

Maintaining Case-Based Expert Systems Using Fuzzy Neural Network

Simon C.K. Shiu, Eric C.C. Tsang, Daniel S. Yeung
Department of Computing, Hong Kong Polytechnic University
Email: {csckshiu | csetsang | csdaniel}@comp.polyu.edu.hk
This project is supported by a Hong Kong Polytechnic University grant PA25

ABSTRACT

There is always a trade off between the number of cases to be stored in the case library of a Case-Based Expert System and the performance of retrieval efficiency. The larger the case library, the more the problem space covered, however, it would also downgrade the system performance if the number of cases grows to an unacceptable high level. In this paper, an approach of maintaining the size of a Case-Based Expert System is proposed. The main idea is using the fuzzy class membership value of each record, determined by a trained Neural Network, to guide the record deletion. These fuzzy membership values are used to calculate the case density of each record, and a deletion policy can then be used to determine the percentage of record to be deleted. Using this approach, we could maintain the size of the Case-base without losing significant amount of information. A testing Case-base consists of 214 records is used as an illustrative example of our approach, the Neural Network software NEURALWORKS PROFESSIONAL II/PLUS[®] is used to develop the Neural Network. It was shown that it could reduce the size of the case library by 28% if we select those records that have an overall class membership of over 0.8 and case density over 0.95. Future work includes integrating adaptation rules for building deletion policy.

1. INTRODUCTION

Expert System is one of the branches of Artificial Intelligence that has successfully moved from laboratories to real life applications. Among the various Expert System paradigms, Case-based Reasoning (CBR) is a relatively recent technique that is attracting increasing attention. According to [12,17,26], the main reasons of the CBR success are:

- a) It is closer to actual human decision processes. (i.e. solving the problem by comparing similar situations in the past).
- b) Automation of the process of incorporating new knowledge in the knowledge base.
- c) Better explanation and justification by showing previous examples.
- d) CBR does not require an explicit domain model and so knowledge elicitation becomes a task of gathering case histories.

A CBR system typically consists of four processes [17,26]:

- a) Retrieve the most similar case.
- b) Reuse the case to attempt to solve the problem.
- c) Revise the proposed solution if necessary.
- d) Retain the new solution as a part of a new case.

When the number of cases increases over time, inductive learning methods could be used to induce general rules from the cases. These rules may be in the form of a decision tree, where each leaf carries a class name, and each inner node specifies an attribute with a branch corresponding to each possible value. In analyzing the similarities among cases that involve uncertainty, fuzzy production rules are used to calculate the membership degrees for each case, or to find similar cases in the case library [5,19,30,27]. When similar cases have accumulated to warrant maintenance, anomalies may exist in the case library, such as redundant cases, conflicting cases, ambiguous cases, subsumed cases and unreachable cases [18,22]. Besides, performance problems such as retrieval efficiency degrades will become a real issue if uncontrolled case-base growth is allowed. Techniques that can automatically maintaining the size of the case-base as well as detecting problem cases in the case library are therefore crucial to the future success of CBR technologies.

Currently, the CBR community has largely ignored the issue of maintenance although CBR is becoming a more mature knowledge-based technology. This research proposes using a Neural Network approach for maintaining CBR systems. The main idea is based on the concept of case coverage. We use a Neural Network as a classification tool to divide a case base into various classes, all the records belongs to a particular class will have a fuzzy membership value proportional to the activation value generated by the Neural Network. We then compute all the coverage values of the records, and select the record with the highest coverage as the representative case for that class. Future addition and deletion of cases will depend on the coverage value generated by the Neural Network. By using this approach, uncontrolled case-base growth can be avoided, hence the performance and retrieval efficiency could be maintained. This paper is organized into five main sections. The first

section gives the introduction. Section two reviews some of the current work in applying Neural Network in Case Based Expert Systems. Section three describes our methodology, and section four uses a public domain database, i.e. glass database donated by Diagnostic Products Corporation, to illustrate our approach, the experimental results are also described and analyzed in this section. Finally, section five gives the conclusion and future work.

