

Fuzzy Relational Learning: A New Approach to Case-Based Reasoning

Ning Xiong

School of Innovation, Design
and Engineering
Mälardalen University
Västerås, Sweden

Liangjun Ma

Shanghai Data Center
China Life Insurance Company
Shanghai, P.R. China

Shouchuan Zhang

State Nuclear Power Automation
System Engineering Company
Shanghai, P.R. China

Abstract — This paper aims to develop a new approach to case-based reasoning without similarity constraint. The key to this is the case relation model which enables identification of relevant cases from a global perspective. Fuzzy linguistic rules are adopted as powerful means to represent knowledge about relevance between cases in the case relation model. The construction of fuzzy relevance rules can be realized by learning from pairs of cases in the case library. The empirical studies have demonstrated that our CBR system using fuzzy relation model can work with an extremely small number of cases while still yielding competent performance.

Keywords - case-based reasoning; fuzzy rules; relational model; relational learning;

I. INTRODUCTION

A major assumption in case-based reasoning (CBR) [1], [2] is that similar experiences can be reused to guide future reasoning, problem solving and learning. This is referred to as the similarity assumption which plays a central role in CBR processes, where similar cases are retrieved from the case base to offer approximate solutions the new problem at hand.

The notion of similarity is closely related to proximity or adjacency. The similarity-based learning methods aim to gather local knowledge from neighboring cases for problem solving. They are supported by the rule of thumb that similar problems have similar solutions. However, this rule is important only in a heuristic sense and will become invalid when truly useful solutions are not residing within cases in the neighborhood. The possibility of similar cases not conveying useful information for problem solving presents a crucial challenge for the CBR research [3], [4].

On the other hand, the similarity assumption appears too restrictive for applying CBR in situations when no instances in the case base are similar to the target problem. So far this issue has not received serious attention by the research community. All CBR papers in literature took it for granted that a set of similar cases can be found from the case base for further processing. Nevertheless, the availability of similar cases is not warranted in reality due to practical limitations, in particular when the cases in the case base are sparsely distributed or not representative of the entire problem space [5].

This paper proposes a new approach to CBR without similarity constraint. The basic idea is to construct a case relation model to predict the relevance between two arbitrary cases. We consider two cases to be relevant if the solution of one case is directly usable for another case. As case relevance is not constrained by adjacency in locations, we utilize global information from varied cases including those remote ones in solving a new problem. Exploiting remote cases would be beneficial to overcome the weakness of conventional CBR approaches in situations when similar experience to new problems is not available.

Further the case relation model is represented as a set of fuzzy linguistic rules. We believe that fuzzy if-then rules present a powerful and flexible means to represent the rich domain knowledge for the relation of pairs of cases. Fuzzy rule-based reasoning can be performed to predict whether and to which extent a case from the library is relevant to the problem in query. The construction of fuzzy relevance rules can be realized by learning from the case library as a valuable resource. Our empirical studies have demonstrated that the learned fuzzy relation model enables the CBR system to work with an extremely small number of cases while still yielding competent performance.

The remaining of the paper is organized as follows: Section II surveys related works. In section III we outline our novel CBR system using (fuzzy) case relation model for prediction of case relevance. The learning of fuzzy rules for the relation model is addressed in section IV. Section V presents experimental results for evaluation of the proposed method. Finally, concluding remarks are given in section VI.

II. RELATED WORKS

Identifying local similar relations between cases is a key task for conventional CBR systems. So far the main stream of the works for construction of similarity models has been focused on feature weighting [6]. Features are assigned with different weights in accordance with their importance, and the global similarity metric is defined as a weighted sum of the local matching values in single attributes. Different approaches of interest have been proposed for identifying such weights automatically. Incremental learning attempts to modify feature weights according to success/failure feedback

of retrieval results [7]. The probability of ranking principle was utilized in [8] for the assignment of weight values to features. Case-ranking information was utilized in [9] for weight adaptation so that the ranking of retrieved cases will be consistent with a desired order. Accuracy improvement represents another way for adapting the set of weights as discussed in [10] and [11]. Nevertheless, no matter how the values of weights are derived, the capability of these similarity learning methods is inherently constrained by weighted combination of the local matching degrees. This limitation in the structure of similarity makes it hard to represent more general knowledge and criteria for assessing case relations from a global perspective.

