scheme is introduced into the CBR system to initiate the case library, providing improved performance in the stress diagnosis task.

In fuzzy logic, exact reasoning is treated as a special case of approximate reasoning. Everything in fuzzy logic appears as a matter of some degree i.e. degrees of membership function or degrees of truth.

Table 1. Rules for the fuzzy inference system

Fuzzy rules for classification		
Rule no.	**Antecedent** ============= ***Percentage_Negative_Slope***	**Consequent** ============= ***Stress_Class***
1.	VeryHigh	VeryStress
2.	High	Stress
3.	Medium	Normal/Stable
4.	Low	Relax
5.	VeryLow	VeryRelax

A single-input single-output Mamdani fuzzy model is implemented in which the *percentage of negative slope* features is taken as the input variable and the corresponding *stress class* as the output. The parameters of the IF–THEN rules (known as antecedents or premise in fuzzy modeling) define a fuzzy region of the input space, and the output parameters (known as consequent in fuzzy modeling) specify a corresponding output as shown in table 1.

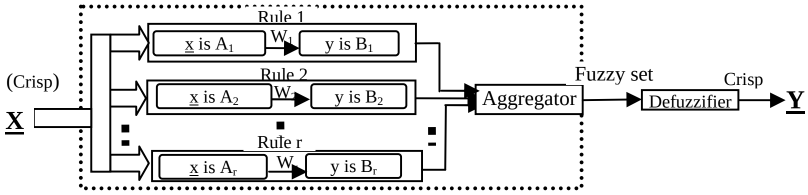


Fig. 2. Block diagram of a fuzzy inference system [9]

The basic structure of fuzzy logic expert systems, commonly known as fuzzy inference system (FIS) shown in Fig. 2, is a rule-based system consisting of three conceptual components: a rule base that consists of a collection of fuzzy IF–THEN rules; a database that defines the membership functions (*mf*) used in the fuzzy rules; and a reasoning mechanism that combines these rules into a mapping routine from the inputs to the outputs of the system, to derive a reasonable conclusion as output. *Percentage_Negative_Slope* and *Stress_Class* are the linguistic variables with the universe of discourse {0, 100} and {1, 5} respectively. VeryHigh, High, Medium, Low and VeryLow are the linguistic values determined by the fuzzy sets "TriangleFuzzySet" on the universe of discourse of *Percentage_Negative_slope*; VeryStress, Stress, Normal/Stable, Relax and VeryRelax are the linguistic values determined by the fuzzy sets "SingletonFuzzySet" on the universe of discourse of *Stress_Class*.

2.3 Textual information retrieval

Clinicians are also considering other factors such as patients feelings, behaviours, social facts, working environments, lifestyle and so on in diagnosing individual stress levels. Such information can be presented by a patient using natural text format and visual analogue scale. Textual data of patients capture important indication not contained in measurements and also provide useful supplementary information. Therefore the system added textual features in the case vector which helps to better interpret and understand the sensor readings and transferring valuable experience between clinicians [3]. For the textual cases, the *tf-idf* (term frequency–inverse document frequency) [15] weighting scheme is used in the vector space model [16] together with cosine similarity to determine the similarity between two cases. Additional domain information often improves results, i.e., a list of words and their synonyms or a dictionary provides comparable words and relationships within the words using class and subclass. It uses domain specific ontology that represents specific knowledge, i.e., relation between words. The different steps in retrieval of similar case(s) in the system are described in Fig. 3.

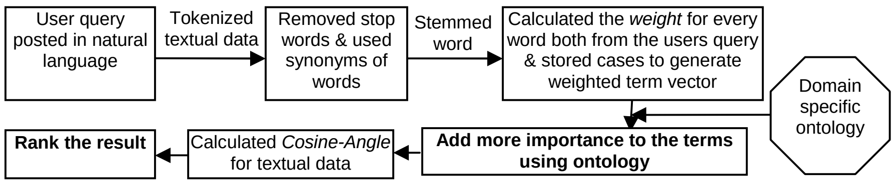


Fig. 3. Different steps for the case retrieval

The text tokenizer algorithm decomposes the whole textual information into sentences, and then into individual words. A filtering step is required to improve retrieval effectiveness due to the huge amount of words. The following three steps are applied to extract the important textual features:

1. Remove the stop-words and special characters blacklist both from the users' query and patients' record.
2. A list of synonyms of the words is used to reduce the number of terms and Porter stemming algorithm helps stemming the words that provide the ways of finding morphological variants of search term. After calculating the weight for each word, these words are represented as terms in a vector space.
3. Improve the importance assessments for candidate terms before measuring the cosine similarity value for the textual information between the stored case and user's query case by using domain specific ontology.

