

A HYBRID EXPERT SYSTEM SUPPORTING DIAGNOSIS OF HEART DISEASES

Abdel-Badeeh M. Salem and Rania A. HodHod

Faculty of computer & Information Sciences

Ain Shams University, Abbassia, Cairo, Egypt

e-mail:absalem@asunet.shams.edu.eg

e-mail:rania_adel@hotmail.com

Abstract: In this paper, we have developed a hybrid expert system prototype used for supporting diagnosis of heart diseases. The system merges uncertainty management techniques and case-based reasoning. The system is able to give appropriate diagnosis for 4 heart diseases namely; mitral stenosis, left-sided heart failure, stable angina pectoris and essential hypertension with the corresponding certainty factor. The system has been implemented in visual prolog for windows and has trained set of 42 cases for Egyptian cardiac patients and has been tested by another 13 different cases. Each case contains 33 significant attributes resulted from the statistical analysis performed to 110 cases.

Keywords: *Expert Systems, Case-Based Reasoning, Uncertainty, Medical Informatics.*

1. INTRODUCTION

There are many problems concerning the building of medical expert systems, such as dealing with uncertainty, fuzzy terms and doctors' heuristic knowledge. To overcome these problems, hybrid expert systems allow for the synergistic combination of more than one inference technique with more strength and less weakness than either one alone [1, 2]. In this paper, we developed a hybrid expert system prototype that merges case-based reasoning

methodology (CBR) [3] and uncertainty management techniques [4] for supporting diagnosis for 4 heart diseases. One of the main advantages of CBR is dealing with situations where solutions are not clear cut and its ability to reason from old experiences or cases. In addition to its ability to overcome the problems addressed in rule based expert systems e.g. knowledge acquisition, performance experience, adaptive solutions and maintaining.

2. KNOWLEDGE ACQUISITION & REPRESENTATION

The knowledge has been collected from expert doctors in heart domain from EL-Maadi Military Egyptian hospital, health insurance Egyptian institute, El-Azhar University doctors and medical reference books [4].

Knowledge was collected for these diseases only: Mitral Stenosis (MS), Left-Sided Heart Failure (LSHF), Stable Angina Pectoris (SAP) and Essential Hypertension (EHT). Knowledge was represented using the concept of frames (where each frame represents one case). There were 110 collected cases each comprises 207 attributes.

The architecture of the developed system consists of two knowledge structures (KS). Case memory KS represents the episodic memory and contains the demographic and clinical data. Similarity rules KS, which are used to support retrieving of cases from the case memory. In addition to a working database contains the case specific data and explanation mechanism.

3. CASE INDEXING & RETRIEVAL

One-way to assign importance values to case features is to have a human expert assign them as the case library is being built. Another way is to do a statistical evaluation of a known corpus of cases to determine which dimension predicts different outcomes and/or solutions best. Both methods are applied to assign importance values to different features. Figure1 illustrates a sample of the assigned values of the different features. The features of the input case are assigned as indices characterizing the case. These indices are used to retrieve several closest match cases from the case memory. The system uses the nearest-neighbor algorithm, which determines how similar two cases are by comparing their features [3].

4. SYSTEM VERIFICATION & RESULTS

System verification includes the following tests [6]: (a) check case retrieval accuracy, (b) check retrieval consistency, (c) check for case duplication and (d) global tests.

(a) Check case retrieval accuracy, which means that if the case-base is queried with one of its cases, it should give the same case with distance measure, equals 100%. Figure2 shows the distance measure against the case identification number, where it is clear that only one case hits the value 1.0 which is the query case

(b) Check retrieval consistency, which means that if exactly the same search has been performed twice, the same source cases should be retrieved with the same accuracy.

(c) Check for case duplication in which a case should exactly match itself, but should not be identical to other cases.

(d) Global tests that are important to verify the overall performance of the system. The steps of this test are as follows: (1) did the system retrieve a useful set of cases? (2) Was the retrieval time acceptable?. Table 1 shows the answer to the above questions, for 5 patients.