A new similarity model without feature weighting was proposed in [12] and [13] as an effort to seek more powerful representation of knowledge for case retrieval. The idea was to encode the information about feature importance into local compatibility measures such that feature weighting is no longer needed. Later, in [14], it was analyzed and demonstrated that the parameters of such compatibility measures can be learned from the case library in favor of coherent matching, i.e. to maximize the supportive evidence while minimize the amount of inconsistency derived from pairwise matching of cases from the case base.

The integration of fuzzy theory with CBR methodology has been addressed by some researchers. Yager [15] explained that there was a close connection between fuzzy system modeling and case based reasoning. Dubois and Prade [16] formalized the fundamental hypothesis of CBR in the context of fuzzy rules. They established a formal framework in which case-based inference can be implemented as a special type of fuzzy set-based approximate reasoning.

Fuzzy set and fuzzy logic have also been used for case representation and case matching in a CBR process. Fuzzy sets were used to depict imprecise case features in a fuzzy case-based reasoning system [17]. In [18] the fuzzy subset “small” was defined on attribute differences for numerical evaluation of similarity between cases. Moreover, fuzzy linguistic rules were adopted in [19] and [20] as a flexible means to express the criteria for similarity assessment.

III. CBR USING CASE RELATION MODEL

An overview of the proposed CBR system using case relation model is depicted in Fig. 1, in which traditional similarity evaluation is replaced by global assessment of case relevance that is no longer subject to local neighborhood. Given a query problem Q , we look in the case library for all relevant cases across the whole problem space. The relevance of a known case C with respect to query Q is predicted by the case relation model. Every case in the case library is evaluated according to this model for how it is relevant for solving query problem Q , and those cases with their degrees of relevance larger than a specified threshold are then selected. Hence the set of cases retrieved from the case base is formulated as

$$G = \{C \in CB \mid Relev(C, Q) \geq \alpha\} \quad (1)$$

where CB is the case base and $Relev(C, Q)$ denotes the degree of relevance for case C with respect to query problem Q . Subsequently all the cases in G are delivered to the next block of “decision fusion”, where solutions of the retrieved cases are combined and adapted to make a solution to the new problem.

The purpose of the decision fusion step is to find a new solution to the query problem by modifying and aggregating known solutions of the retrieved cases. Here we consider the relevance values of cases as indicators of the utility or appropriateness of their solutions for solving the new problem. Thus cases with higher relevance degrees will have more influence in determining the final solutions. For instance, for numerical prediction problems, the outcome of the query problem Q is predicted as the weighted average of the outcomes of the retrieved cases as given by

$$Out(Q) = \frac{\sum_{C_i \in G} Relev(C_i, Q) \cdot Out(C_i)}{\sum_{C_i \in G} Relev(C_i, Q)} \quad (2)$$

where $Out(C_i)$ is the output value for case C_i . If the task is to solve classification problems, we need to launch a voting procedure to choose the most plausible class from a set of candidates. The values of relevance of the retrieved cases that have the same outcome can be accumulated into a voting score (VS) for the associated class. In general, the voting score for a candidate class h is calculated by

$$VS(h) = \sum_{C_i \in G} \begin{cases} Relev(C_i, Q), & \text{if } Class(C_i) = h \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Then we select the class with the largest voting score as the predicted class for query problem Q , i.e.,