2. LITERATURE REVIEW

In traditional CBR systems, case retrieval is mainly relied on algorithms such as the nearest neighbor search. It looks for cases stored in memory that consist of the greatest number of characteristics that are the same or similar as the current case. There are many limitations of such an approach:

- a) Difficult to determine which characteristics are more important in retrieval of cases.
- b) Matching of selected feature is an all or nothing affair.
- c) It requires a very large case base to cover the entire problem space.
- d) Features that define a case can be of different types, which must be indexed or represented in different ways.

Recently, Fuzzy Neural Networks are being used for indexing and retrieval of cases, [16] uses fuzzy feature vectors and Neural Network to improve the indexing and retrieval steps in case-based systems. They used a supervised Neural Network to accept inputs of various formats, such as boolean, continuous, multi-valued and fuzzy. They have shown that the use of fuzzy representation for some features enhanced the accuracy of retrieval because the cases retrieved tended to match most closely on the fuzzy attributes. Other works in using fuzzy logic and Neural Networks for case indexing and retrieval could be found in [4,8,11,15,24]

Case-base maintenance is defined by [13] as the process of refining a CBR system's case-base to improve the system's performance, i.e. Case-base maintenance implements policies for revising the organization or contents (representation, domain content, accounting information, or implementation) of the case-base in order to facilitate future reasoning for a particular set of performance objectives. [23] proposed the use of case-based density and the concept of case coverage to determine the maintenance policy, in this paper, we will integrate [23]'s idea with the use of an Neural Network to achieve a better approach to determine the deletion policy.

3. METHODOLOGY

Our case maintenance methodology is divided into four main phases as illustrated in the following diagram 1. In

the first step, case data needed to be converted to computer readable format and ready to be inputted into the Neural Network for training. This includes the determination of feature representation strategy, i.e. which feature is best described by numeric, symbol, boolean or fuzzy attribute. In the case of fuzzy attribute, expert's advice on the determination of fuzzy sets may be necessary. After all the information is in a usable form, we use a typical Neural Network for classification of class membership in next step. We could use the following networks for classification, i.e. Adaptive Resonance Theory Networks (ART), Back-Propagation Neural Network, Probabilistic Neural Network, Fuzzy ARTMAP, Learning Vector Quantization and Self-Organizing Maps (SOM). In step three, we use fuzzy membership to define class for each record, and computed the case density for deletion policy establishment finally. A more detail explanation is given in the following section with an example.

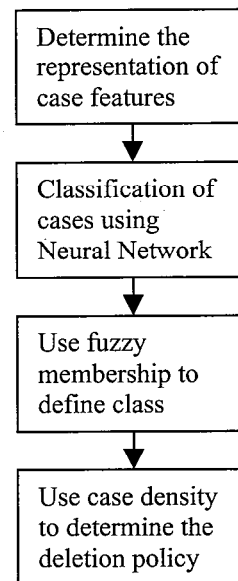


Diagram 1 Case-based Maintenance Steps

4. GLASS IDENTIFICATION DATABASE

In order to illustrate our approach further, it is useful to apply it to a testing database. The system chosen was donated by [10]. The glass identification database, as shown in Table 1, consists of 214 records, and each record has 11 attributes as follows:

1. Id number: 1 to 214
2. RI: refractive index
3. Na: Sodium (unit measurement is weight percent in corresponding oxide, as are attributes 4 to 10)
4. Mg: Magnesium
5. Al: Aluminum
6. Si: Silicon

7. K: Potassium
8. Ca: Calcium
9. Ba: Barium
10. Fe: Iron
11. Type of glass: window glass / non-window glass

Table 1 Sample Cases

RI	Na	Mg	Al	Si	K	Ca	Ba	Fe	Type
1.51215	12.99	3.47	1.12	72.98	0.62	8.35	0	0.31	W
1.51768	12.56	3.52	1.43	73.15	0.57	8.54	0	0	W
1.51652	13.56	3.57	1.47	72.45	0.64	7.96	0	0	W
1.51969	12.64	0	1.65	73.75	0.38	11.53	0	0	N
1.51754	13.39	3.66	1.19	72.79	0.57	8.27	0	0.11	W
1.51911	13.9	3.73	1.18	72.12	0.06	8.89	0	0	W