2.4 Biofeedback treatment

The basics of biofeedback is that a patient gets feedback in a clear way (patient observes the graph and knows from preceding education how it should change) and with this feedback can behaviourally train the body and/or mind to biologically respond in a different better way. Biofeedback often focused on relaxation and how the patient can practice relaxation while observing, e.g. the changes in skin temperature. The intention of the system is to enable a patient to train himself/herself without particular supervision. After discussion with clinicians it has been figured out that most of the sensor based biofeedback applications comprised of three phases, 1) analyze and classify patient and make a risk assessment, 2) determine individual levels and parameters, and finally 3) adapt and start the biofeedback training. If the clinician only uses sensor readings shown on a screen then the classification is highly experience based.

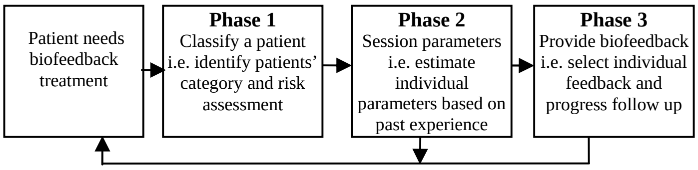


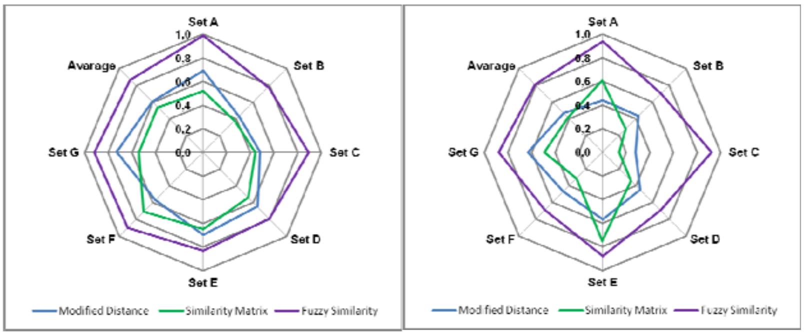
Fig. 4. General architecture of a three-phase biofeedback system

In the first phase as shown in fig 4, a clinician normally asks a number of questions and makes a number of more or less systematic measurements/calculations and then classify a patient depends on the risk and risk-reduction (e.g. stress reactivity and recovery/capacity) of stress. In the second phase, a number of measurements have been done to find out parameters such as, baseline, ceiling, floor temperature etc. needed to tailor the biofeedback session to a patient in order to achieve as good results as possible. Finally, the third phase generates recommendations for a biofeedback training session.

3 Experimental results

The performance of the system has been validated in a prototypical system and close collaboration with experts. Note that, in our previous study, the experiment has been conducted considering 39 cases within the 24 subjects but this paper presents experimental work considering 53 cases from 31 subjects. Moreover, more than one test data sets i.e. for the similarity matching evaluation 7 test sets and for the rest of the evaluation 2 test sets have been considered.

Similarity matching in CBR: Seven subsets of cases and seven query cases, for example: Set A: {7 cases} with query case id 4, Set B: {11 cases} with query case id 16, Set C: {10 cases} with query case id 28, and so on are chosen randomly. All the test sets have been sorted according to the similarity with a query case decided by a domain expert (human reasoning). The sorted cases are then converted to the rank numbers, i.e., the position of a case in the ranking. The top six cases from each set according to the expert's ranking use as standard for the evaluation process where both the similarity values and the ranking numbers are considered. Main goal of this experiment is to investigate the best similarity algorithm for the CBR system compare to the expert's opinion.



a) R^2 value in ranking   b) R^2 value in similarity

Fig. 5. Goodness-of-fit both in ranking and similarity to compare three algorithms

Figure 5 illustrates the comparison between the three algorithms (i.e. modified distance, similarity matrix and fuzzy similarity) for the seven test subsets. The goodness-of-fit (R^2), to evaluate both the ranking numbers and the similarity values compare to an expert, is calculated for all the subsets. The value of R^2 lies between the 0 and 1. A value close to 1 indicates that both the algorithm's and expert's propose a similar decision value. As can be seen from figure 5, both in "ranking number" and "similarity value" criteria, fuzzy similarity algorithm is more reliable than the other algorithms. Fuzzy similarity algorithm performs better in all the seven test subsets.