$$Class(Q) = \underset{h}{\arg \max} [VS(h)] \quad (4)$$

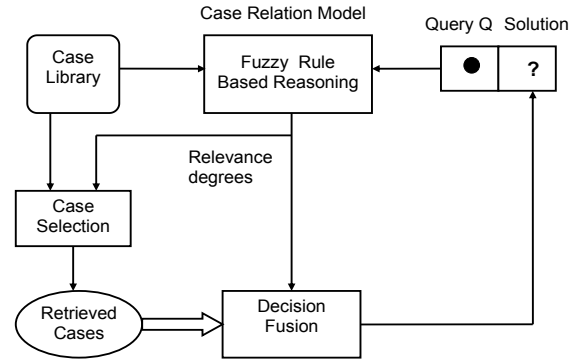


Figure 1. A CBR system using case relation model

Further, the case relation model consists of a set of fuzzy rules and fuzzy-rule based reasoning is performed to determine whether and to which extent a case in the case base is relevant to the new problem at hand. Since fuzzy linguistic rules are more powerful and flexible means to represent the knowledge for assessing global relation between cases, larger

numbers of relevant cases will be identified than case retrieval in terms of local similarity criterion.

Next we consider the structure of fuzzy rules that are employed for the case relation model. Suppose there are n features for problems in the underlying domain. A case C in the case base is indexed by an $(n+1)$ tuple: $C = (x_1, x_2, \dots, x_n, S)$ where x_1, x_2, \dots, x_n denote the attribute values in this case and S is the corresponding solution. Likewise we use an n -tuple (y_1, y_2, \dots, y_n) to represent a query problem Q with y_j referring to the value of the j th attribute in the problem. Let FS_j be the set of fuzzy subsets (linguistic terms) for describing attribute j . A fuzzy rule employed for assessing case relevance can be formulated as follows:

$$\text{If } [(x_1 = A_1) \wedge (y_1 = B_1)] \text{ and } [(x_2 = A_2) \wedge (y_2 = B_2)] \text{ and } \dots \text{ and } [(x_n = A_n) \wedge (y_n = B_n)] \text{ Then Relevance} = V \quad (5)$$

with $A_j \in FS_j$ and $B_j \in FS_j$ for $j=1 \dots n$, and $V \in \{1.0, 0\}$. Note that the conclusion of this rule is a singleton being either unity or zero, it can be regarded as a zero-order Sugeno fuzzy rule.

A special property of the rule in (5) is that it takes into account a pair of attribute values in every sub-condition part. The firing strength of the rule is defined by:

$$t = \min_{j=1 \dots n} (\mu_{A_j}(x_j) \cdot \mu_{B_j}(y_j)) \quad (6)$$

Finally, with a set of fuzzy rules in the form of (5) in the case relation model, the degree of relevance between case C and query problem Q can be calculated as follows:

$$\text{Relev}(C, Q) = \frac{\sum_{\forall k} t_k(C, Q) \cdot V_k}{\sum_{\forall k} t_k(C, Q)} \quad (7)$$

where V_k is the singleton conclusion for rule R_k , and t_k denotes the firing strength of rule R_k given the case C and query problem Q .

III. LEARNING FUZZY RULES FROM THE CASE BASE

With the structure of fuzzy rules having been formulated in the preceding section, we now turn to discussing how to generate such fuzzy rules for assessing relevance between cases. Recall that two cases are relevant if the solution of one case is directly usable for solving another case. Our aim is to elicit relevance values that can precisely approximate utility information. This means that we desire the relation $\text{Relev}(C, Q) = \text{Utility}(S, Q)$ for any query problem Q from the domain and any case C from the case base, with S denoting the solution for case C .

Supervised learning is performed in this paper to generate fuzzy rules for the case relation model. We need a “teacher” to specify samples of desired relevance values for various pairs of cases as training examples. Hiring a domain expert to examine the true relations for many pairs of cases is usually too costly or even unfeasible. The approach we adopt in this paper is to rely on the case base to create adequate training samples for generating fuzzy rules. We will explain how the

training examples can be derived in subsection A , followed by fuzzy rule learning using the derived training samples in subsection B .