We used the Neural Network package NERUALWORKS PROFESSIONAL II/PLUS© version 5.3 for testing various types of Neural Network's performance on fuzzy classification of data. We have tried the following network models: Back-Propagation model, Probabilistic Model and SOM model. All of them are suitable for our purpose, and we have finally chosen the Back-Propagation Network to perform detail analysis of the case base. The learning rule algorithm is Delta-Rule, and the transfer function is a sigmoid function. After about 50,000 cycles of training, the network converged and the RMS error was 0.1322, which was considered to be successfully trained.

Furthermore, the confusion matrix graph showed in Figure 1 represents the following results: the x-axis representing the desired output and the y-axis representing the actual output. The confusion matrix breaks the diagram into a grid. If the probe point produce an output of 0.7 and the desired output was 0.5, then the bin around the intersection of 0.7 from the y-axis and 0.5 from the x-axis receives a count. A bar within the bin displays counts, and the bar grows as counts accumulated. The bin that received the most counts is shown at full height, while all of the other bins are scaled in relation to it. The confusion matrix is also equipped with a pair of histograms.

The histogram that runs across the top of the instrument shows the distribution of the desired outputs. The histogram along the right shows the distribution of the actual outputs. Any actual outputs that lay outside the range of the graph will be added to the top or bottom bins along the right (depending on their value). By looking at the two confusion matrix, the desired outputs and actual outputs intercepts quite well, and this also give us satisfactory indicator that the network was trained satisfactory.

We then tested the network by the original set of data, and only accept correct classification if the fuzzy membership value is higher than 0.8, even with such a high membership degree, the overall accuracy of correct

classification is 94%. We expected that it would go up to as high as 99% if more tolerance of fuzziness were allowed. The typical output after training is shown in Table 2 and the network architecture is shown in Diagram 2.

Table 2 Sample output from training

Expected Value 1=Window Glass 0=Non Windows Glass	Fuzzy membership for Window Glass	Fuzzy Membership for Non Window Glass
1	0	0.999998
1	0	0.997225
1	0	0.99802
0	1	0.007992
1	0	0.999834
1	0	0.999047
0	1	0.457469
1	0	0.999737

We select those records having WINDOW-GLASS membership value higher than 0.8 and label them as WINDOW-GLASS class and those records having a NON-WINDOW-GLASS membership value higher than 0.8 and label them as NON-WINDOW-GLASS class. The remaining records are labeled as ODD class, i.e. they neither belong to WINDOW-GLASS nor NON-WINDOW-GLASS to a satisfactory degree. The results are as follows:

Table 3 Classification of the records

Class Name	No. of cases in this class
WINDOW-GLASS	158
NON-WINDOW-GLASS	43
ODD(the remaining)	13

Based on the above result, we can calculated the case density by applying the formula suggested by [23]:

$$\text{CaseDensity}(c, G) = \frac{\sum_{c' \in G - \{c\}} \text{sim}(c, c')}{|G| - 1}$$

$$\text{Sim}(c, c') = 1 - \frac{[\sum_i (c_i - c_i')^2]^{\frac{1}{2}}}{n}$$

|G| = no. of cases in class G
 c = an individual case
 n = total no. of features

For each class, we could use the similarity measurement and case density calculation to compute every case's case density, as shown in table 4.