Fuzzy rule-based classification to CBR: the performance of the CBR system depends on the number of available cases in a case library. So, the goal of this evaluation is to see the improvement of the CBR system adding these artificial cases into the CBR library. Experiment has been done by defining two different case libraries as: LibraryA, with real cases only, classified by the expert and LibraryB being twice as big as LibraryA with hybrid cases, classified by the expert and the fuzzy rule-based classification. As shown in table 2, for the two tests (test1 and test2) on an average the LibraryB indicates the classification accuracy 87% while the LibraryA reaches 74% of fitness compared to expert classification. So, there is 13% increase in the R^2 value and 22% (Mean absolute difference) decrease in the error rate when the system uses the LibraryB (hybrid cases) i.e. case library containing enough cases. For the two tests (using two case libraries) the number of correctly classified cases on average is presented in percentage (see 4th column) in table 2. Here, the CBR system can correctly classify 83% using LibraryB whereas using LibaryA the system can only correctly classify 61% of the cases.

Table 2. Comparison results among the case libraries

Average result for test1 and test2	**Goodness-of-fit (R^2)**	**Mean Absolute Difference**	**Correctly classified cases**
LibraryA	0.74	0.38	61%
LibraryB	0.87	0.16	83%

System performance vs. junior clinicians: for the testing purpose an experienced clinician and two junior clinicians (*JC1* and *JC2*) are involved, 2 subsets of cases (setA and setB) are created randomly with the 11 and 14 expert approved cases. The cases of both subsets are classified by the two junior clinicians who have less experience in this domain. The main goal is to see how good the system can classify compare to the junior clinicians i.e. whether the system can be useful to assist the junior clinician in the classification task.

Table 3. Comparison results between the system and junior clinicians for the two test data sets.

Evaluation Method	**Test setA**			**Test setB**		
	JC1	*JC2*	*The System*	*JC1*	*JC2*	*The System*
Correctly Classified Cases	64%	55%	81%	57%	57%	79%
Goodness-of-fit (R^2)	0.86	0.88	0.92	0.80	0.81	0.83
Absolute Mean Difference	0.36	0.45	0.18	0.43	0.43	0.28

From the table 3 it can be seen that the system using fuzzy similarity matching algorithm can classify correctly better than all the junior clinicians. The test group SetA with 11 cases, the system classifies correctly 81% and the junior clinicians classify correctly 64% and 55% respectively. The number of the correctly classified cases for setB with 14 cases in percentage is 79 by the system whereas the junior clinicians have succeeded to classify correctly as 57 in percentage. The Goodness-of-fit (R^2) value for both the test groups (setA and setB) are 92% and 83% by the system against the senior clinician, the R^2 values are almost the same or little better than the junior clinicians as 86% and 88% for setA, 80% and 81% respectively for setB. The absolute mean difference or error rates in classification for both the test groups are comparatively lower (0.18 and 0.28) than the junior clinicians.

4 Conclusions

Clinical systems have proven to be able to extend the capability of clinicians in their decision making task. But reliability is often a concern in clinical applications. The system presented in this paper supports a clinician in a number of complex tasks in stress management by combining more than one artificial intelligence techniques where CBR is applied as the core technique. Reliability of clinical systems based on sensor readings could certainly be increased by providing contextual information supporting the reasoning tasks. Therefore, the system considers additional information in textual format applying textual information retrieval with ontology. Here, it is also illustrated that it is possible to increased accuracy in the classification task, by extending the case library with artificial cases. A case study also shows that the system provides results close to a human expert. Today the system is based on one physiological parameter i.e. finger temperature sensor signal, in future several other parameters such as heart rate variability, breathing rate etc. could be investigated as a reference of the work for more reliable and efficient decision support in stress management.