A. Deriving Training Examples from Known Cases

Supervised learning is performed in this paper to generate fuzzy rules for the case relation model. We need a “teacher” to specify samples of desired relevance values for various pairs of cases as training examples. Hiring a domain expert to examine the true relations for many pairs of cases is usually too costly or even unfeasible. The approach we adopt in this paper is to rely on the case base to create adequate training samples for generating fuzzy rules. We will explain how the training examples can be derived in subsection A , followed by fuzzy rule learning using the derived training samples in subsection B .

We consider the available case base as a dispensable resource for deriving training samples for learning fuzzy rules. As all cases in the case library have known and optimal solutions, it is a straight forward matter to derive the utility of one case with respect to another by comparing their solutions. Given two cases C_i and C_j , the utility of C_i with respect to C_j can be determined by examining the relation between their solutions, S_i and S_j respectively. The closer solution S_i appears to solution S_j , the more useful solution S_i will be for problem solving in case C_j . In view of this, we define utility between cases as equivalent to the vicinity between their solutions. Thus we can write:

$$\text{Utility}(C_i, C_j) = \text{Vic}(S_i, S_j) \quad (8)$$

The criterion of vicinity between solutions is usually domain dependent, thus we cannot further concretize equation (8) without considering problem context and specifics. Nevertheless, for some common CBR applications such as classification and numerical prediction, the following match functions can be recommended as reasonable metrics in respective circumstances.

1. In classification problems with symbolic classes without orders, the vicinity between two classes can be defined by a binary function as

$$\text{Vic}(S_i, S_j) = \begin{cases} 1 & \text{if } S_i = S_j \\ 0 & \text{if } S_i \neq S_j \end{cases} \quad (9)$$

2. In classification problems with symbolic classes having ordinal values, the vicinity between two classes should reflect the relative distance in the order. Thus we have

$$\text{Vic}(S_i, S_j) = \begin{cases} 1 & \text{if } S_i = S_j \\ 1 - \frac{e(S_i, S_j)}{K} & \text{if } S_i \neq S_j \end{cases} \quad (10)$$

where K is the total number of classes and $e(s_i, s_j)$ denotes the number of classes between S_i and S_j in the order.

3. For problems of numerical prediction, the vicinity between two outputs can be defined by measuring the syntactical differences such that

$$Vic(S_i, S_j) = 1 - \frac{|S_i - S_j|}{\max_k \{S_k\} - \min_k \{S_k\}} \quad (11)$$

The utility derivation stated above enables us to acquire many utility values for pairs of cases from the case library. They are then used as desired relevance values (for the corresponding case pairs) in the training examples for learning fuzzy rules. Since we can yield a degree of utility for every pair of cases in the case library, a much larger multitude of training samples than the number of cases can be created.

B. Generating Fuzzy Rules from Training Examples

We extend the Wang-Mendel algorithm [21] to automatically generate fuzzy relevance rules from the training examples. A training example is a pair of cases along with the desired relevance values between them. As our task is to build fuzzy rules in form of (5), the original Wang-Mendel algorithm is slightly modified to deal with pairs of attribute values in the conditions of the rules. Initially we build candidate fuzzy rules from pairs of cases included in the training data set, one rule from each case pair. Subsequently these candidate rules are evaluated for their truth values in light of their original case pairs. Finally rule reduction is done on each group of candidate rules with the same condition, i.e. only the rule with the highest truth value within the group is kept and retained into the case relation model while other candidate rules are discarded.