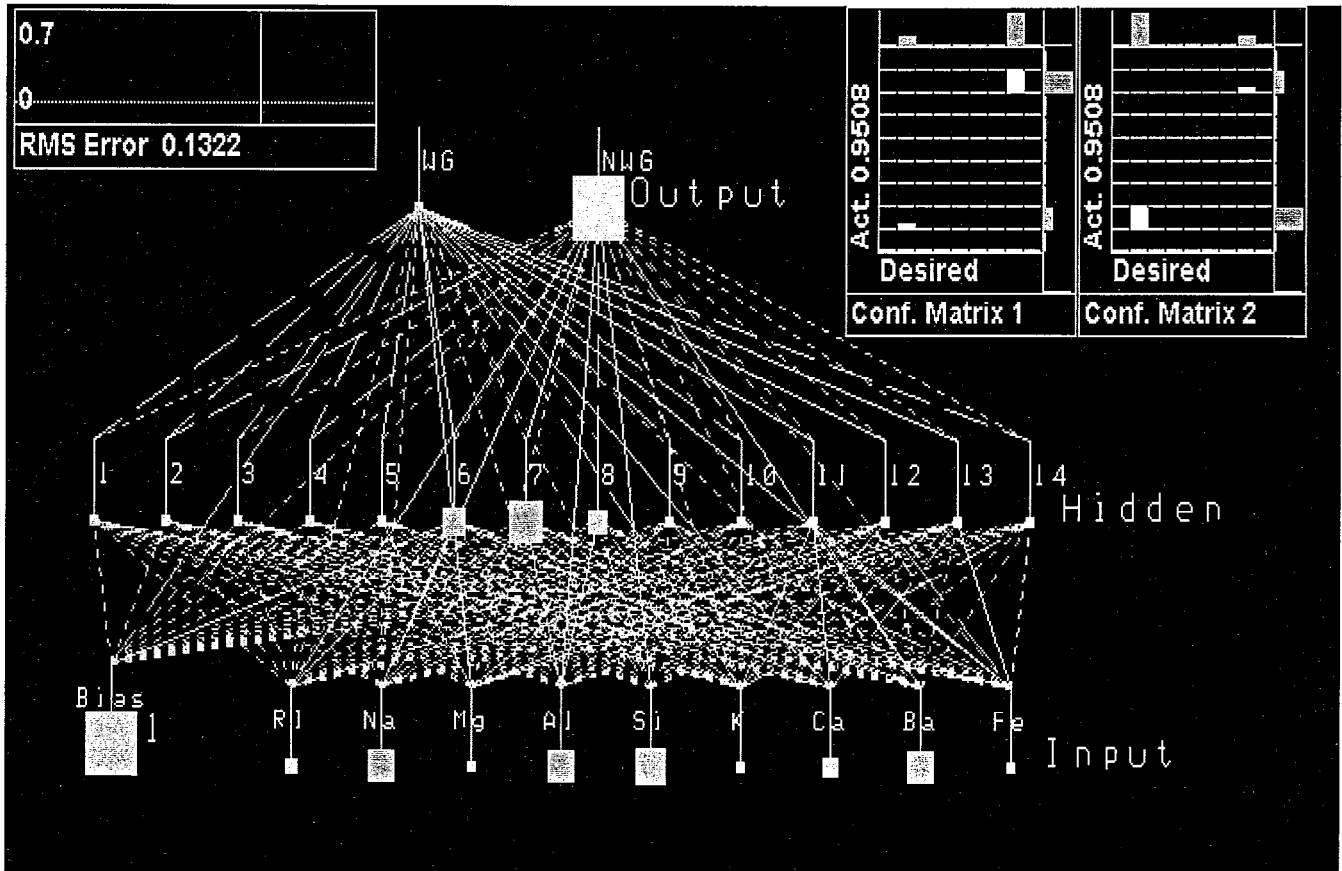


Diagram 2 Neural Network Architecture for classification

Table 4 Sample Result of Case density

Fuzzy membership for Window Glass	Fuzzy Membership for Non Window Glass	Class we defined according to fuzzy membership value	Case density
0.999834	0.000166	Window Glass	0.963236
0.999047	0.000954	Window Glass	0.959489
0.999737	0.000263	Window Glass	0.964335
0.412578	0.587422	Not Sure	0.940094
0.999975	0.000025	Window Glass	0.939792
0.999985	0.000015	Window Glass	0.939759
0.995205	0.004795	Window Glass	0.9396
0.000004	0.999996	Non Window Glass	0.938816
0.999962	0.000038	Window Glass	0.93853
0.000023	0.999977	Non Window Glass	0.938372