References

1. Holt, A., Bichindaritz, I., Schmidt, R., Perner, P.: Medical applications in case-based reasoning, The Knowledge Engineering Review, Vol. 20:3, (2005) 289 – 292

2. Ahmed, M. U., Begum, S, Funk, P., Xiong, N., Schéele, B. V.: A Three Phase Computer Assisted Biofeedback Training System Using Case-Based Reasoning, In 9th European Conference on Case-based Reasoning workshop proceedings. Trier, Germany (2008)
3. Ahmed, M. U., Begum, S, Funk, P., Xiong, N., Schéele, B. V,: Case-based Reasoning for Diagnosis of Stress using Enhanced Cosine and Fuzzy Similarity, Transactions on Case-Based Reasoning on Multimedia Data, vol Volume 1, nr Number 1, IBaI Publishing, (2008a), ISSN: 1864-9734
4. Bichindaritz, I., Marling, C.: Case-based reasoning in the health sciences: What's next? In Artificial Intelligence in Medicine. 36(2), (2006), pp 127-135
5. Bonissone, P., Cheetham, W.,: Fuzzy Case-Based Reasoning for Residential Property Valuation, Handbook on Fuzzy Computing (G 15.1), Oxford University Press. (1998).
6. Bradburn, J., Zeleznikow: The application of case-based reasoning to the tasks of health care planning. In proceedings of topics in case-based reasoning: 1st European workshop, EWCBR-93. Springer, Berlin, (1994), pp 365–378
7. Dvir, G., Langholz, G, Schneider, M: Matching attributes in a fuzzy case based reasoning. Fuzzy Information Processing Society, pp. 33–36. (1999).
8. Gierl, Schmidt, R: CBR in Medicine. In Case-Based Reasoning Technology, From Foundations to Applications. Springer-verlag. (1998), pp. 273 – 298. ISBN:3-540-64572-1
9. Jang, S, R, Sun, C.T, Mizutani, E: Neuro-fuzzy and Soft Computing. A computional approach to learning and machine intelligence. Prentice Hall, NJ. ISBN 0-13261066-3. (1997)
10. Begum, S., Ahmed, M.U., Funk, P., Xiong, N., Schéele, B.V: A case-based decision support system for individual stress diagnosis using fuzzy similarity matching, In Computational Intelligence (CI), in press, (2008). Blackwell.
11. Montani, S: Exploring new roles for case-based reasoning in heterogeneous AI systems for medical decision support. In Applied Intelligence. 2007, pp 275–285
12. Montani, S., Portinale, L., Leonardi, G., Bellazzi, R., Bellazzi, Case-based retrieval to support the treatment of end stage renal failure patients, In Artificial Intelligence in Medicine 37 , (2006), pp 31-42
13. Nilsson, M., Sollenborn: Advancements and Trends in medical case-based reasoning: An overview of systems and system development. In proceedings of the 17th International FLAIRS Conference, (2004), pp. 178-183.
14. O'Sullivan, D., Bertolotto, M., Wilson, D., McLoghlin, E.: Fusing Mobile Case-Based Decision Support with Intelligent Patient Knowledge Management. In Workshop on CBR in the Health Sciences, 151-160. (2006). ECCBR'06
15. Salton, C., Buckley, Term Weighting Approaches in Automatic Text Retrieval, Technical Report. UMI Order Number: TR87-881., Cornell University 1987.
16. Salton, A. Wong, C. S. Yang, A Vector Space Model for Automatic Indexing, Communications of the ACM, vol.18, nr. 11, 1975. pp 613–620.
17. http://www.iworx.com/LabExercises/lockedexercises/LockedSkinTempNL.pdf, last referred April 2009
18. Wang, W. J: New similarity measures on fuzzy sets and on elements. Fuzzy Sets and Systems, 1997, pp. 305–309.
19. Wu, D.D., Weber, R., Abramson, F. D. A Case-Based Framework for Leveraging NutriGenomics Knowledge and Personalized Nutrition Counseling. In Workshop on CBR in the Health Sciences, 2004,71-80. ECCBR'04
20. Perner, P.: Introduction to Case-Based Reasoning for Signals and Images, In: P. Perner (Ed.) Case-Based Reasoning on Signals and Images, Springer Verlag (2007)
21. Marling, C., Shubrook, J., Schwartz F.: Case-Based Decision Support for Patients with Type 1 Diabetes on Insulin Pump Therapy. In Advances in Case-Based Reasoning: 9th European Conference, ECCBR 2008 Proceedings, Springer, Berlin. 2(008).