More concretely, the learning algorithm for generating fuzzy relevance rules works in the following three steps:

Step 1: Creating fuzzy rules separately from individual case pairs included in the training data set. Given a pair of cases: $C_i = (x_1^i, x_2^i, \dots, x_n^i, S_i)$ and $C_j = (x_1^j, x_2^j, \dots, x_n^j, S_j)$, we produce a candidate fuzzy rule R_{ij} as follows:

$$\text{If } [(x_1 = A_1^{ij}) \wedge (y_1 = B_1^{ij})] \text{ and } \dots \text{ and} \quad (12)$$

$$[(x_n = A_n^{ij}) \wedge (y_n = B_n^{ij})] \text{ Then } Relevance = Vic(S_i, S_j)$$

$$\text{where } A_1^{ij} = \arg \max_{A \in FS_1} \mu_A(x_1^i)$$

$$B_1^{ij} = \arg \max_{B \in FS_1} \mu_B(x_1^j) \quad (13)$$

$$\dots \quad \dots \quad \dots \quad \dots$$

$$A_n^{ij} = \arg \max_{A \in FS_n} \mu_A(x_n^i)$$

$$B_n^{ij} = \arg \max_{B \in FS_n} \mu_B(x_n^j)$$

Step 2: Evaluating the truth value of every candidate fuzzy rule in terms of the case pair from which the rule is created. The truth value of the candidate fuzzy rule as formulated in (12) and (13) is given by:

$$D(R_{ij}) = \min_{k=1 \dots n} (\mu_{A_k^{ij}}(x_k^i) \cdot \mu_{B_k^{ij}}(x_k^j)) \quad (14)$$

Step 3: Clustering all the candidate fuzzy rules into groups such that rules in the same group have the same condition and rules in different groups have distinct conditions. Then, for every rule group, we select the rule with the highest truth value

while discarding the remaining ones to remove conflicting and redundant information in the group.

IV. EXPERIMENTAL EVALUATIONS

In this section we show the evaluation results on two well-known benchmark data sets: IRIS data and Wine data, which can be downloaded from the webpage: <ftp.ics.uci.edu/pub/machine-learning-databases>. The 150 examples in the IRIS data set are characterized by four attributes. The WINE data set consists of 178 samples with 13 attributes. Both data sets have examples classified in three classes without order.

A. Learning from Small Case Bases

In experiments, both the IRIS data set and WINE data set were randomly divided into three parts of equal size: one part was used as case base for learning fuzzy relevance rules and the remaining two parts were used as the test data offering query problems. The fuzzy rules learnt from case bases were employed as case relation model to guide the retrieval of relevant cases for classification of problems in the test data. We did such experiments three times for both data sets (IRIS and WINE), with each time taking a different part of examples as the case base for learning. Tables I and II below indicate the classification accuracy on test problems for the IRIS and WINE data respectively.

TABLE I
TEST ACCURACY ON IRIS DATA USING OUR METHOD

Numbers of trials	Classification accuracy
1	0.940
2	0.960
3	0.960
Average	0.953

TABLE II
TEST ACCURACY ON WINE DATA USING OUR METHOD

Numbers of trials	Classification accuracy
1	0.958
2	0.958
3	0.958
Average	0.958

We can see from the above tables that excellent classification accuracy was achieved by our new CBR system despite the small number of cases (50-60) in the case bases. This can be attributed to the pair-wise examinations of cases in the case base, which produces multiplication of training patterns for fuzzy relational learning of the case relevance model.

Moreover, we compare the performance of our system with some other machine learning approaches in terms of the mean accuracy (on test data) and the number of cases used for learning. Table III illustrates the results for the IRIS data, and the figures for the WINE data are given in Table IV. It is obviously seen that the accuracy obtained by our method is the second best among the results from other papers.

Nevertheless, we used a much lower number of cases for learning than any other works as indicated in the tables.