From the result, in WINDOW-CLASS class, we select those cases whose case density is greater than 0.95 and has a WINDOW-CLASS membership value greater than 0.95 as the representative cases, and delete the others in this class. Similarly, in NON-WINDOW-GLASS class, we select those cases whose case density is greater than 0.95 and has a NON-WINDOW-GLASS membership value greater than 0.95 as the representative cases, and delete the others in this class. In addition, we retain all the cases in ODD class. We could also use the case density to select some cases as representative cases in the ODD class as well. (In this experiment, we choose to retain all the

cases in the ODD class). The final case maintenance result is as follows:

Table 5 Case Maintenance Result

Class Name	No. of cases originally	No. of cases remain
WINDOW-GLASS	158	112
NON-WINDOW-GLASS	43	29
ODD class	13	13
TOTAL	214	154

Using the above approach, we could delete 60 cases, having a 28% decrease in the size of the case base.

5. CONCLUSION

There is always a trade off between the number of cases to be stored in the Case library of a Case-Based Expert System and the retrieval efficiency encountered. In this paper, we have developed an approach of maintaining the size of a Case-Based Expert System. The main idea is based on the use of the fuzzy class membership value of each record determined by a trained Neural Network. The fuzzy membership values are then used to calculate the Case Density of each record. A deletion policy can then be used to determine to percentage of record to be deleted. A testing Glass/Non-Glass Case-base consists of 214 records is used as an illustrative example of our approach, and NEURALWORKS PROFESSIONAL II/PLUS® is used to

develop the Neural Network. It was shown that we could reduce the size of the Case library by 28% by selecting those records that have an overall class membership of over 0.8 and Case density of over 0.95. Future work includes extending the fuzzy feature selection concepts for identify important case features for ranking. Secondly, unsupervised Neural Networks will be used to determine the classes membership, and finally, integrating of data mining techniques such as discovering of adaption rules for guiding deletion policy will be investigated.