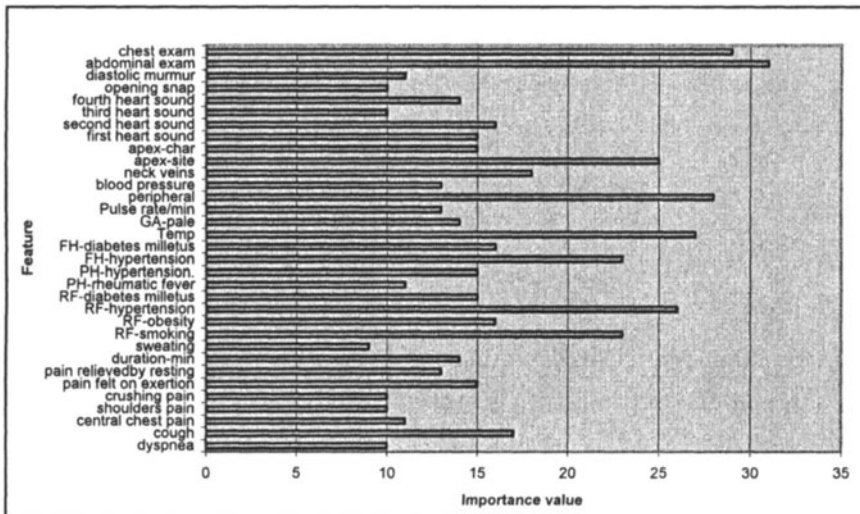


Figure 1: Importance values of different features

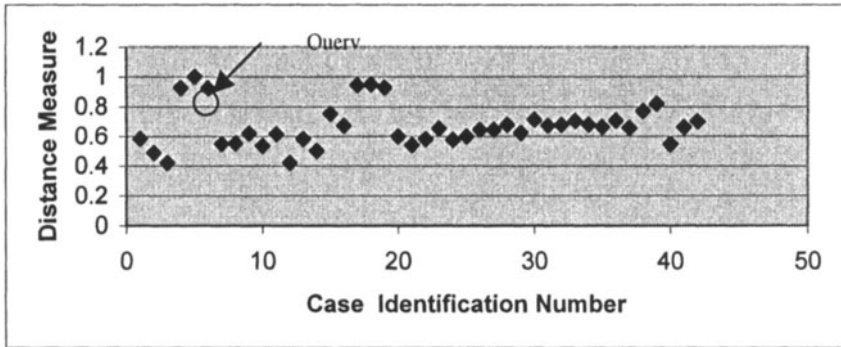


Figure 2: Check of retrieval accuracy

From the Table 1 the answer of the first question is yes. In experiments 1, 2, 3 and 5 the resulted diagnosis of all the closest retrieved cases are the same so this estimates that the query case has the same diagnosis, which is true. But, in experiment 4 there is a confused answer where essential hypertension (EHT) and stable angina pectoris (SAP) are two different diagnoses suggested for the query case. We overlapped this by using K-Nearest Neighbor algorithm, where K represents number of retrieved cases which is always an odd number and then voting for the most probable class (diagnosis) takes place, so there will be 1 vote for stable angine pectoris and 2 votes for essential hypertension tension which is the correct result. The answer of the second question, concerning the acceptable retrieval time, is less than 1 second to get all cases sorted according to their distance measure.

Table 1: Global system verification check

Exp #	Input Case (Query)	Result(R) & Accuracy(A) of the Retrieved Cases					
		Case 1		Case 2		Case 3	
		R	A	R	A	R	A
1	MS	MS	0.984	MS	0.978	MS	0.923
2	SAP	SAP	0.869	SAP	0.868	SAP	0.847
3	SAP	SAP	0.829	SAP	0.821	SAP	0.820
4	EHT	EHT	0.779	SAP	0.764	EHT	0.756
5	LSHF	LSHF	0.758	LSHF	0.740	LSHF	0.694

5. CONCLUSION

In this paper, a hybrid medical expert system that supports diagnosis of four heart diseases is presented. Knowledge structure is represented via

formalism integrating cases and rules. The system was successful in getting correct estimation to thirteen different patients (cases). The system was verified by conducting several tests in order to measure retrieval accuracy, retrieval consistency, case duplication and the overall system performance. System verification shows that both the system and the methodology are successful to estimate a correct diagnosis for heart disease patient. Cardiologists evaluate the system performance by testing it practically where the system succeeded in estimating the correct diagnosis for 13 foreign cases. For future work, more cases will be added to the case memory and it will be clinically tested.

5. REFERENCES

- [1] D.E. Brown, J.J. Pomykalski, Truszkowski, "Expert Systems", Wiley Encyclopedia for Electrical and Electronics Engineering, 1999.
- [2] I. Hatzilgeroudis, J. Prentzas, "HYMES: A Hybrid Modular Expert System with Efficient Inference and Explanation", in Proceedings of the 8th Panhellenic Conference on Informatics, Nicosia, Cyprus, 2001.
- [3] J. L. Kolodner (Ed.), Case-Based Reasoning, Morgan Kaufmann Publishers: California, 1993.
- [4] J. Macleod (Ed.), Davidson's Principles and Practice of Medicine, The English language book society and Churchill Livingstone, thirteenth edition, 1981.
- [5] D. A. Waterman (Ed.), A Guide to Expert Systems, Reading, MA: Addison-Wisley, 1986.
- [6] I. Watson (Ed.), Applying case-based reasoning: Techniques for Enterprise Systems. Morgan kauffmann Publishers: California, 1997.