TABLE III
COMPARISON ON IRIS DATA

The learning methods	Accuracy on test data	Number of cases used for learning
This paper	0.953	50
C4.5 [22]	0.947	135
IGA classifier [23]	0.951	135
Ref. [24]	0.953	135
Ref [25]	0.967	144

TABLE IV
COMPARISON ON WINE DATA

The learning methods	Accuracy on test data	Number of cases used for learning
This paper	0.958	59 ~ 60
C4.5 [22]	0.901	160 ~ 161
IGA classifier [23]	0.937	160 ~ 161
Ref. [24]	0.916	160 ~ 161
Ref [26]	0.944	160 ~ 161
SOP-3 [27]	0.935	160 ~ 161
MOP-3 [27]	0.970	160 ~ 161

B. Robustness against Sharp Case Reduction

An interesting phenomenon observed in experiments is that, after fuzzy rule learning, the original case base can be further reduced to an extremely small size while still yielding good performance of our CBR system. This can be explained by the fact that the learned case relation model conveys information in a global perspective such that cases relevant to query problems can be identified from remote regions given sparsely distributed cases.

TABLE V
ACCURACY ON IRIS DATA USING OUR METHOD UNDER CASE REDUCTION

Numbers of trials	Original case base size	Reduced case base size	Accuracy on test data
1	50	10	0.940
2	50	10	0.940
3	50	10	0.960
Average	50	10	0.947

TABLE VI
ACCURACY ON WINE DATA USING OUR METHOD UNDER CASE REDUCTION

Numbers of trials	Original case base size	Reduced case base size	Accuracy on test data
1	60	6	0.771
2	59	6	0.815
3	59	6	0.975
Average	59.3	6	0.854

Tables V and VI show the classification accuracy in the new experiments, in which the original cases bases were first used to construct the fuzzy relevance rules and then they got a sharp reduction in size for being utilized to classify problems in test data. The results given in these tables demonstrate the robustness of our CBR system against case reduction as long as case relation models are successfully established. A small

case base saves memory space and computational cost and is very beneficial for applications of CBR in real-time systems.

C. Comparison with the KNN Method

We compare our method with the K-nearest neighbor (KNN) method in terms of classification accuracy on the same test problems using the same case bases. Tables VII and VIII illustrate the accuracy obtained by KNN in the same experiments as those described in subsection A. We can see that KNN achieved slightly better accuracy on the IRIS data than our method. This is not surprising since nearest neighbors can sometimes provide very useful local information when they are sufficiently close to a problem in query.

However, as demonstrated in tables IX and X, the performance of KNN degraded more quickly than our method with sharply reduced case bases. This can be explained by the fact that, with a sharp case reduction, the cases in the case base are more likely sparsely distributed such that the nearest cases retrieved contain less meaningful information in a local sense. In contrast, by searching for relevant cases globally, our new CBR system is able to detect and utilize geographically “remote” cases that are relevant. Hence its performance is less affected by the reduction of the size of the case base.

TABLE VII
TEST ACCURACY ON IRIS DATA USING KNN METHOD

Numbers of trials	Accuracy (K=1)	Accuracy (K=3)	Accuracy (K=5)
1	0.940	0.950	0.960
2	0.980	0.980	0.990
3	0.960	0.930	0.980
Average	0.960	0.953	0.977

TABLE VIII
TEST ACCURACY ON WINE DATA USING KNN METHOD

Numbers of trials	Accuracy (K=1)	Accuracy (K=3)	Accuracy (K=5)
1	0.737	0.712	0.729
2	0.681	0.639	0.630
3	0.748	0.681	0.739
Average	0.722	0.677	0.699

TABLE IX
COMPARISON OF OUR METHOD WITH KNN ON IRIS DATA UNDER CASE REDUCTION

Numbers of trials	Reduced case base size	Accuracy of our method	Accuracy of KNN (K=5)
1	10	0.940	0.860
2	10	0.940	0.670
3	10	0.960	0.870
Average	10	0.947	0.800

TABLE X
COMPARISON OF OUR METHOD WITH KNN ON WINE DATA UNDER CASE REDUCTION

Numbers of trials	Reduced case base size	Accuracy of our method	Accuracy of KNN (K=5)
1	6	0.771	0.373
2	6	0.815	0.403
3	6	0.975	0
Average	6	0.854	0.259