6. REFERENCES

- [1] Aha, D. W. & Wettschereck D., "Case-Based Learning: Beyond Classification of Feature Vectors," in Machine Learning ECML-97, eds. by Someren M. & Widmer, G., pp. 329-336.
- [2] Campbell J.A., & Wolstencroft J. "Structure and Significance of Analogical Reasoning," Artificial Intelligence in Medicine, Vol. 2 (1990) pp. 103-118.
- [3] Chen, M.S., et al., "Data Mining: An Overview from a Database Perspective," IEEE Transaction on Knowledge and Data Engineering, Vol. 8, No. 6, pp.866-883.
- [4] De, Rajat K. and Pal, Sankar K., "Case-based Classification using Fuzziness and Neural Networks," Knowledge Discovery and Data Mining (Digest No. 1998/310), IEE Colloquium, 1998, pp. 6/1-6/3.
- [5] Dubitzky W. et al., "An Advanced Case-knowledge Architecture Based on Fuzzy Objects," Applied Intelligence, Vol. 7, No.3, July, 1997, pp. 187-204.
- [6] Durkin, J., "Expert Systems: Catalogue of Applications," Intelligent Computer Systems, University of Akron, Ohio, USA, 1993.
- [7] Ebert, C., "Fuzzy Classification for Software Criticality Analysis," Expert Systems With Applications, Vol. 11, No. 3, pp. 323-342, 1996.
- [8] Egri, P.A. and Underwood, P.F., "HILDA: Knowledge Extraction From Neural Networks in Legal Rule Based and Case Based Reasoning," in IEEE Int. Conf. on Neural Networks, 1995, Vol.4. pp.1800-1805.
- [9] Gebhardt, F., "Survey on Structure-Based Case Retrieval," The Knowledge Engineering Review, Vol. 12:1, pp.41-58, 1997.
- [10] Glass Identification Database, donated by Diagnostic Products Corporation. UML machine learning repository database.
- [11] Jeng, B.C. & Liang, T.P., "Fuzzy Indexing and Retrieval in Case-Based Systems," Expert Systems With Applications, Vol. 8, No. 1, pp.135-142, 1995.
- [12] Kettler, K., "Case-Based Reasoning: An Introduction," Expert Systems With Applications, Vol. 6, pp. 3-8, 1993.
- [13] Leake, D.B. and Wilson, D.C., "Categorizing Case-Base Maintenance: Dimensions and Directions," in Proceedings of 4th European Workshop, EXCBR-98, pp.196-207.
- [14] Liebowitz, J., "Worldwide Perspectives and Trends in Expert Systems," AI Magazine, Summer 1997, pp.115-119.
- [15] Liu, Zhi Qiang and Yan, Francis, "Fuzzy Neural Network in Case-Based Diagnostic System," IEEE Trans on Fuzzy Systems, Vol.5., No.2., May, 1997.
- [16] Main, J., Dillon T.S. and Khosla R. "Use of Neural Networks for Case-Retrieval in a System For Fashion Shoe Design," in Proceedings of Eight International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, Melbourne, Australia, June 1995, pp.151-158.
- [17] Marir, F. & Watson, I., "Case-Based Reasoning: A Categorised Bibliography," The Knowledge Engineering Review, Vol. 9:4, 1994, pp.355-381.
- [18] O'Leary, D.E., "Verification and Validation of Case-Based Systems," Expert Systems With Applications, Vol. 6, pp.57-66, 1993.
- [19] Petersen, J., "Similarity of fuzzy data in a case-based fuzzy system in anaesthesia," Fuzzy Sets and Systems, Vol. 85 (1997), pp. 247-262.
- [20] Ram A. & Santamaria J.C., "Continuous Case-based Reasoning," Artificial Intelligence, Vol. 90 (1997) pp. 25-77.
- [21] Salzberg S., et al., "Best-Case Results for Nearest-Neighbour Learning," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No. 6., June, 1995, pp. 599-608.
- [22] Shiu, S.C.K. "Formal Description Techniques for the Verification of Expert Systems," PhD Thesis, Department of Computing, HK Polytechnic University, 1997.
- [23] Smyth, B. and McKenna E., "Modelling the Competence of Case-Bases," in Proceedings of 4th European Workshop, EXCBR-98, pp.207-220.
- [24] Vasudevan, C., Smith, S.M. and Ganesan, K., "Fuzzy Logic in Case-Based Reasoning," NASA Joint Technology Workshop on Neural Network and Fuzzy Logic, NAFIPS/IFIS/NASA'94, pp.301-302
- [25] Voss, A., "Case Reusing Systems - Survey, Framework and Guidelines," The Knowledge Engineering Review, Vol. 12:1, pp.59-89, 1997.
- [26] Watson, I. & Marir, F., "Case-Based Reasoning: A Review," The Knowledge Engineering Review, Vol. 9:4, 1994, pp.327-354.
- [27] Yeung, D.S. & Tsang, E.C.C., "A Comparative Study on Similarity-Based Fuzzy Reasoning Methods," IEEE Transactions on SMC, Vol. 27, No. 2, April, 1997, pp. 216-227.
- [28] Yeung, D.S. & Tsang, E.C.C., "A Multi-level Weighted Fuzzy Reasoning Algorithm for Expert Systems," to appear in IEEE Transactions on SMC.
- [29] Yeung, D.S. & Tsang, E.C.C., "Fuzzy Knowledge Representation and Reasoning Using Petri Nets," Expert Systems with Applications, Vol. 7, No.2, pp. 281-289, 1994.
- [30] Yeung, D.S. & Tsang, E.C.C., "Improved Fuzzy Knowledge Representation and Rule Evaluation Using Fuzzy Petri Nets and Degree of Subsethood," International J. of Intelligent Systems, Vol. 9, pp.1083-1100, 1994.
- [31] Yeung, D.S. & Tsang, E.C.C., "Weighted Fuzzy Production Rules," Fuzzy Sets & Systems, 88 (1997) pp. 299-213.