V. CONCLUSION

Traditional CBR systems only utilize local similar cases for solving new problems. However, the availability of similar cases is not always guaranteed in practical situations, in particular when the cases in the case base are sparsely distributed. This paper aims to develop a new CBR approach that is not subject to similarity constraint. To this end we need to construct a global case relation model to determine the relevance between arbitrary cases in the whole problem space. Further we suggest that fuzzy linguistic rules being employed as powerful means to represent knowledge in the relation model. Fuzzy relational learning is conducted to discover fuzzy relevance rules from pairs of cases in the case library. Finally, the results of evaluations reveal that, with knowledge support by the case relation model, our CBR system can work competently with a very low number of cases.

As more general significance (beyond CBR), this paper demonstrates the feasibility of building a relational knowledge base from a repository of single objects. To our knowledge, the work presented here represents the first effort of research in the new avenue of fuzzy relational learning. In future we plan to extend the current work by revising the membership functions of fuzzy sets that are used to build fuzzy relevance rules. For that purpose, the extended Wang-Mendel algorithm will be embedded in an overall optimization framework for fitness evaluation of possible specifications of membership functions. The other research direction would be to create flexible structured fuzzy rules for construction of a compact fuzzy relation model with high generalization. We intend to further develop our previous method of premise learning [28], [29] to deal with pairs of cases in a flexible structured fuzzy relevance rule.

ACKNOWLEDGEMENT

The work is partly within the EMOPAC project sponsored by the Swedish Knowledge Foundation. The main author is also grateful to ABB FACTS, Prevas, and VG Power for their co-financing of the project.

REFERENCES

- [1] A. Aamodt, and E. Plaza, "Case-based reasoning: foundational issues, methodological variations, and system approaches," *Artificial Intelligence Com.*, vol. 7, pp. 39-59, 1994.
- [2] R. L. D. Mantaras et al., "Retrieval, reuse, revision and retention in case-based reasoning," *The Knowledge Engineering Review*, vol. 20, pp. 215-240, 2005.
- [3] S. Massie, S. Craw, and N. Wiratunga, "When similar cases don't have similar solutions," In: *Proc. 7th Int. Conf. Case-Based Reasoning*, Belfast, Northern Ireland, 2007, pp.92-106.
- [4] X. Zhou, Z. Shi, and H. Zhao, "Reexamination of CBR hypothesis", In: *Proc. 18th Int. Conf. Case-based Reasoning*, 2010, pp. 332-345.
- [5] M. W. Floyd, and B. Esfandiari, "An active approach to automatic case generation," In: *Proc. International Conference on Case-Based Reasoning*, 2009, pp. 150-164.
- [6] D. Wettschereck, and D. Aha, "Weighting features," In: *Proc. 1st Int. Conf. on Case-based Reasoning*, 1995, pp. 347-358.
- [7] A. Bonzano, P. Cunningham, and B. Smith, "Using introspective learning to improve retrieval in CBR: A case study in air traffic control," In: *Proc. 2nd Int. Conf. Case-based Reasoning*, 1997, pp. 291-302.
- [8] N. Cercone, A. An, and C. Chan, "Rule-induction and case-based reasoning: Hybrid architectures appear advantageous," *IEEE Trans. Knowledge and Data Engineering*, vol. 11, pp. 166-174, 1999.
- [9] L. Coyle, and P. Cunningham, "Improving recommendation ranking by learning personal feature weights," In: *Proc. 7th European Conference on Case-Based Reasoning*, 2004, pp. 560-572.
- [10] J. Jarmulak, S. Craw, and R. Rowe, "Genetic algorithms to optimize CBR retrieval," In: *Proc. European Workshop on Case-Based Reasoning (EWCBR 2000)*, 2000, pp. 136-147.
- [11] H. Ahn, K. Kim, and I. Han, "Global optimization of feature weights and the number of neighbors that combine in a case-based reasoning system," *Expert Systems*, vol. 23, pp. 290-301, 2006.
- [12] N. Xiong, and P. Funk, "Building similarity metrics reflecting utility in case-based reasoning," *Intelligent & Fuzzy Systems*, vol. 17, pp. 407-416, 2006.
- [13] N. Xiong, and P. Funk, "Combined feature selection and similarity modeling in case-based reasoning using hierarchical memetic algorithm," In: *Proc. IEEE World Congress on Computational Intelligence*, 2010, pp. 1537-1542.
- [14] N. Xiong, "Towards coherent matching in case-based classification," *Cybernetics and Systems*, vol. 42, pp. 198-214, 2011.
- [15] R. R. Yager, "Case-based reasoning, fuzzy systems modeling and solution composition," In: *Proc. 2nd Int. Conf. Case-Based Reasoning*, Providence, RI, USA, 1997, pp. 633-643.
- [16] D. Dubois, and H. Prade, "Fuzzy set modeling in case-based reasoning," *International Journal of Intelligent Systems*, vol. 13, pp. 345-373, 1998.
- [17] V. Marques, J. T. Farinha, and A. Brito, "Case-based reasoning and fuzzy logic in fault diagnosis," *WSEAS Transactions on Computers*, vol. 8, pp. 1408-1417, 2009.
- [18] N. Xiong, and P. Funk, "Construction of fuzzy knowledge bases incorporating feature selection," *Soft Computing*, vol. 10, pp. 796- 804, 2006.
- [19] N. Xiong, "Fuzzy rule-based similarity model enables learning from small case bases," *Applied Soft Computing*, vol. 13, pp. 2057-2064, 2013.
- [20] N. Xiong, "Learning fuzzy rules for similarity assessment in case-based reasoning," *Expert Systems and Applications*, vol. 38, pp. 10780-10786, 2011.
- [21] L. X. Wang, and J. M. Mendel, "Generating fuzzy rules by learning from examples," *IEEE Trans. Systems, Man, & Cybernetics*, vol. 22, pp. 1414-1427, 1992.
- [22] J. R. Quinlan, *C4.5: Programs for Machine Learning*, San Mateo, CA, Morgan Kaufman, 1993.
- [23] S. Y. Ho, H. M. Chen, and S. J. Ho, "Design of accurate classifiers with a compact fuzzy-rule base using an evolutionary scatter partition of feature space," *IEEE Trans. Syst., Man, Cybern., Part B*, vol. 34 pp. 1031-1043, 2004.
- [24] Y. C. Hu, "Finding useful fuzzy concepts for pattern classification using genetic algorithm," *Information Sciences*, vol. 175, pp.1-19, 2005.
- [25] N. Xiong, L. Litz, and H. Ransom, "Learning premises of fuzzy rules for knowledge acquisition in classification problems," *Knowledge and Information Systems*, vol. 4, pp. 96-111, 2002.
- [26] T. Elomaa, and J. Rousu, "General and efficient multisplitting of numerical attributes," *Machine Learning*, vol. 36, pp. 201-244, 1999.
- [27] H. Ishibuchi, and Y. Nojima, "Analysis of interpretability-accuracy tradeoff of fuzzy systems by multiobjective fuzzy genetics-based machine learning," *International Journal of Approximate Reasoning*, vol. 44, pp. 4-31, 2007.
- [28] N. Xiong, and L. Litz, "Reduction of fuzzy control rules by means of premise learning --- method and case study," *Fuzzy Sets and Systems*, vol. 132, pp. 217-231, 2002.
- [29] N. Xiong, "Evolutionary learning of rule premises for fuzzy modeling," *International Journal of Systems Science*, vol. 32, pp. 1109-1118, 2001.